Neural and cognitive processes underlying numerosity perception and mental arithmetic

André Knops

Université de Paris

Mémoire de synthèse en vue d'une Habilitation à Diriger des Recherches

Neural and cognitive processes underlying numerosity perception and mental arithmetic

Présenté par André Knops

Chargé de Recherche, Université de Paris, CNRS

2021

date de soutenance: 14-01-2022
composition du jury:
présidante du jury: Karine Doré-Mazars
rapporteurs: Jérôme Prado, Catherine Thevenot
membres: Wim Fias, Anja Ischebeck

CURRICULUM VITAE

André Knops
80 rue Saint Dominique, 75007 Paris, France
Born in Viersen (Germany), 29/10/1973
1 daughter (*2012)
Knops.andre@gmail.com
www.andreknops.com
Languages: German (mother tongue), English, French

RESEARCH INTERESTS

My research is situated at the interface between Experimental Psychology, Cognitive Neuroscience and Developmental Psychology. It is dedicated to the understanding of the cognitive and neural underpinnings of numerical cognition. I combine classical psychophysical methods with the analysis of structural and functional neuroimaging data from healthy participants across the life-span.

HIGHER EDUCATION

2015	**Habilitation & Venia Legendi Psychologie** Humboldt-Universität zu Berlin
2002 – 2006	**PhD (Dr. phil.) Psychology (summa cum laude)** RWTH Aachen Universität
1995 – 2001	**Studies of Psychology (Dipl.-Psych.; 'very good')** Philipps-Universität Marburg

PROFESSIONAL EXPERIENCE

Since 2018	**CNRS Researcher (permanent)** UMR 8240, Université Paris-Descartes / Université de Paris
2016 - 2017	**Stand-In Professor for Experimental Psychology** Department of Psychology, Humboldt-Universität zu Berlin, Germany
2012 - 2017	**Research Group Leader** Department of Psychology, Humboldt-Universität zu Berlin, Germany
2010 - 2012	**Postdoctoral Research Fellow** Department of Neurology, Section Neuropsychology, RWTH Aachen University, Germany
2009 - 2010	**Postdoctoral Research Fellow** Center for Mind/Brain Sciences, University of Trento, Italy
2006 - 2009	**Postdoctoral Research Fellow** INSERM Unit 562, "NeuroSpin" Center, Saclay, France
2005 - 2006	**PhD Candidate** Institute of Medical Psychology and Medical Sociology, RWTH Aachen University, Germany
2002 - 2006	**PhD Candidate** Department of Neurology, Section Neuropsychology, RWTH Aachen University, Germany

PRIZES & AWARDS

2014	*Journal of Vision* rates review as "Exceptionally Good Review"
2014	Nominated for the Award of Good Teaching at Humboldt-Universität zu Berlin
2006	Best interdisciplinary thesis of the Medical Faculty, RWTH Aachen University
2006	Borchers-Medal, RWTH Aachen University for outstanding thesis

SCIENTIFIC SERVICE

AD-HOC REVIEWER FOR JOURNALS (32)

Acta Psychologica; Advances in Cognitive Psychology; Attention, Perception & Psychophysics; Autism Research; Behavioral and Brain Functions; Brain Research; Cerebral Cortex; Cognition; Cortex; European Journal of Developmental Psychology; Experimental Psychology; Frontiers in Psychology; Human Brain Mapping; Journal of Cognitive Neuroscience; Journal of Experimental Child Psychology; Journal of Cognitive Psychology; Journal of Experimental Psychology: Learning, Memory & Cognition; Journal of Neuroscience; Journal of Vision; Language and Cognitive Processes; Nature Human Behaviour; Neuroimage; Neuropsychologia; Neuroreport; Perception; PLoS One; Psychological Research; Psychological Science; Psychonomic Bulletin & Review; Social Cognitive and Affective Neuroscience; Quarterly Journal of Experimental Psychology; Scientific Reports

AD-HOC REVIEWER FOR CONFERENCES AND ORGANIZATIONS (8)

Bologna.lab (Humboldt-Universität zu Berlin)

German Research Council (DFG)

European Research Council (ERC)

Fondation Maison des Sciences de l'Homme (FMSH; Marie Curie Cofund)

Fonds Wetenschappelijk Onderzoek – Vlaanderen (Research Foundation Flanders (FWO))

German-Israeli Foundation for Scientific Research and Development (G.I.F.)

Italian Ministry of Health

Psychologie & Gehirn (26 – 28/05/2016, Berlin)

ORGANIZED SYMPOSIA & SPECIAL ISSUES

2019	Numerosity Perception; ESCoP Tenerife (25 – 28/09/2019)
2017	Guest Editor *Journal of Numerical Cognition:* „Domain-general Factors Influencing Numerical and Arithmetic Processing"
2016	Numerical Cognition, 58th TeaP, Heidelberg (21 – 23/03/2016)
2015	Between Embodiment and Neuronal Recycling – Contribution of the Dorsal Stream to Visual Perception, Frankfurt/Main (7 – 8/06/2015, with Guido Hesselmann)
2014	Numerical Cognition I & II, 57th TeaP, Hildesheim (8 – 11/03/2015)
2010 – 2012	Weekly Colloquium "Brain & Number Group Aachen"
2011	Numerical Cognition and its Dysfunctions, 26th Annual Conference of German Society of Neuropsychology, Aachen (23/11/2011)

MEMBERSHIPS

Mathematical Cognition and Learning Society (MCLS)

German Psychological Society (DGPs)

European Society of Cognitive Psychology (ESCoP)

FUNDING

2018 - 2022	How Domain-general Functions Contribute to the Development of Numerical Competencies (Agence National de la Recherche, ANR; DFCDNC)	433,000 €
2018 - 2020	On the structure of arithmetic facts in memory and its interaction with numerical magnitude representation (main applicant: Daniele Didino)	150,000 €
2015	Symposium ‚Contributions of dorsal stream to visual perception'. DFG (German Research Foundation) HE 6244/3-1 with Guido Hesselmann	10,000 €
2012 - 2016	The impact of the number sense and spatial processing on mental arithmetic. DFG KN 959/2-1 [Emmy Noether Group]	840,000 €
2009 - 2010	Neural Dynamics of mental arithmetic – An fMRI guided TMS-Study. DFG KN 959/1-1	113,000 €
2010	*Attentional shifts in mental arithmetic – nature and relevance. Autonomous Region Trentino (ERC Marie Curie Action Cofund) – declined*	*131,000 €*
2006 - 2007	Emotional Control of Cognition – On the Influence of Mood on Executive Functions in Patients with Biploar Mood Disorder. Nachwuchsprogram der Medizinischen Fakultät der RWTH Aachen, AZ 34/06	36,500 €

TRAINING AND PROFESSIONAL DEVELOPMENT

2013 – 2016	Workshop "Setting up and leading a research unit"
	Workshop „EU funding"
	Workshop "Peer-review and new developments in scientific publishing"
	Annual meeting of Emmy Noether grantees, DFG, Potsdam
2013	"Learning to teach" – didactic and methodological foundations in teaching
	Humboldt-Universität zu Berlin
2013	„Advanced User"
	Berlin Center for Advanced Neuroimaging (BCAN)
2011	BrainVisa/Anatomist – imaging software platform
	Neurospin Center, Saclay, France
2005	BrainVoyager – imaging software
	Brain Innovation, Maastricht, Netherlands

ACADEMIC SELF-ADMINISTRATION

2013 – 2015	Press agent at the Department of Psychology, Humboldt-Universität zu Berlin
2014	Committee member ‚Best Diploma Thesis' at the Department of Psychology, Humboldt-Universität zu Berlin
2013; 2014	Selection of Mentors for Q-Teams at Bologna.lab of the Humboldt-Universität zu Berlin

SOFTWARE AND PROGRAMMING SKILLS

R, SPSS, Matlab, python, PsychtoolBox, BrainVoyager, SPM, MS Office

INVITED PRESENTATIONS

Université Lyon 1, Journee des 'Dys' – Dyscalculie. Lyon, France (12/10/2019)

Department of Psychology, University of York, United Kingdom (24/04/2018)

Brandenburg Medical School, Neuruppin, Germany (07/11/2017)

Laboratory for the Psychology of Child Development and Education (LaPsyDÉ), Université Paris Descartes, Paris, Frankreich (10/10/2016).

Workshop "Heterogeneous Contributions to Numerical Cognition", University of Ghent, Belgium, (01 – 03/06/2016)

Laboratoire Psychologie de la Perception (LPP), Université Descartes, Paris, France (22.02.2016)

Colloquium of the Clinic for Psychiatry und Psychotherapy, Tübingen, Germany (10/02/2016)

Colloquium Experimental Psychology, Albert Ludwigs University Freiburg, Germany (04/11/2015)

Colloquium of the Department of Psychology, Eberhard-Karls University Tübingen, Germany (12/05/2015)

Leibniz Institut für Wissensmedien (IWM/KMC), Tübingen, Germany(19/11/2013)

Colloquium Cognitive Sciences, University of Potsdam, Germany (07/06/2013).

Colloquium "Visual information processing in perception and action", University of Hamburg, Germany (15/11/2012)

Formal DCCN Colloquium, Donders Center for Cognitive Neuroimaging, Nijmegen, Netherlands, (06/12/2011)

Institute for Neuroscience and Medicine, Forschungszentrum Jülich, Germany (29/09/2010)

Symposium "The neural basis of number processing" University of Ghent, Belgium (08/04/2010 – 09/04/2010)

Computational Cognitive Neuroscience Lab, Padua, Italy (28/05/2009)

Center for Mind/Brain Sciences (CIMeC), Rovereto, Italy (09/12/2008)

Stanford Cognitive Sciences and Neurosciences Lab, Stanford, USA (02/09/2008)

Psycholinguistics Laboratory, University of the Basque Country, Vitoria-Gasteiz, Spain (25/04/2008)

INSERM Unité 562, Orsay, France (19.02.2006)

TEACHING

I am teaching on a regular basis, covering a variety of topics in Experimental Psychology from seminars (SE, max 25 participants) for master students on attention and decision making to lectures (VL, max 120 participants) on history and theoretical foundations of Experimental Psychology for Bachelor students in Psychology.

2017	Motivational Psychology	VL
2017	Empirical-experimental Internship (2x)	SE
2016/2017	Psychological Basis of Neuroscience	VL
2016/2017	History and theoretical foundations of Experimental Psychology	VL
2016/2017	Psychological Models of arithmetic and Dyscalculia	SE

2015/2016	Neurocognitive Basis of numerical cognition (28.10.2015)	VL
2015	The role of numerical competencies in decision making	SE
2014/2015	Attention and working memory from a neurocognitive perspective	SE
2014	(Non)conscious processing from neurocognitive perspective	SE
2013/2014	Neurocognitive aspects of arithmetic and dyscalculia	SE
2013	Functional imaging in numerical cognition	SE
2012/2013	Numerical cognition (06.02.2013)	VL
2008/2009	Cognitive und neural Foundations of numerical competencies and reading, University of Trento (23.03.2009), Numbers and Space	VL
2006	Communication skills for dentists, RWTH Aachen	SE
2006	Clinics and psychology of mental disorders, RWTH Aachen	SE
2005	Clinics and psychology of mental disorders, RWTH Aachen	SE

Paris, 11.01.2021

PUBLICATIONS

Peer-reviewed journals [* = open access]

1. * Pinho, A.L., Amadon, A., Gauthier, B., Clairis, N., **Knops, A.** et al. (2020). Individual Brain Charting dataset extensio n, second release of high-resolution fMRI data for cognitive mapping. *Scientific Data* 7, 353
2. * Rączy, K., Czarnecka, M., Paplińska, M., Hesselmann, G., **Knops, A.** & Szwed, M. (2020). Tactile to visual number priming in the left Intraparietal Cortex of sighted Braille readers. *Scientific Reports*, 10:17571, 1-11.
3. Glaser, M. & **Knops, A.** (2020). When adding is right – temporal order judgments reveal spatial attention shifts during two-digit mental arithmetic. *Quarterly Journal of Experimental Psychology*, 73(7), 1115–1132.
4. * Rączy, K., Czarnecka, M., Zaremba, D., Paplińska, M., Hesselmann, G., **Knops, A.** & Szwed, M. (2020). A shared code for Braille and Arabic digits revealed by cross-modal priming in sighted Braille readers. *Acta Psychologica*, 202.
5. *** Knops. A.** (2019). Author accepted manuscript: Understanding mathematical thinking–the key for improving mathematical performance and a test case for exploring the human mind, The *Quarterly Journal of Experimental Psychology*, https://doi.org/10.1177/1747021819882893
6. Breil, C. Didino, D. & **Knops, A.** (2019). The influence of semantic processing and response latency on the SNARC effect. *Acta Psychologica, 196*: 75-86.
7. * Didino, D., Pinheiro-Chagas, P., Wood, G. & **Knops, A.** (2019) Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect. *Front. Psychol. 10*:160. doi: 10.3389/fpsyg.2019.00160
8. Cavdaroglu, S. & **Knops, A.** (2018). Evidence for a posterior parietal cortex contribution to spatial but not temporal numerosity perception, *Cerebral Cortex, 29(7)*, 2965–2977, https://doi.org/10.1093/cercor/bhy163
9. * Pinheiro-Chagas, P., Didino, D. Haase, V. G., Wood, G. & **Knops, A.** (2018). The developmental trajectory of the operational momentum effect, *Front. Psychol. 9*:1062. doi: 10.3389/fpsyg.2018.01062
10. *** Knops, A.**, Nuerk, H.-C., & Göbel, S.M. (2017). Domain-General Factors Influencing Numerical and Arithmetic Processing. *Journal of Numerical Cognition, 3(2),* 112 - 132.
11. * Katz, C. Hoesterey, H. & **Knops, A.** (2017). A role for attentional reorienting during approximate multiplication and division. *Journal of Numerical Cognition, 3(2),* 246 - 269.
12. **Knops, A.** (2017). Probing the neural correlates of number processing. *The Neuroscientist*, *23*(3), 264 – 274.
13. Katz, C. & **Knops, A.** (2016). Decreased cerebellar-cerebral connectivity contributes to complex task performance. *Journal of Neurophysiology, 116,* 1434–1448.
14. Cavdaroglu, S. & **Knops, A.** (2016). Mental subtraction and multiplication recruit both phonological and visuospatial resources: evidence from a symmetric dual-task design. *Psychological Research*, *80(4)*: 608-624.
15. Klein, E., Suchan, J., Moeller, K., Karnath, H.-O., **Knops, A.**, Wood, G., Nuerk, H.-C., & Willmes, K. (2016). Considering structural connectivity in the triple code model of numerical cognition – Differential connectivity for magnitude processing and arithmetic facts. *Brain Structure & Function, 221(2): 979-95.*

16. Cavdaroglu, S., Katz, C. & **Knops, A.** (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception, *Neuroimage*, *116*: 135-148.
17. Didino, D., **Knops, A.**, Vespignani, F., & Kornpetpanee, S. (2015). Asymmetric activation spreading in the multiplication associative network due to asymmetric overlap between numerosities semantic representations? *Cognition, 141*, 1-8.
18. Hesselmann, G., Darcy, N., Sterzer, P. & **Knops, A.** (2015). Exploring the boundary conditions of unconscious numerical priming effects with continuous flash suppression. *Consciousness and Cognition, 31*, 60 – 72.
19. * Fischer, M.H. & **Knops, A.** (2014). Attentional Cueing in Numerical Cognition – A Commentary on Zanolie & Pecher (2014). *Frontiers in Psychology, 5*:1381.
20. * Pinheiro-Chargas, P., Wood, G., **Knops, A.**, Krinzinger, H., Lonnemann, J., Startling-Alves, I., Willmes, K., & Haase, V. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One. 9(11)*: e111155.
21. Hesselmann, G. & **Knops, A.** (2014). No conclusive evidence for numerical priming under interocular suppression. *Psychological Science. 25(11)*:2116-2119. doi:10.1177/0956797614548876
22. *** Katz, C.** & **Knops, A.** (2014). Operational momentum in approximate multiplication and division? *PLoS One; 9(8):e104777. doi: 10.1371/journal.pone.0104777.*
23. **Knops, A.,** Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short term memory and enumeration. *The Journal of Neuroscience*, *34(30)*: 9857-9866.
24. **Knops, A.** , Dehaene, S., Bertelletti, I., Zorzi, M. (2014). Can Approximate Mental Calculation Account for Operational Momentum in Addition and Subtraction? *The Quarterly Journal of Experimental Psychology, 67(8)*: 1541-1556.
25. **Knops, A.** & Willmes, K. (2014). Numerical ordering and symbolic arithmetic share frontal and parietal circuits in the right hemisphere, *Neuroimage, 84:* 786-795.
26. *** Knops, A**., Zitzmann, S. & McCrink, K. (2013). Examining the presence and determinants of operational momentum in childhood. *Frontiers in Psychology*, 4:325.
27. * Koten Jr J.W., Lonnemann, J., Willmes, K., **Knops, A.** (2011). Micro and macro pattern analyses of FMRI data support both early and late interaction of numerical and spatial information. *Frontiers in Human Neuroscience, 5:115*.
28. Boecker, M., Drueke, B., Vorhold, V., **Knops, A.**, Philippen, B., Gauggel, S. (2011). When response inhibition is followed by response reengagement: An event-related fMRI study. *Human Brain Mapping, 32(1)*: 94 -106.
29. **Knops, A.,** Thirion, B., Hubbard, E.M., Michel, V. & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science, 324(5934)*:1583-5.
30. **Knops, A.,** Viarouge A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and non-symbolic calculation: Evidence from the operational momentum effect. *Attention, Perception & Psychophysics, 71(4)*, 803-821.
31. Klein, E., Nuerk, H.-C., Wood, G., **Knops, A.** & Willmes, K. (2009). The exact vs. approximate distinction in numerical cognition is not exact, but only approximate: How different processes work together in multi-digit addition, *Brain and Cognition, 69*, 369 – 381.
32. Huckauf, A., **Knops, A.**, Nuerk, H.-C., & Willmes, K. (2008). Semantic processing of crowded stimuli? *Psychological Research, 72(6)*: 648 – 656.
33. Lonnemann, J., Krinzinger, H., **Knops, A.**, & Willmes, K. (2008). Spatial representations of numbers in children and their connection with calculation abilities. *Cortex, 44(4)*, 420-428.
34. **Knops A.**, Nuerk H.-C., Sparing R., Foltys H., Willmes K. (2006). On the functional role of the parietal cortex in number processing: How gender mediates the impact of a 'virtual lesion' induced by rTMS. *Neuropsychologia, 44(12)*, 2270 – 2283.

35. **Knops A.**, Nuerk H.-C., Fimm B., Vohn R., Willmes K. (2006). A Special Role for Numbers in Working Memory? – An fMRI Study. *Neuroimage, 29*, 1 – 14.

Books

Knops, A. (2020). Bien débuter en CE1 – Compréhension des nombres. Paris, Éditions Nathan.

Knops, A. (2019). Numerical Cognition – The Basics. Routledge.

Book chapters

1. Aite, A., Alterelli, I., Borst, G. & **Knops, A.** (in press). The learning brain: a new perspective for education. In: O'Donnell, A. (ed.). *The Oxford Handbook of Educational Psychology*, Oxford University Press, Oxford, UK.
2. **Knops, A.** (in press). Neuroimaging. In: Zeigler-Hill, V. & Shackelford, T. (eds). Encyclopedia of Personality and Individual Differences. Springer.
3. **Knops, A.** (2018). Neurocognitive evidence for spatial contributions to numerical cognition. In: Henik, A. & Fias, W. (Eds.): *Heterogeneity of Function in Numerical Cognition*, Academic Press.
4. Gauggel, S., **Knops, A.**, & Staedtgen, M. (2007). Transkranielle Magnetstimulation (TMS). In S. Gauggel & M. Herrmann (Eds.), *Handbuch der Neuro- und Biopsychologie* (S. 220 – 228), Göttingen, Hogrefe Publishers.

Theses

1. **Knops, A.** (2015). Neurocognitive evidence for cultural recycling of cortical maps in numerical cognition. Habilitation, Humboldt-Universität zu Berlin.
2. **Knops, A.** (2006). On the structure and neural correlates of the numerical magnitude representation and its influence in the assessment of verbal working memory. Doctoral Thesis, RWTH Aachen University.
3. **Knops, A.** (2001). Der Einfluss der Reizkomplexität auf den Prozess der mentalen Rotation und seine elektrophysiologischen Korrelate. Diploma Thesis, Philipps University Marburg

Conference Presentations (peer reviewed)

1. **Knops, A**. (2019). The Triple Code Model of number processing – a critical evaluation. 7th Meeting of the Federation of the European Societies of Neuropsychology (FESN), Milan, IT, 15.
2. Glaser, M. & Knops, A. (2019). Complex addition orients attention: evidence from temporal order judgments. Abstracts of the Conference European Society for Cognitive Psychology, Tenerife, 332.
3. Czarnecka, M., Rączy, K., Szewczyk, J., Paplińska, M.,Hesselmann, G., **Knops, A.**, & Szwed, M. (2019). High MVPA decoding accuracy for abstract tactile braille numbers in the intraparietal sulcus (IPS) of sighted Braille readers. Abstracts of the Conference European Society for Cognitive Psychology, Tenerife, 264.
4. Didino, D. & **Knops, A.** (2019). Decade-consistency effect in multiplication: the influence of the place-value structure in the retrieval process. Abstracts of the Conference European Society for Cognitive Psychology, Tenerife, 234.
5. Rączy, K., Czarnecka, M., Zaremba, D., Paplińska, M., Hesselmann, G., **Knops, A.**, & Szwed, M. (2019). Tactile to visual number priming in sighted Braille readers: behavioral and fMRI studies. Abstracts of the Conference European Society for Cognitive Psychology, Tenerife, 223.
6. Rączy, K., Czarnecka,M., Szewczyk, J., Zaremba, D., Paplińska, M., Hesselmann, G., **Knops, A.**, & Szwed, M. (2019). Touching numbers. Investigating the neural representation of numbers in dual visual/Braille alphabet readers, Abstracts of the Conference European Society for Cognitive Psychology, Tenerife, 78.
7. Glaser, M. & **Knops, A.** (2019). Spatial biases induced by mental arithmetic and the impact of task difficulty. Conference of The Mathematical Cognition and Learning Society, Ottawa, CA.
8. Didino, D., Breil, C. & **Knops, A.** (2019). The role of semantic processing and response latency in the SNARC effect. Conference of Experimental Psychology (TeaP, 2019), London, GB.
9. Glaser, M. & Knops, A. (2019). Spatial biases induced by mental arithmetic and the impact of task difficulty. Conference of Experimental Psychology (TeaP, 2019), London, GB.
10. Glaser, M. & Knops, A. (2018). Visuospatial attention shifts during mental arithmetic can be measured via a temporal order judgment task In: Abstracts of the 60th Conference of Experimental Psychology, Lengerich, Pabst.
11. Didino, D., & **Knops**, **A**. (2017). The numerical magnitude system influences the retrieval of multiplication facts. Program of the 20th Conference of the European Society for Cognitive Psychology, p. 61.
12. Vagharchakian, L., Izard, V., van den Bos, W., **& Knops, A.** (2016). On the impact of age and format on proportional reasoning. TeaP 2016 – Abstracts (p. 356), Lengerich: Pabst.
13. Cavdaroglu, S. & **Knops, A.** (2015). An fMRI investigation on the role of parietal and sensory cortices in time-variant enumeration processes. In Bermeitinger, C., Mojzisch, A., Greve, W. (Eds.) Teap 2015 - Abstracts. (p. 56), Lengerich: Pabst.
14. Katz, C. & **Knops, A.** (2015). Cerebellar-parietal BOLD activity versus functional connectivity in simple and complex mental arithmetic. In Bermeitinger, C., Mojzisch, A., Greve, W. (Eds.) In Bermeitinger, C., Mojzisch, A., Greve, W. (Eds.) Teap 2015 - Abstracts. (p. 132), Lengerich: Pabst.
15. Cavdaroglu, S. & **Knops., A.** (2014). Limitations of Abstract Number Perception: an fMRI study on the temporal integration of auditory and visual number information. Program No. 624.07/DD18. Neuroscience Abstracts. Washington, DC: Society for Neuroscience, 2014. Online.
16. Katz, C. & **Knops, A.** (2014). Changes in cerebellar-parietal connectivity contribute to differential anterior parietal activity for complex and simple mental calculation. Program No. 624.06/DD17. Neuroscience Abstracts. Washington, DC: Society for Neuroscience, 2014. Online.
17. **Knops, A.,** Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture underlies capacity limits in both working memory and enumeration. In O.

Güntürkün (Eds.): *49. Kongress der Deutschen Gesellschaft für Psychologie: Supplement to Psychological Test and Assessment Modeling* (p. 570). Lengerich: Pabst.

18. Hesselmann, G., Sterzer, P. & **Knops, A.** (2014) The scope and limits of numerical priming under continuous flash suppression. The European Conference on Visual Perception, August 24-28, Belgrad, Serbia
19. **Knops, A.**, Piazza, M., Sengupta, R., Eger, E. & Melcher, D. (2014). A shared, flexible neural map architecture underlies capacity limits in both working memory and enumeration. 20th Annual Meeting of the Organization for Human Brain Mapping, June 8 - 12, Hamburg, Germany.
20. **Knops, A.**, Willmes, K. (2013). Linking ordinality and cardinality – fMRI evidence for a right hemispheric network. 19th Annual Meeting of the Organization for Human Brain Mapping, June 16-20, Seattle, USA.
21. Horoufchin H., Borghi, A., Chen, H., **Knops, A.**, Goebel, R., Binkofski, F. (2012). Embodied Language - The differential role of canonical and mirror neurons in the processing of words. 18th annual Meeting of the Organization for human Brain Mapping, 79.
22. **Knops, A.,** Lonnemann, J., Koten, J., Willmes, K. (2011). What the distance congruity effect can tell us about the locus of numerical-spatial interactions. In: Bittrich, K., Blankenberger, S., Lukas, J. (Eds.) Experimentelle Psychologie. Beiträge zur 53. Tagung experimentell arbeitender Psychologen (p. 92), Lengerich: Pabst.
23. **Knops, A.**, Lonnemann, J., Willmes, K., Koten, J. (2011). Micro- and macro pattern analyses of fMRI data support both early and late late interaction of numerical and spatial information. Zeitschrift für Neuropsychologie, 22(3), 200.
24. **Knops, A.**, Nuerk, H.-C., Olsen, N., & Willmes, K. (2005). Der Kompatibilitätseffekt im Japanischen: Notations- und Kontexteffekte in Größenvergleichsaufgaben mit zweistelligen Zahlen. Lange, K.W., Bäuml, K.H., Greenlee, M.W., Hammerl, M., Zimmer, A., (Eds.) Experimentelle Psychologie: Beiträge zur 47. Tagung experimentell arbeitender Psychologen, Lengerich: Pabst.
25. **Knops, A.**, Wood, G., Nürk, H.-C., Sparing, R., Willmes K. (2005). Die Rolle des Parietalkortex bei der Verarbeitung numerischer Informationen. Tagung Deutsche Gesellschaft für Neurologie, Wiesbaden.
26. Klein, E., Nuerk, H.-C., **Knops, A.**, Wood, G., Weber, J., Willmes, K. (2004). Can „approximate" calculation processes be distinguished in a meaningful way? An eFMRI study about number processing. XXIX Congress of the European Society of Neuroradiology and XIII Advanced Course. Aachen
27. **Knops, A.**, Nuerk, H.-C., Clemens, S., Waberski, T. Willmes, K. (2004). Der Kompatibilitätseffekt: semantische oder motorische Interferenz? In: D. Kerzel, V. Franz, K. Gegenfurtner (Eds.). Experimentelle Psychologie. Beiträge zur 46. Tagung experimentell arbeitender Psychologen (p. 143), Lengerich: Pabst.
28. **Knops, A.**, Vohn, R., Nuerk, H.-C., Willmes, K. (2004). Letters or Numbers – How to assess verbal Working Memory. 10th Annual Meeting of the Organization for Human Brain Mapping, NeuroImage, 22, p.32.
29. Nuerk, H.-C., **Knops, A.,** Foltys, H., Dambeck, N., Willmes, K. (2004). Parietale Aktivierung für Zahlen: Pflicht oder Kür? Eine TMS-Studie. In: D. Kerzel, V. Franz, K. Gegenfurtner (Eds.). Experimentelle Psychologie. Beiträge zur 46. Tagung experimentell arbeitender Psychologen (p. 183), Lengerich: Pabst.
30. Vohn, R., Kraus, J**., Knops, A.,** Willmes, K., Sturm, W. (2004). Testing Paivio's Dual Encoding Theory for verbal and non-verbal learning: an fMRI study. 10th Annual Meeting of the Organization for Human Brain Mapping, NeuroImage, 22, p.33.
31. **Knops, A.,** Nuerk, H.-C., & Willmes, K. (2003). Die semantische Größenrepräsentation: Existiert mehr als nur ein numerischer Zahlenstrahl. In: J. Golz, F. Faul, R. Mausfeld (Eds.). Experimentelle Psychologie. Abstracts der 45. Tagung experimentell arbeitender Psychologen (p. 105), Regensburg: Roderer

32. Nuerk, H.-C., Wood, G., Graf, M., **Knops, A.,** & Willmes, K. (2003). The SNARC effect 10 years later: Not simply spatial, not purely numerical and not just automatic, but really amodal. Proceedings of the XII Conference of the European Society for Cognitive Psychology, Granada, Spain.

HUMBOLDT UNIVERSITÄT ZU BERLIN

Le Conseil de la faculté des Sciences de la Vie

de la Humboldt Universität zu Berlin

confère à

Monsieur André Knops

né le 29 octobre 1973 à Viersen,

suite à la procédure d'habilitation conformément au règlement de l'habilitation de la faculté de mathématiques et de sciences naturelles II du 17 janvier 2005,

L'AUTORISATION D'ENSEIGNER

dans la matière

psychologie.

Le Dr Knops a ainsi fourni la preuve qu'il peut, de manière autonome,
exercer la psychologie en tant que matière dans les domaines de l'enseignement et de la recherche.

Thème de la thèse écrite :

Neurocognitive evidence for cultural recycling of cortical maps in numeral cognition

Thème de l'exposé public :

Psychologie de la transmission d'informations erronée

Berlin, le 11 février 2015

[Signature illisible]

Prof Dr Jan-Hendrik Olbertz

Président de la Humboldt Universität zu Berlin

[Tampon de la Humboldt Universität zu Berlin]

[Signature illisible]

Prof: Dr. Richard Lucius

Décan de la Faculté
des Sciences de la Vie

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Der Rat der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin hat

Herrn André Knops
geb. am 29. Oktober 1973 in Viersen

nach einem Habilitationsverfahren gemäß der Habilitationsordnung der Mathematisch-Naturwissenschaftlichen Fakultät II vom 17. Januar 2005 die

LEHRBEFÄHIGUNG

für das Fach

Psychologie

zuerkannt.

Herr Dr. Knops hat damit den Nachweis erbracht, dass er das Fach Psychologie selbständig in Forschung und Lehre vertreten kann.

Thema der schriftlichen Habilitationsleistung:
Neurocognitive evidence for cultural recycling of cortical maps in numerical cognition

Thema des öffentlichen Vortrages:
Psychologie der fehlerhaften Informationsverarbeitung

Berlin, den 11. Februar 2015

Prof. Dr. Jan-Hendrik Olbertz
Präsident der Humboldt-Universität zu Berlin

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Prof. Dr. Richard Lucius
Dekan der Lebenswissenschaftlichen Fakultät

Je soussignée Virginie Varlet, traductrice assermentée auprès des tribunaux et des notaires de Berlin, certifie la conformité de la traduction précédente en français au document d'origine en langue allemande qui m'a été présenté.

Berlin, le 4 décembre 2015

RHINE-
-WESTPHALIA
TECHNICAL
UNIVERSITY OF
AACHEN

URKUNDE

The department of Philosophy
of the Rhine-Westphalia Technical University of Aachen

under the rectorate of Professor
of Statistics and Economic Mathematics, Dr. Burkhard Rauhut

And under the deanship of Professor of

German Linguistics, Christian Stetter, PhD

hereby awards

Mr. André Knops

born on 29 October 1973 in Viersen,

Germany, hereby receives the academic

degree

Doctor of Philosophy

(PhD)

after following proper doctoral procedures with involvement of evaluator
Professor Dr. Klaus Willmes-von Hinckeldey
Prof. Dr. Wilhelmus Spijkers (Netherlands)

With his dissertation

"On the Structure and Neural Correlates of the Numerical Magnitude Representation and its Influence in the Assessment of Verbal Working Memory"

and with an oral examination has proved his scientific competence and hereby receives the final grade of

summa cum laude.

Aachen, October 16, 2006

Rector

Dean

Department of Psychology, Philipps-Universität Marburg

DIPLOM

(DIPLOMA)

André Knops

born October 10., 1973 at Viersen

On October 08, 2001 he passed

the Examination for the Diploma in Psychology at the Philipps-Universität Marburg in accordance with the Examination Regulations introduced on 14 December, 1994

"VERY GOOD"

On 13.08.1996 he had succesfully passed the Preliminary Examination for the Diploma in Psychology at the Philipps-Universität Marburg in accordance with the Examination Regulations introduced on 14 December, 1994, with the overall result "Satisfactory".

On the basis of the above Examinations he achieved the Degree

Diplom - Psychologen

Marburg, 08 October 2001

PHILIPPS-UNIVERSITÄT MARBURG-LAHN DEKAN DES FACHBEREICHS PSYCHOLOGIE

Head of
Department of Psychology

Chair
Examination Committee

Publications presented in this thesis in reverse chronological order

1. Glaser, M. & **Knops, A.** (2020). When adding is right – temporal order judgments reveal spatial attention shifts during two-digit mental arithmetic. *The Quarterly Journal of Experimental Psychology, 73*(7), 1115–1132.
2. Didino, D., Pinheiro-Chagas, P., Wood, G., & **Knops, A. (2019).** Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect. *Frontiers in Psychology* 6;10:160. doi: 10.3389/fpsyg.2019.00160.
3. Cavdaroglu, S. & **Knops, A.** (2018). Evidence for a posterior parietal cortex contribution to spatial but not temporal numerosity perception, *Cerebral Cortex, 29(7), 2965–2977.*
4. Pinheiro-Chagas, P., Didino, D. Haase, V. G., Wood, G. & **Knops, A.** (2018). The developmental trajectory of the operational momentum effect, *Frontiers in Psychology – Developmental Psychology, 9:1062. doi: 10.3389/fpsyg.2018.01062*.
5. **Knops, A.**, Nuerk, H.-C., & Göbel, S.M. (2017). Domain-General Factors Influencing Numerical and Arithmetic Processing. *Journal of Numerical Cognition, 3(2),* 112 – 132*.*
6. Katz, C. Hoesterey, H. & **Knops, A.** (2017). A role for attentional reorienting during approximate multiplication and division. *Journal of Numerical Cognition, 3(2), 246 – 269.*
7. Cavdaroglu, S., Katz, C. & **Knops, A.** (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception, *Neuroimage*, *116*: 135-148.
8. Hesselmann, G., Darcy, N., Sterzer, P. & **Knops, A.** (2015). Exploring the boundary conditions of unconscious numerical priming effects with continuous flash suppression. *Consciousness and Cognition, 31, 60 – 72*.
9. Pinheiro-Chargas, P., Wood, G., **Knops, A.**, Krinzinger, H., Lonnemann, J., Startling-Alves, I., Willmes, K., & Haase, V. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One*. 9(11): e111155. doi:10.1371/journal.pone.0111155
10. Hesselmann, G. & **Knops, A.** (2014). No conclusive evidence for numerical priming under interocular suppression. *Psychological Science*. *25(11)*:2116-2119. *doi:10.1177/0956797614548876*
11. Katz, C. & **Knops, A.** (2014). Operational momentum in approximate multiplication and division? *PLoS One*; 9(8):e104777. doi: 10.1371/journal.pone.0104777.
12. **Knops, A.**, Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short term memory and enumeration. *The Journal of Neuroscience*, *34(30)*; 9857-9866.
13. **Knops, A.**, Thirion, B., Hubbard, E.M., Michel, V. & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. Science, 324(5934):1583-5.
14. **Knops, A.**, Viarouge A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and non-symbolic calculation: Evidence from the operational momentum effect. *Attention, Perception & Psychophysics, 71*(4), 803-821.

Synthèse de travaux

Résumé

La vie dans les sociétés industrialisées modernes exige le traitement constant d'informations numériques. Nous rencontrons des chiffres lorsque nous faisons des achats (prix, poids et tailles), que nous tapons un numéro de téléphone, que nous choisissons le quai d'une gare ou que nous vérifions la durée de retard d'un vol donné. Nous devons également surveiller en permanence et additionner nos dépenses au cours du mois pour ne pas nous endetter. C'est pourquoi le calcul a été identifié comme une compétence clé dans les sociétés modernes. La numératie est associée à des facteurs qui sont essentiels au bien-être de la société tout entière (Butterworth & Kovas, 2013 ; Gross, Hudson, & Price, 2009), y compris les progrès de l'éducation (Duncan et al., 2007). Il a été démontré qu'il permet de prédire le statut socio-économique (Ritchie & Bates, 2013), l'emploi, le salaire, la santé mentale et physique (Parsons & Bynner, 2005), et même le risque de faire face à des difficultés financières (Gerardi, Goette, & Meier, 2013).

Dans ce contexte et compte tenu de l'importance primordiale du calcul, il est surprenant que nous ne comprenions pas encore pleinement les processus cognitifs et neuraux qui sont impliqués. Mon travail vise à mieux comprendre les processus fondamentaux qui sont impliqués dans la cognition numérique. Il peut être divisé en deux grandes catégories : la perception de la numérosité et les processus fondamentaux qui sous-tendent le calcul mental.

Les humains partagent avec de nombreuses espèces la capacité d'énumérer approximativement le nombre d'éléments dans un ensemble, souvent appelé système de nombres approximatifs (SNA). Le SNA adhère à la loi de Weber, est présent très tôt dans la vie et augmente en précision avec l'âge (Halberda & Feigenson, 2008 ; Halberda, Ly, Wilmer, Naiman, & Germine, 2012 ; Piazza, 2010 ; Xu & Spelke, 2000 ; Xu, Spelke, & Goddard, 2005). Au niveau neural, l'extraction de la numérosité d'un ensemble d'éléments est corticalement mise en œuvre dans un réseau de zones occipitales et pariétales qui forment un gradient avec une spécificité de numérotation croissante au fur et à mesure que l'on passe du cortex occipital au cortex pariétal. Dans plusieurs études utilisant l'imagerie par résonance magnétique fonctionnelle (IRMf), j'ai pu démontrer que (1) la sensibilité pour les informations numériques présentées simultanément (c'est-à-dire des ensembles de points) augmente le long d'un chemin cortical allant du cortex occipital au cortex pariétal où elle est maximale. (2) Comme la signature neurale des informations numériques devient de plus en plus prononcée le long du trajet, les caractéristiques non numériques ne sont pas représentées de manière distincte dans le cortex pariétal. Par conséquent, nous n'avons pas pu classer de manière fiable les informations non numériques dans le cortex pariétal alors que la numérosité était décodable

tout au long de la voie corticale. (3) Au niveau des voxels, nous avons trouvé des preuves de l'idée que la représentation de la magnitude numérique dans le cortex pariétal postérieur suit un schéma de codage de lieu, reflétant les résultats précédents au niveau neuronale. (4) Notre paradigme nous a permis d'établir ces résultats en l'absence d'exigences de tâches liées à une réponse active. (6) Bien que cela témoigne du rôle important des zones pariétales dans la représentation des informations cardinales provenant d'ensembles d'éléments, le codage des informations numériques à partir des informations numériques séquentielles reste insaisissable. Dans deux études d'IRMf, nous n'avons observé aucune activation en réponse à des informations numériques séquentielles, ni en mode auditif (série de bips) ni en mode visuel (série de disques clignotants).

Bien que la signature neurale de la numérosité dans le cortex pariétal postérieur soit indépendante des processus liés à la réponse, le traitement de la numérosité n'est pas entièrement automatique dans toutes les circonstances. Dans une série d'expériences, nous avons rendu les amorces de la numérotation invisibles grâce à la suppression continue des flashs (continuous flash suppression). Nous n'avons trouvé aucun impact fiable des informations numériques présentées de manière non consciente sur le jugement de la magnitude d'ensembles d'objets clairement visibles. L'amorçage d'identité observé est plutôt susceptible de refléter l'impact de conditions de déclenchement d'action cognitive prédéfinies.

Il a été proposé que le traitement des ensembles d'items ne dépassant pas trois ou quatre items implique des routines cognitives et neurales qui sont résumées par le terme "subitizing". Dans cette étendu numérique, les temps de réaction et les taux d'erreur restent constants et n'augmentent pas avec le nombre d’éléments. Une théorie récente postule que la subitisation repose sur des processus neuraux qui sont également utilisés pour conserver des informations visuelles sur une courte période de temps (mémoire visuelle à court terme, MVCT). Contrairement aux conceptions précédentes de la MVCT qui supposent un nombre fixe de fentes pouvant être remplies par des éléments particuliers, Melcher et Piazza supposent que le système visuel peut affecter de manière flexible une quantité variable de ressources limitées à des éléments individuels - en fonction des exigences de la tâche et de l'importance des éléments (Franconeri, Alvarez, & Cavanagh, 2013 ; Todd & Marois, 2004). Conformément à cette hypothèse, la capacité individuelle de MVCT a été positivement corrélée avec le niveau individuel de subitisation (Piazza, Fumarola, Chinello et Melcher, 2011). En manipulant la difficulté des tâches, Melcher et Piazza ont pu influencer la quantité de ressources qui ont été affectées aux éléments individuels. Dans les tâches qui exigeaient une représentation précise et riche, plus de ressources ont été attribué aux éléments que dans les situations où il suffisait de coder la présence ou l'absence d'un élément. La capacité de subitiser était réduite lorsqu'il restait peu de ressources par rapport aux situations où les exigences de tâches étaient faibles, ce qui laissait plus de ressources pour la subitisation. En utilisant l'IRMf, nous avons pu

démontrer que le cortex pariétal postérieur supérieur contient des informations topographiques d'objets individuels que le système visuel peut exploiter de manière flexible pour les objectifs actuels. Nous montrons que les limites de capacité spécifiques à la tâche (trois à quatre objets dans le dénombrement et deux à trois dans la MVCT) est proportionnelle au niveau d'activité du cortex pariétal postérieur (CPP) : un ensemble identique de voxels dans cette région, communément activés pendant les deux tâches, a modifié son profil de réponse global reflétant les limites de capacité spécifiques à la tâche. Ce comportement témoigne de la présence d'une cartographie des caractéristiques saillantes où un petit nombre d'éléments peut être présenté avec une grande précision et un minimum de bruit pour permettre un codage riche des caractéristiques du stimulus, telles que l'orientation et la position spatiale, qui étaient nécessaires dans notre contexte de MSTV. Avec une précision plus faible, un plus grand nombre d'éléments serait représenté au prix d'une résolution plus faible des caractéristiques, bien suffisante pour une simple énumération des éléments dans un ensemble donné (Melcher & Piazza, 2011).

L'ensemble de ces résultats nécessite une mise à jour des notions théoriques centrales concernant le traitement des numérosités.

Les processus psychologiques qui nous permettent de résoudre des problèmes arithmétiques peuvent être classés en deux grandes catégories : l'élaboration sémantique et la récupération directe à partir de la mémoire à long terme. La récupération directe à partir de la mémoire à long terme caractérise certains problèmes arithmétiques, notamment celui des petites additions et des multiplications. Qu'est-ce qui caractérise l'élaboration sémantique lors du calcul mental ? Une proposition récente suggère que l'élaboration sémantique dans le contexte de la résolution de problèmes arithmétiques repose sur des processus visuo-spatiaux qui ont évolué pour des fonctions sensorielles de niveau inférieur. Le code de population représentant la position centrée sur la tête d'un objet dans le champ visuel est une combinaison additive probabiliste optimale des coordonnées centrées sur l'œil et des coordonnées de position de l'œil (Beck, Latham, & Pouget, 2011). Comme ces coordonnées sont principalement codées en PPC, ces zones accueillent donc des circuits neuronaux perceptifs de base qui sont capables d'intégrer de manière additive les résultats entrants de deux populations neuronales séparées. Par conséquent, grâce au "recyclage neuronal", l'arithmétique mentale peut coopter ces circuits pour une cognition de haut niveau. L'arithmétique mentale peut être conçue comme un processus au cours duquel plusieurs informations numériques (opérandes) doivent être intégrées (par exemple, additionnés). Chaque opérande peut correspondre à la position sur une représentation interne de la magnitude spatiale (MNL). L'intégration de ces codes de position peut alors être facilitée par les mécanismes qui nous permettent de planifier les mouvements des yeux dans l'espace, en reliant l'arithmétique mentale au domaine de l'attention spatiale.

L'effet de momentum opérationnel (MO) décrit un biais systématique dans l'estimation des résultats des problèmes d'addition et de soustraction simples. Les résultats des problèmes d'addition sont surestimés alors que les résultats des problèmes de soustraction sont sous-estimés (Knops, Viarouge, & Dehaene, 2009 ; McCrink, Dehaene, & Dehaene-Lambertz, 2007). La MO a été observée à la fois en notation symbolique (c'est-à-dire en chiffres arabes) et non symbolique (Knops, Viarouge, et al., 2009 ; Pinhas & Fischer, 2008). La MO a été interprétée comme un excellent exemple de recyclage neuronal puisqu'elle est considérée comme une conséquence de la réutilisation des circuits corticaux qui ont évolué pour le traitement spatial au cours de l'arithmétique mentale (Fischer & Shaki, 2014). Dans le même ordre d'idées, nous avons démontré que l'activité cérébrale provoquée par les saccades vers la gauche et vers la droite, qui s'accompagnent de décalages attentionnels respectifs, permet de prédire si les participants résoudraient des problèmes d'addition ou de soustraction (Knops, Thirion, Hubbard, Michel et Dehaene, 2009). Nous avons constaté que les problèmes d'addition correspondaient à l'activité neuronale associée aux saccades de droite, probablement parce que les participants s'orientaient vers des nombres plus grands du côté droit de la LMN. L'établissement des effets de base du domaine de l'attention déclenché par un calcul mental fournirait une preuve unanime de cette notion. Il a été constaté que l'attention modulait la perception à des niveaux très bas de la hiérarchie du traitement visuel. Par exemple, Carrasco et ses collègues ont observé à plusieurs reprises une augmentation de la sensibilité aux contrastes dans les endroits dans le focus d’attention au détriment d'une diminution de la sensibilité aux contrastes dans les endroits en dehors du focus attentionnel (Carrasco, 2011).

Dans deux expériences, nous avons utilisé un paradigme de jugement d'ordre temporel pour déterminer si nous pouvons observer des décalages d'attention vers la droite et la gauche pendant la phase de calcul du traitement complexe (à deux chiffres) d'addition et de soustraction, respectivement. Les deux expériences ont révélé que l’addition déplace l'attention davantage vers la droite par rapport à la tâche de base, ce qui corrobore les résultats précédents (S. X. Li & Cai, 2014 ; Liu, Cai, Verguts, & Chen, 2017 ; Masson, Letesson, & Pesenti, 2018) et soutient donc l'idée que l'attention spatiale module l'apparence (temporelle) des événements visuels. La deuxième observation principale est l'absence de décalage de l'attention spatiale vers la gauche pour la soustraction (par rapport à la tâche de base). Une possibilité est que les soustractions pourraient être résolues par des stratégies d'addition, comme cela a déjà été suggéré par certains (Masson et al., 2017). Cela aurait pu conduire à des décalages attentionnels vers la droite qui auraient masqué des décalages vers la gauche en fonction des opérations. Une autre explication possible de l'atténuation du biais spatial dans la soustraction est que la charge accrue de la mémoire de travail (c'est-à-dire la charge

cognitive) dans le traitement de la soustraction est associée à un traitement spatial réduit qui pourrait avoir conduit à des biais spatiaux moins importants.

Nous avons étudié le développement de l'effet MO dans un groupe de 162 enfants âgés de 8 à 12 ans. Alors qu'aucun biais de réponse n'a été observé pour le groupe d'âge le plus jeune, les enfants de 9 ans et plus ont montré un biais de réponse clair, c'est-à-dire qu'ils ont montré un effet MO. L'effet MO a augmenté de façon monotone avec l'âge. L'augmentation de l'effet MO s'est accompagnée d'une augmentation de la précision globale. En d'autres termes, alors que les jeunes enfants ont fait plus d'erreurs non systématiques, les enfants plus âgés ont fait moins d'erreurs, mais des erreurs systématiques. Le compte de décalage de l'attention fournit une explication possible de ces résultats en se basant sur la relation fonctionnelle entre l'attention visuospatiale et le calcul mental et sur l'influence de la scolarisation formelle. Nous proposons que l'acquisition de compétences arithmétiques pourrait renforcer le recours systématique à la ligne de nombres mentale spatiale et aux mécanismes attentionnels qui médient le déplacement le long de cette métrique. Nos résultats fournissent une étape dans la compréhension des mécanismes qui sous-tendent le calcul approximatif et une contrainte empirique importante pour les comptes courants sur l'origine de l'effet de la MO.

D'une manière plus générale, ces résultats mettent en évidence la contribution des mécanismes perceptifs et attentionnels de base qui ont évolué pour le codage spatial de l'espace visuel et le guidage des mouvements oculaires à l'arithmétique mentale de base.

Summary

Living in modern industrialized societies requires the constant processing of numerical information. We come across numbers when we go shopping (prizes, weights, and sizes), type a phone number, choose the platform in a train station or check on how long a given flight is delayed. We also need to constantly monitor and add up our expenses over the month so not to go into debt. Hence, numeracy has been identified as a key competency in modern societies. Numeracy is associated with factors that are central to the welfare of the entire society (Butterworth & Kovas, 2013; Gross, Hudson, & Price, 2009) including educational progress (Duncan et al., 2007). It has been shown to predict socioeconomic status (Ritchie & Bates, 2013), employment, salary mental and physical health (Parsons & Bynner, 2005), and even the risk of facing financial difficulties (Gerardi, Goette, & Meier, 2013).

Against this background and given the paramount importance of numeracy, it is surprising that we still do not fully understand the cognitive and neural processes that are involved. My work aims at better understanding the core processes that are involved in basic numerical cognition. It can broadly be divided into the following two topics: numerosity perception and core processes underlying mental arithmetic.

Humans share with many species the capacity to approximately enumerate the number of items in a set, often referred to as the approximate number system (ANS). The ANS adheres to Weber's law, is present very early in life, and increases in precision with increasing age (Halberda & Feigenson, 2008; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Piazza, 2010; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). At the neural level, extracting the numerosity from a set of items is cortically implemented in a network of occipital and parietal areas that form a gradient with increasing numerosity specificity as one moves from occipital to parietal cortex. In several studies using functional magnetic resonance imaging, I was able to demonstrate that (1) sensitivity for simultaneously presented numerosity information (i.e. sets of dots) increases along a cortical pathway from occipital to parietal cortex where it is maximal. (2) As the neural signature of numerical information becomes more and more pronounced along the pathway, non-numerical features are not distinguishably represented in parietal cortex. Hence, we were unable to reliably classify non-numerical information in parietal cortex while numerosity was decodable along the entire cortical pathway. (3) At the voxel level, we found evidence for the idea that numerical magnitude representation in posterior parietal cortex follows a place-coding scheme, mirroring previous findings at the single unit level. (4) Our paradigm allowed us to establish these findings in the absence of active response-related task requirements. (6) While this speaks for an important role of parietal areas in the representation of cardinal information from sets of items, the coding of numerical information

from sequential numerosity information remains elusive. Across two fMRI studies, we did not observe any activation in response to sequentially presented numerosity information, neither in auditory (series of beeps) nor in visual modality (series of flashing disks).

Despite neural signature of numerosity in posterior parietal cortex being independent from response-related processes, numerosity processing is not entirely automatic under all circumstances. In a series of experiments, we rendered numerosity primes invisible using continuous flash suppression. We found no reliable impact of non-consciously presented numerosity information on the magnitude judgment of clearly visible sets of items. Rather, the observed identity priming is likely to reflect the impact of pre-specified cognitive action-trigger conditions.

Sets of items that do not exceed three or four items have been proposed to be processed via an independent set of cognitive and neural routines that are summarized by the term subitizing. In this numerical range reaction times and error rates remain constant and do not increase with numerical magnitude. A recent theory postulates that subitizing relies on neural processes that are also used to maintain visual information over a short period of time (visual short term memory, vSTM). In contrast to previous conceptions of vSTM that assume a fixed number of slots that can be filled with particular items, Melcher and Piazza assume that the visual system can flexibly assign a variable amount of limited resources to individual items – depending on task demands and the saliency of the items (Franconeri, Alvarez, & Cavanagh, 2013; Todd & Marois, 2004). In line with this assumption, individual vSTM capacity was positively correlated with subitizing range (Piazza, Fumarola, Chinello, & Melcher, 2011). By manipulating task demands, Melcher and Piazza were able to influence the amount of resources that were assigned to individual items. In tasks that required a precise and rich representation, items were assigned more resources compared to situations in which coding the mere presence or absence of an items sufficed. Subitizing was reduced when little resources were left compared to situations where task demands were low, hence leaving more resources for subitizing. Using fMRI, we were able to demonstrate that posterior superior parietal cortex contains topographic information of individuated objects that the visual system can flexibly exploit for the current goals. We show that task-specific capacity limits (three to four items in enumeration and two to three in vSTM) are neurally reflected in the activity of the posterior parietal cortex (PPC): an identical set of voxels in this region, commonly activated during the two tasks, changed its overall response profile reflecting task-specific capacity limitations. This behavior speaks for the presence of a flexible saliency/priority map where a small number of items can be presented with high precision with minimal noise to allow for rich encoding of stimulus features, such as orientation and spatial position, that were required in our vSTM context. With lower precision, more items would be represented at the cost of lower

feature resolution, albeit sufficient for mere enumeration of items in a given set (Melcher & Piazza, 2011).

Together, these results require an update of central theoretical notions concerning the processing of numerosity information.

The psychologic processes leading to the successful mental solving of arithmetic problems can be broadly categorized into semantic elaboration and direct retrieval from long-term memory. Direct retrieval from long-term memory characterizes some arithmetic problems, notably that of small additions and multiplications. What characterizes semantic elaboration during mental arithmetic? A recent proposal suggests that semantic elaboration in the context of arithmetic problem solving relies on visuo-spatial processes that have evolved for lower-level sensory functions. The population code representing the head centered position of an object in the visual field is an optimal probabilistic additive combination of the eye-centered coordinates and the eye position coordinates (Beck, Latham, & Pouget, 2011). Since these coordinates are primarily coded in PPC, these areas thus host basic perceptual neural circuits that are capable of integrating the incoming output of two separate neuronal populations in an additive way. Consequently, via 'neuronal recycling' mental arithmetic may co-opt these circuits for high-level cognition. Mental arithmetic can be conceived of as a process during which several numerical pieces of information (operands) need to be integrated (e.g. summed up). Each operand may correspond to the position on an internal spatial magnitude representation (MNL). The integration of these positional codes may then be mediated by exactly those mechanisms that allow us to plan eye movements in space, linking mental arithmetic to the domain of spatial attention.

The operational momentum (OM) effect describes a systematic bias in estimating the outcomes of simple addition and subtraction problems. Outcomes of addition problems are overestimated while outcomes of subtraction problems are underestimated (Knops, Viarouge, & Dehaene, 2009; McCrink, Dehaene, & Dehaene-Lambertz, 2007). The OM has been observed with addition and subtraction, both in symbolic (i.e. Arabic digits) and non-symbolic notation (Knops, Viarouge, et al., 2009; Pinhas & Fischer, 2008). OM has been interpreted as a prime example of neuronal recycling since it is considered a consequence of the reuse of cortical circuits that have evolved for spatial processing during the course of mental arithmetic (Fischer & Shaki, 2014). In line with this we demonstrated that brain activity elicited by left- and rightward saccades, which are accompanied by respective attentional shifts, predict whether participants were performing centrally-presented addition or subtraction problems (Knops, Thirion, Hubbard, Michel, & Dehaene, 2009). We found that addition problems corresponded to the neural activity associated with right-ward saccades, presumably since participants oriented towards larger numbers on the right side of the MNL. Establishing the hall mark effects

from the attention domain as a consequence of mental arithmetic would provide unanimous evidence for this notion. Attention has been found to modulate the perception at very low levels of the visual processing hierarchy. For example, Carrasco and colleagues repeatedly observed increased contrast sensitivity at attended locations at the expense of decreased contrast sensitivity at unattended locations (Carrasco, 2011).

In two experiments, we used a temporal order judgment paradigm to investigate whether we can observe attentional shifts to the right and left during the calculation phase of complex (two-digit) addition and subtraction processing, respectively. Both experiments revealed that addition processing shifts attention more rightward compared to the baseline task which corroborates earlier findings (S. X. Li & Cai, 2014; Liu, Cai, Verguts, & Chen, 2017; Masson, Letesson, & Pesenti, 2018) and hence supporting the idea that spatial attention modulates the (temporal) appearance of visual events. The second main observation of the present study is the absence of spatial attention shifts to the left for subtraction (compared to the baseline). One possibility is that subtraction tasks might be solved via addition strategies as it has already been suggested by some (Masson, Pesenti, & Dormal, 2017). This might have led to attentional shifts to the right which masked operation-congruent shifts to the left. Another possible explanation for the attenuated spatial bias in subtraction is that the increased working memory load (i.e. cognitive load) in subtraction processing is associated with reduced spatial processing which might have led to less spatial biases.

We investigated the development of the OM effect in a group of 162 children from 8 to 12 years old. While no response bias was observed for the youngest age group, children aged 9 and older showed a clear response bias, that is they showed an OM effect. The OM effect monotonically increased with age. The increase of the OM effect was accompanied by an increase in overall accuracy. That is, while younger children made more and non-systematic errors, older children made less but systematic errors. The attentional shift account provides a possible explanation of these results based on the functional relationship between visuospatial attention and mental calculation and on the influence of formal schooling. We propose that the acquisition of arithmetical skills could reinforce the systematic reliance on the spatial mental number line and attentional mechanisms that mediate the displacement along this metric. Our results provide a step in the understanding of the mechanisms underlying approximate calculation and an important empirical constraint for current accounts on the origin of the OM effect.

On a more general note, these results point to the contribution of basic perceptual and attentional mechanisms that have evolved for spatial coding of visual space and guidance of eye movements to basic mental arithmetic.

List of 5 scientific activities as PI

Publications

Since I finished my PhD in 2006, I published 34 articles in international journals, 8 first-author papers, 14 senior author papers and 12 co-authored papers. I also published one book and one work book for elementary school children.

Selected publications:

1. **Knops, A.**, Thirion, B., Hubbard, E.M., Michel, V. & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. Science, 324(5934):1583-5.

   In this paper, I used innovative fMRI analysis techniques to test whether activity in attention-related brain circuits is involved during mental arithmetic. This represents a foundation for my subsequent work.

2. Cavdaroglu, S. & **Knops, A.** (2018). Evidence for a posterior parietal cortex contribution to spatial but not temporal numerosity perception, *Cerebral Cortex, 29(7), 2965–2977.*

   This paper allowed me to test a corner stone of the Triple-Code Model of Number Processing and presents a building block for future work.

3. Glaser, M. & **Knops, A.** (in press). When adding is right – temporal order judgments reveal spatial attention shifts during two-digit mental arithmetic. *The Quarterly Journal of Experimental Psychology*.

   This experiment allowed me to test whether well-known characteristics of spatial attention can be observed in the context of a 'cognitive' task. This is an important proof of concept for my theory that spatial attention contributes to mental arithmetic.

Book

Knops, A. (2019). *Numerical Cognition.* Abingdon, Oxon; New York, NY: Routledge.

In this book I summarize the basic knowledge the community has accumulated over the last years in a digest and accessible manner that is suited for a large, not necessary academic readership. This helped me gain more visibility inside and outside the academic community.

Organized Workshop

Between Embodiment and Neuronal Recycling – Contribution of the Dorsal Stream to Visual Perception, Workshop, Frankfurt/Main (7 – 8/06/2015). Organizers: **André Knops** & Guido Hesselmann

This international workshop brought together 8 speakers from 5 European countries and inspired a discussion of the underlying organizational and functional principles that characterize dorsal stream. It was also important for enlarging my international network beyond my own research domain.

Introduction

Living in modern industrialized societies requires the constant processing of numerical information. We come across numbers when we go shopping (prizes, weights, and sizes), type a phone number, choose the platform in a train station or check on how long a given flight is delayed. We also need to constantly monitor and add up our expenses over the month so not to go into debt. Hence, numeracy has been identified as a key competency in modern societies. Numeracy is associated with factors that are central to the welfare of the entire society (Butterworth & Kovas, 2013; Gross et al., 2009), including educational progress (Duncan et al., 2007). It has been shown to predict socioeconomic status (Ritchie & Bates, 2013), employment, salary mental and physical health (Parsons & Bynner, 2005), and even the risk of facing financial difficulties (Gerardi et al., 2013).

Numeracy is often taken as a direct index for general intelligence. People who struggle with mathematics fear being intellectually insufficient. Also, understanding of mathematics is also perceived as a stable and persistent trait - irremediable. Contrary to this common thinking, various intervention studies have demonstrated that mathematical understanding is malleable. Yet - the key lies in a better understanding of the cognitive mechanisms that underlie mathematical cognition. Only when we know what process or competency we need to tackle and in which way, can we expect significant progress.

Numerical cognition is a multi-facetted domain, comprising the investigation of diverse but complementary processes that range from the core mechanisms involved in the perception of numerical information and assigning a number to the perceived information all the way to the algorithmic process of solving arithmetic problems (Figure 1).

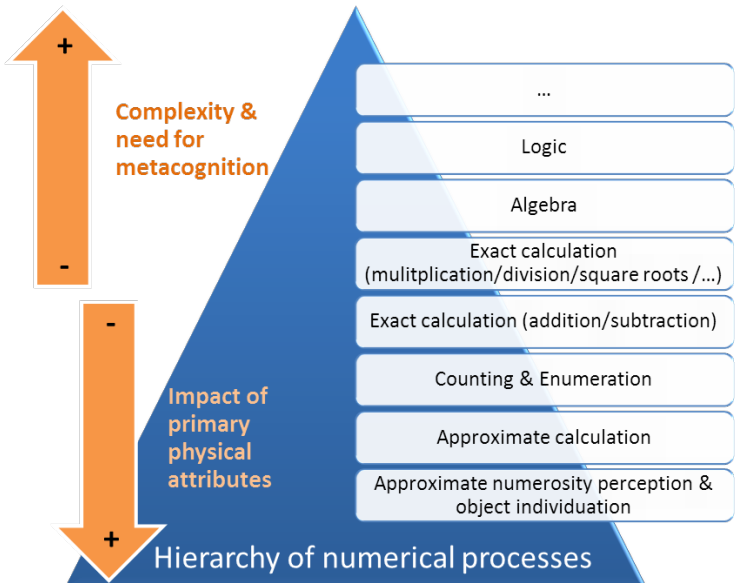

Figure 1: Schematic hierarchy of numerical cognition processes according to their complexity.

Each of these skills comprises the involvement of numerous domain-general and domain-specific processes. For example, when asked to verify the outcome of a simple mental arithmetic problem such as 8 * 4 = 28, one has to access the numerical meaning of the operands and retrieve the correct outcome (32) from long-term memory. Arithmetic fact knowledge is conceptualized as an associative network, in which the representations of operands and results are highly interconnected (see for example, Ashcraft, 1987; Campbell, 1995; Verguts & Fias, 2005). The retrieval process is thought to be driven by an automatic activation spreading within the network (Galfano, Penolazzi, Vervaeck, Angrilli, & Umilta, 2009; Galfano, Rusconi, & Umilta, 2003; Niedeggen, Rosler, & Jost, 1999; Rusconi, Galfano, Rebonato, & Umilta, 2006; Rusconi, Galfano, Speriani, & Umilta, 2004). Namely, following the presentation of two operands (e.g., 8 x 4), activation is spread so that a series of possible results (likely the product (32) and the multiples of the operands close to it (e.g. 24, 28, 36, and 40)) is activated. The highest activated result is retrieved as the actual result. Since the network is a highly interconnected associative network, activation can spread both from operands to results and between results themselves (see, for example, the network interference model, Campbell, 1995). As a consequence, one has to inhibit the response tendency to approve the provided but incorrect solution, although it is numerically related (i.e. 28 is a multiple of 4). Interestingly, we recently demonstrated that this process is more difficult for incorrect solutions that are related but larger compared to incorrect solutions that are related but smaller than the correct outcome (Didino, Knops, Vespignani, & Kornpetpanee, 2015). This implies that the functional architecture of the nodes within the network reflects some of the organizing principles of the numerical magnitude representation which is thought to take the form of a spatially organized mental number line (MNL). One key characteristic of the MNL is

its logarithmic compression, meaning that the distance between neighboring entries becomes increasingly smaller as the numerical magnitude increases. In turn, this means that co-activation of neighboring entries increases as numerical magnitude increases. Before this background the observed difficulty to reject related but larger outcomes may reflect stronger co-activation within the associative network due to the semantic relationship of the entries on the MNL.

This example illustrates the complexity of cognitive processes characterizing allegedly simple arithmetic fact retrieval. It also illustrates how the analog and spatial character of the MNL may influence mental arithmetic.

Over the last years I have pursued the characterization of (a) the nature and neural implementation of numerical core processes and (b) how they change across life-span. As illustrated in Figure 1, higher order numerical competencies are supposed to be grounded in cognitive primitives, that is, core capacities that develop early in life and allow the approximation of numerical quantities. In the following, I will first summarize my work that can broadly be divided into the topics of (1) numerosity perception and (2) core processes underlying mental arithmetic. Then, I will delineate my research plans for the years to come.

Numerosity Perception

[Publications on this topic:

- *Cavdaroglu, S. & Knops, A. (2018). Evidence for a posterior parietal cortex contribution to spatial but not temporal numerosity perception, Cerebral Cortex, 29(7), 2965–2977.*
- *Knops, A. (2017). Probing the neural correlates of number processing. The Neuroscientist, 23 (3), 264-274.*
- *Cavdaroglu, S., Katz, C. & Knops, A. (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception, Neuroimage, 116: 135-148.*
- *Hesselmann, G., Darcy, N., Sterzer, P. & Knops, A. (2015). Exploring the boundary conditions of unconscious numerical priming effects with continuous flash suppression. Consciousness and Cognition, 31, 60 – 72.*
- *Hesselmann, G. & Knops, A. (2014). No conclusive evidence for numerical priming under interocular suppression. Psychological Science. 25(11):2116-2119. doi:10.1177/0956797614548876*
- *Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short term memory and enumeration. The Journal of Neuroscience, 34(30); 9857-9866.]*

Approximate Number System

Humans share with many species the capacity to approximately enumerate the number of items in a set, often referred to as the approximate number system (ANS). The ANS adheres to Weber's law. When participants compare the magnitude of two numbers or two numerosities (number of items in a set) their performance decreases (i.e. longer reaction times and higher error rate) as the numerical distance between to-be-compared numbers decreases – the so-called distance effect (Dehaene, Dupoux, & Mehler, 1990; Moyer & Landauer, 1967). Moreover, as numerical magnitude of to-be-compared numbers increases, the distance between numbers must increase proportionally to achieve comparable performance (size effect). For numerosities, the ratio between two items defines performance in comparison tasks. The ANS is present very early in life and increases in precision with increasing age (Halberda & Feigenson, 2008; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Piazza, 2010; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). For example, while 6-month-olds are sensitive to numerical changes with a ratio of 1:2 (e.g. 6 vs. 16 items), infants at the age of 9 months can already discriminate numerosities differing by a ratio of 2:3 (e.g. 8 vs. 12 items; (Lipton & Spelke, 2003). Adults usually succeed in discriminating numerosities differing by a ratio of 7:8 (Knops, Dehaene, Berteletti, & Zorzi, 2014). Importantly, some conceive of the ANS as a unitary module that is activated for numerical information from different notations (e.g. Arabic digits or set of dots (Dehaene & Cohen, 1995; Eger et al., 2009b; Piazza & Eger, 2015; Piazza, Pinel, Le Bihan, & Dehaene, 2007) or modes (sets of dots or series of sequentially presented dots (Castelli, Glaser, & Butterworth, 2006; Nieder & Dehaene, 2009; Nieder, Diester, & Tudusciuc, 2006). Others argue, however, that numerosity information is indirectly inferred from correlated non-numerical stimulus features, such as density, convex hull (the shortest contour that can be drawn around a given set of item), or the individual item size

(Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011), (Bulthé, Smedt, & Op de Beeck, 2014; Cohen Kadosh, Cohen Kadosh, Kaas, Henik, & Goebel, 2007; Lyons, Ansari, & Beilock, 2015).

Neural and cognitive mechanisms of numerosity perception

Currently, there is a lively debate concerning the extent to which the estimation of a numerosity depends on non-numerical features. On the one hand there is mounting evidence that suggests that numerical estimates are directly perceived, independent from visual characteristics. At the neural level, a number of studies supports the idea that numerosity is extracted along an occipital-parietal pathway – both at the neural and the neuronal level (Eger et al., 2009a; Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2004). In an adaptation fMRI study, Piazza and colleagues repeatedly presented dot displays with either 20 or 50 dots to human participants in the scanner (Piazza et al., 2007). They observed two effects in bilateral areas around the IPS. First, activity decreased as one numerosity was repeated over and over again (in fact they changed all non-numerical parameters of the dot sets except the number of dots), reflecting the adaptation of the neural system. Second, when they changed numerosity (e.g. from 50 to 20 dots), activity increased and reached pre-adaptation level. Hence, bilateral areas around the IPS were sensitive to changes in numerosity. The organization of parietal cortex may even bear more resemblance with the metaphor of the MNL. At the macroscopic neural level, a recent study recently found a topographic organization of numerosity-tuned voxels in fMRI. That is, in human parietal cortex, a spatially organized gradient of numerosity-specificity was observed (Harvey, Klein, Petridou, & Dumoulin, 2013). Harvey and colleagues, too, presented participants with sets of dots in various changing layouts while recording brain activity with a high-field (7 Tesla) scanner. They found that specific voxels at one end of the topographically organized parietal area were tuned to numerosity one, while adjacent voxels were tuned to two. This spatial organization continued until numerosity 7, being located at the other extreme of the area. The cortical area devoted theses numerosities decreased as numerosity increased. Most area was devoted to numerosity one while numerosity seven showed minimal cortical extend. In a nutshell, these results can be described as a cortical instantiation of the number line, including logarithmic compression and independent from non-numerical features. This suggests that the cortical organization of numerical magnitude follows similar principles as other visual dimensions such as orientation or contrast which are known to form cortical maps.

On the other hand, it has been argued that numerosity is indirectly derived by weighing up the quantity information from different dimensions including the ratio of energy in different spatial frequency bands (Dakin et al., 2011), or all available sensory information such as overall occupied area, density, size and other dimensions together (Gebuis, Cohen Kadosh, & Gevers, 2016). Supporting evidence for this idea comes from studies that find numerosity

discrimination to be influenced by the variation of non-numerical stimulus parameters, both in behavioral ((Smets, Sasanguie, Szücs, & Reynvoet, 2015; Tokita & Ishiguchi, 2010, 2012) and in electroencephalographic (Gebuis & Reynvoet, 2014) measures. Other studies did not converge on these findings, however (Guillaume, Mejias, Rossion, Dzhelyova, & Schiltz, 2018; Pinheiro-Chagas et al., 2014). Neural overlap of activations in response to quantity changes along various dimensions such as physical size (Pinel, Piazza, Le, & Dehaene, 2004), have been taken as evidence for a holistic processing of numerosities (Leibovich, Katzin, Harel, & Henik, 2016). In fact, a theory of magnitude has been proposed that assumes that space, time and number share common cortical circuits in parietal cortex (Bueti & Walsh, 2009).

Most of these alternative approaches do not yet provide the level of theoretical detail that the direct approach provides and/or have not been extensively tested. As an exception, a recent study directly pitted these ideas against each other (Harvey & Dumoulin, 2017). In their ultra-high field fMRI study, Harvey and Dumoulin presented participants with sets of dots that systematically varied with respect to which low-level, non-numerical feature was confounded with numerosity (e.g. overall area, density, item perimeter). Using a population receptive field model, they then fitted which parameter would best describe the observed data in parietal areas (i.e. explain maximal amount of variance). When comparing the amount of variance explained by the different model, numerosity captured significantly more variance compared to visual feature models. Moreover, several biologically plausible computational models successfully simulated the evolution and performance of a numerosity system that shares central features with the assumption of an approximate number system (Miller, 2013; I. Stoianov & Zorzi, 2012; I. P. Stoianov & Zorzi, 2017; Verguts & Fias, 2004b).

Computational models of numerosity perception propose a sequence of dedicated mechanisms that are involved to get to this number tuning (Dehaene & Changeux, 1993; Verguts & Fias, 2004a). Dehaene and Changeux proposed the following hierarchy of processing steps: The raw visual input is normalized and represented in a topologically organized map of the scene with activation peaks indicating the position of the elements in the scene. Dedicated units sum the activation peaks from this map, leading to a monotonic increase of activity in this number-sensitive instance activity with increasing number of objects (Dehaene & Changeux, 1993). The summed activation is then fed into the next instance that contains number-selective units. Here, units are tuned to specific numerosities and activity decreases monotonically as numerical distance between preferred and actual numerosity increases.

This sequence has been empirically observed in an fMRI study in humans (Roggeman, Santens, Fias, & Verguts, 2011), adding further empirical evidence to the idea of direct numerosity perception. Roggeman and colleagues tested how these hierarchical instances of

numerosity perception map onto the neural system (Roggeman et al., 2011). Participants were adapted to visually presented dot arrays. After adapting to a given numerosity (e.g. 3), deviants were presented that varied either in location of the dots or in terms of numerosity. By categorizing numerosity deviants into the category smaller-larger (e.g. 1 & 2 vs. 4 & 5) or close-far (e.g. 2 & 4 vs. 1 & 5), this design allowed the mapping of neural circuits that either followed a summation coding or a number-selective scheme. The former would show increased activity for larger compared to smaller deviants. The latter would show increased activity for far compared to close deviants. The authors found a hierarchical organization along an occipital-to-parietal pathway (Roggeman et al., 2011). While occipital regions exhibited the strongest activation recovery for location deviants, areas in superior occipital cortex and the adjacent transition region between occipital and parietal cortex were most sensitive to deviants from the smaller-larger category, indicating a summation coding scheme. Areas in PSPL and IPS exhibited a recovery profile that implied number selectivity. This means that activity in these areas increased as the numerical distance between adapted numerosity and deviant increased – a neural instantiation of the distance effect. Together, this suggests that numerosity is extracted from visual input in areas along an occipital-to-parietal pathway. As the activity travels out from occipital cortex into parietal cortex, it first passes through number sensitive neural circuits in middle occipital and PSPL exhibiting a summation coding schema, before reaching number selective circuits in the IPS.

Cavdaroglu and Knops (2018) extended these findings to a larger numerical range and found – in essence – a similar gradient of increasing numerosity specificity as one travels from occipital to parietal cortex. we tested numerosities outside the subitizing range and isolated perception from decision and response-related processes. We found an increase in the parietal BOLD signal during the presentation of simultaneous numerosities but not during the presentation of sequential numerosities. Using MVPA, we successfully trained a classifier to decode simultaneous numerosity from the BOLD signal in the parietal cortex, providing further confirmation of numerosity selective activity in these areas. No better-than-chance classification was observed for sequential numerosities in the same ROIs (Figure 2). These results imply distinct underlying coding schemes for sequential and simultaneous numerosities. In parietal areas we were able to successfully differentiate between simultaneous and sequential presentation modes using MVPS.

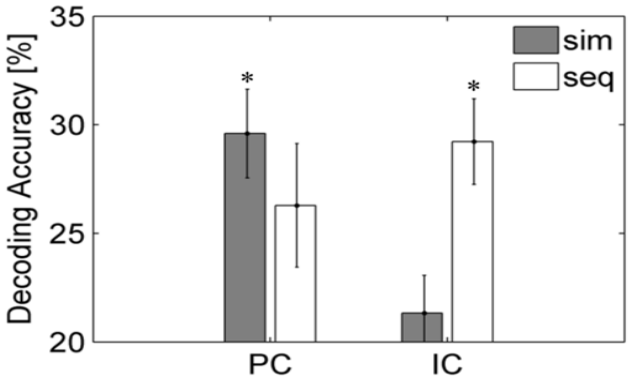

Figure 2: Decoding accuracies for simultaneous (sim) and sequential (seq) numerosities in parietal (PC) and insular cortex (IC) ROIs. Chance level is at 25%. Stars indicate performance significantly better than chance.

Concentrating on parietal ROIs only, we used MVPA to further explore how the encoding of simultaneous numerosity and other visual features (i.e., convex hull, total area, density, and diameter) changes when going from the primary visual cortex to the parietal cortex. Striate and extrastriate areas gave rise to successful classification of both non-numerical visual features and numerosity. In contrast, parietal ROIs allowed for decoding of numerosity and dot diameter only. This suggests a more abstract, higher-level mental representation in the parietal cortex beyond sensory features.

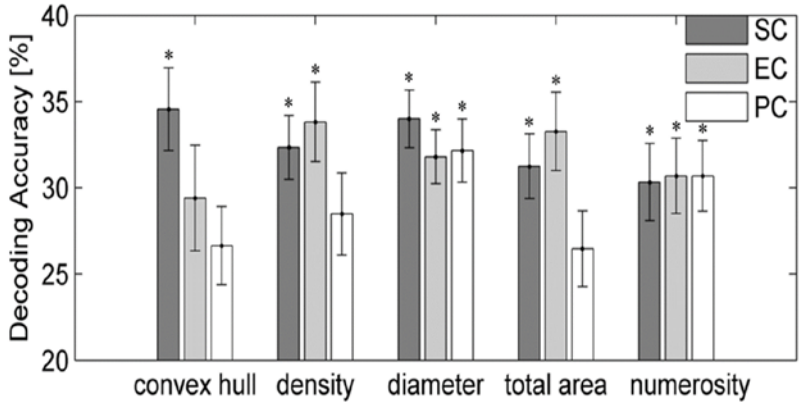

Figure 3: Decoding accuracies for simultaneous numerosities in striate (SC), extra-striate (EC), and parietal cortex (PC), separately for convex hull, density, diameter, total area and numerosity. While decoding accuracy of numerical information increases as one moves from occipital to parietal cortex (see Figure 5C), visual feature classification fails in PC.

Finally, we found voxel-wise numerosity tuning functions for simultaneous numerosities in occipital and parietal ROIs. That is, we re-grouped voxels according to their preferred numerosity (as defined by maximal activity) and analyzed their response to non-preferred

numerosities. In line with the idea of place-coding and paralleling results at the single-unit level, neural response monotonically decreased as numerical distance between preferred and presented numerosity increased (see Figure 4). This numerosity selectivity increased along an occipitoparietal gradient reaching maximal selectivity in parietal areas (see Figure 4C). We observed overlapping summation coding profiles (higher BOLD signal for large numerosities vs. small numerosities) for sequential and simultaneous numerosities in low-level visual areas only.

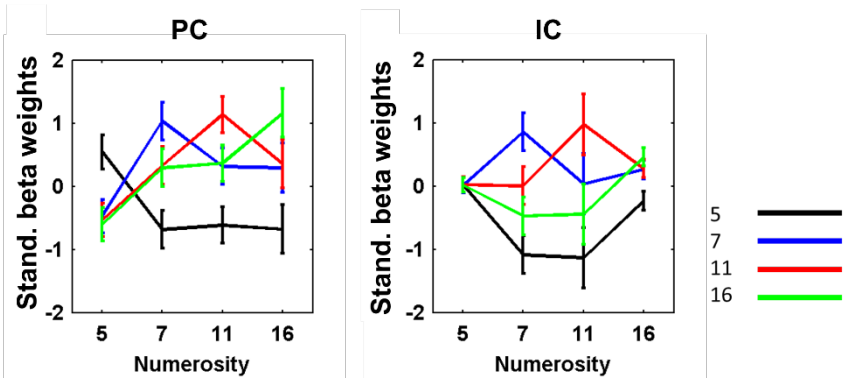

Figure 4: Tuning curves (based on standardized beta weights) for voxels tuned to simultaneous numerosities in parietal cortex (PC, left) and sequential numerosities in insular cortex (IC, right). Error bars represent standard error of the mean.

These data are in line with the idea that numerosity is processed along a dorsal cortical gradient, as proposed by Roggeman and colleagues (2011). We also found that parietal areas around the IPS give rise to lower discriminability of non-numerical stimulus features compared to earlier visual areas. Conversely, the representational precision of numerosity increases along the occipital to parietal gradient. With respect to the question what stimulus parameter drives parietal activity and hence which is, in turn, considered to define behavioural performance (Lasne, Piazza, Dehaene, Kleinschmidt, & Eger, 2019), my findings are in line with a recent study that shows that numerosity is a better predictor of posterior parietal activity compared to low-level, non-numerical features (Harvey & Dumoulin, 2017).

Does numerosity integration over time differ from integration over space?

While evidence is accruing for a hierarchical organization of parietal cortex during the extraction of spatially scattered numeric information from a scene, the mechanisms allowing humans to estimate the number of sequentially presented items remain poorly understood. If the numerical magnitude representation in IPS is truly abstract, sequential and simultaneous numerosities should activate identical cortical circuits. Single-unit recording study in primates

reported shared populations of number-selective neurons in the depth of the IPS for the *maintenance* of both temporal and spatial numerosity (Nieder et al., 2006). However, the same study reported that distinct neuronal populations in the depth of the IPS were recruited during the *extraction* of simultaneous and sequential numerosities, pointing to distinct processes during numerosity extraction. We recently found that distinct cortical areas were associated with temporal and spatial enumeration (Figure 5).

As shown in Figure 5B (Cavdaroglu & Knops, 2018), we observed activation in distinct cortical areas during the processing of simultaneous (yellow) and sequential (blue) numerosities, overlapping mostly in occipital cortex (green). While activity for simultaneous numerosity resembled previously reported patterns of activity and extended into posterior parietal areas, no parietal activity was observed for sequential numerosities. We found that insular cortex was involved in sequential numerosity perception, yet with a qualitatively different coding scheme (see Figure 5 right). While no pronounced place coding was observed for intermediate numerosities (note that the peak at the preferred numerosity is trivial since voxels were chosen as a function of maximal activity for a given numerosity), we observed U-shaped tuning functions for the smallest and largest numerosities. This double-peak tuning profile has recently been predicted theoretically by Hannagan and colleagues (Hannagan, Nieder, Viswanathan, & Dehaene, 2017) and empirically observed at the unit level by Nieder and Merten (Nieder & Merten, 2007). However, our data differ in terms of (a) scale and (b) location. That is, we find this behavior at the voxel level, combining a large number of neuronal populations rather than at the unit level. Hence, the double-peak tuning must be present in a large number of neurons. Also, we observed this behavior in the Insular cortex which is sometimes argued to be part of a saliency network (Menon & Uddin, 2010) or to be involved in discrimination in the context of time processing (Kosillo & Smith, 2010).

Finally, we were able to reliably decode numerosity in insular cortex using MVPA (Figure 5).

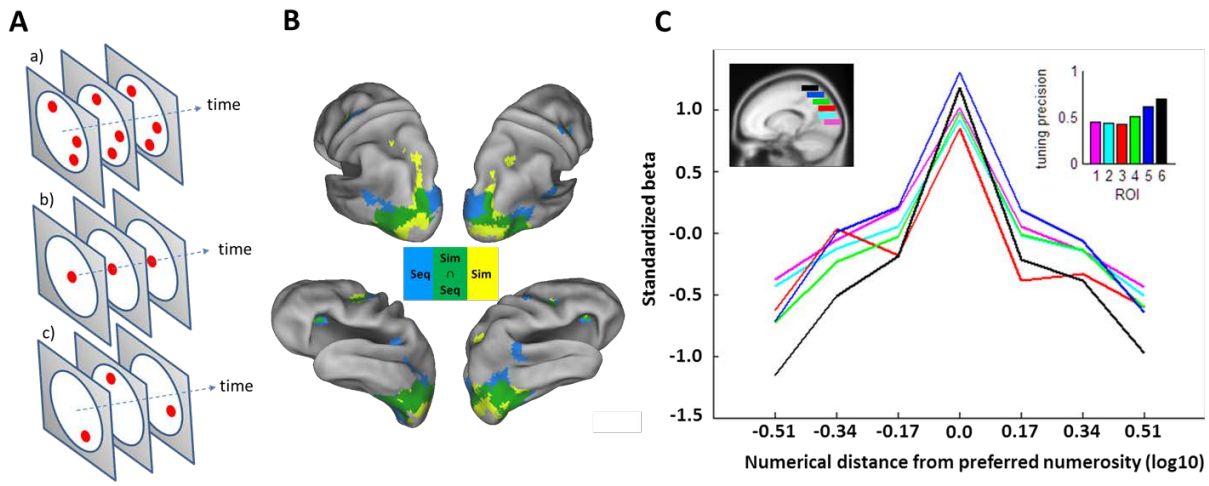

Figure 5: Temporal and spatial enumeration. A – Schematic depiction of simultaneous (a) and sequential (b & c) numerosities. Sequential stimuli in B and C correspond to (b), i.e. centrally flashed disks. B - Brain activation for simultaneous (yellow) and sequential (blue) numerosities projected on inflated templates of left and right hemisphere. Overlapping activity is shown in green. C - Voxel-wise numerical tuning curves along an occipital-parietal gradient (see ROI inset top-left) for simultaneous numerosities. Tuning precision (i.e. width of the tuning curve) increases from occipital to parietal areas, implying increasing number selectivity along this dorsal pathway. The tuning profiles for sequential numerosities (not shown here) did neither exhibit the same regular and mostly monotonic decrease as a function of deviation from preferred numerosity nor any systematic gradient from occipital to parietal areas.

These findings reverberate with previous findings in my lab (Cavdaroglu, Katz, & Knops, 2015). We measured the brain activation of healthy adults performing estimation and/or comparison of sequential visual (series of dots) and auditory (series of beeps) numerosities. Our experimental design allowed us to separate numerosity estimation from comparison and response related factors. The BOLD response in the parietal cortex only increased when participants were engaged in the comparison of two consecutive numerosities that required a response.

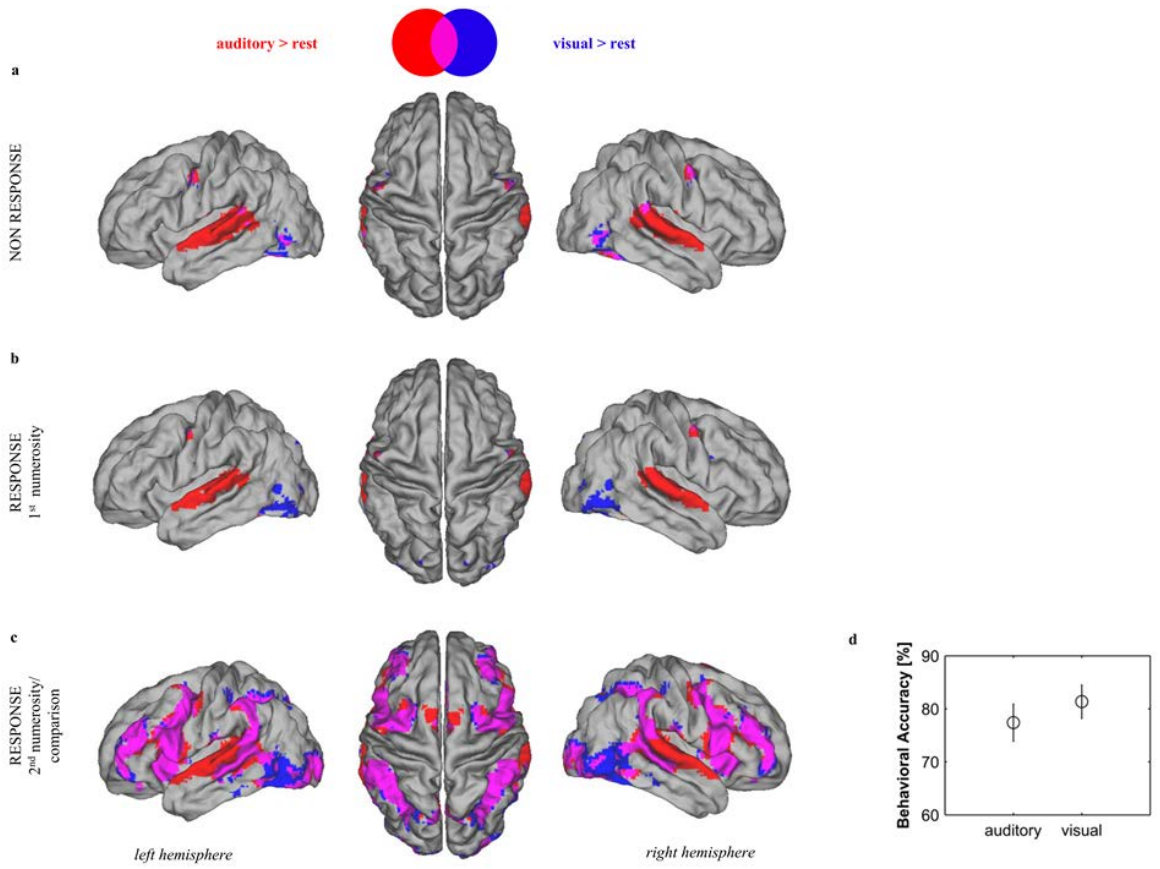

Figure 6. a) The parametric modulation of BOLD activity with increasing numerosity for the numerosity-GLM, for auditory (red) and visual numerosities (blue). The auditory cortex and frontal cortex were activated for auditory numerosities and V5 and frontal cortex were activated for visual numerosities (FDR corrected at $p < .05$ on cluster level). b) The parametric modulation of BOLD activity from the first numerosity (in response trials) > rest contrast, for auditory (red) and visual (blue) modality. c) The parametric modulation of BOLD activity from the probe numerosity (second numerosity in response trials) > rest contrast, for auditory (red) and visual (blue) modality. In addition to sensory cortices, parietal and frontal areas exhibit significant increase in BOLD activity (FDR corrected at $p < .05$ on cluster level). d) Accuracy of participants in comparing the numerosities in response trials; separately for auditory and visual modalities.

Using multivariate pattern analysis, we trained a classifier to decode numerosity in various regions of interest (ROI). We failed to find any parietal ROI where the classifier could decode numerosities during the estimation phase, that is without any response-related processes (selection, preparation, execution) involved. Rather, when participants were not engaged in comparison we were able to decode numerosity in sensory cortices. We decoded numerosity in auditory cortex ROI for auditory stimuli and in a visual cortex ROI for visual stimuli. On the other hand, during response period the classifier successfully decoded numerosity information in a parietal ROI for both visual and auditory numerosities. These results were further

confirmed by support vector regression. In comparison to a support vector classification, support vector regression makes assumptions concerning the relation between stimulus categories. Since numerosity is not merely categorically organized but also inherently ordinal in nature, we applied SVR to test for linearly ordered decoding performance. SVR results were in line with the SVC results. In sum, our study does not support the involvement of the parietal cortex during estimation of sequential numerosity in the absence of an active task with a response requirement.

Hence, the neural mechanisms underlying temporal enumeration remain elusive and point to the involvement of insular cortex. At the very least, the results are at odds with the assumption that sequential and simultaneous numerosity perception rely on redundant cortical networks.

Is non-symbolic numerosity information processed automatically?

At this point one may wonder whether numerosity information is processed automatically or whether we need to focus attention to the processes described previously. Masked priming studies with *symbolic* numbers found that congruent primes would speed up responses while incongruent primes would slow responses down. In masked priming studies, participants are presented with a target digit for which they would need to indicate whether or not it is numerically larger or smaller than a predefined standard (e.g. 5). The target is preceded by the presentation of a prime digit. To prevent participants from consciously perceiving the prime, it is temporally embedded in a random letter string. That is, immediately before the prime comes on screen a letter string (forward mask; e.g. XTKLW) is presented for 70 ms. The prime stays on screen for only a brief period (e.g. 30 ms) before it is replaced by another letter string, the backward mask. This procedure prevents the participant from consciously perceiving the prime digit. Interestingly, when both prime and target are smaller or larger than the standard, reaction times are faster compared to trials in which prime and target are associated with diverging responses (e.g. prime: 2 target: 6)(Naccache & Dehaene, 2001b). This priming was observed across notations (e.g. prime: number word target: Arabic digits) and extended to primes that were not consciously seen throughout the entire experiment (Naccache & Dehaene, 2001b). Together, this corroborates the idea that numerical magnitude is automatically processed – even when not being consciously perceived (see also (Huckauf, Knops, Nuerk, & Willmes, 2008); but see (Kunde, Kiesel, & Hoffmann, 2003) for an alternative explanation).

With respect to the question of whether this automaticity also holds for non-symbolic numerosity information is currently debated. The findings of a recent study (Bahrami et al., 2010) suggested that numerical processing of small quantities (1-3) can escape masking via continuous flash suppression (CFS) and lead to robust numerical priming effects in an

enumeration task. In CFS, high-contrast dynamic patterns shown to one eye disrupt the conscious perception of a low-contrast stimulus shown to the other eye for up to several seconds (Tsuchiya & Koch, 2005). In contrast to binocular rivalry, onset and offset of stimulus suppression can deterministically be controlled by switching the dynamic CFS masks on and off, respectively. Behavioral and neuroimaging studies using CFS have already produced a large but heterogeneous body of evidence regarding the types of visual information that can be processed during this powerful variant of interocular suppression (Hesselmann, 2013; Sterzer, Stein, Ludwig, Rothkirch, & Hesselmann, 2014). The results observed by Bahrami and colleagues contradict earlier findings suggesting a) that binocular rivalry abolishes visual semantic priming (Cave, Blake, & McNamara, 1998; Zimba & Blake, 1983), and b) that semantic analysis does not occur in the absence of awareness induced by CFS (Kang, Blake, & Woodman, 2011). Beyond these inconsistencies, which could simply mean a) that CFS is supported by mechanisms distinct from binocular rivalry (Tsuchiya, Koch, Gilroy, & Blake, 2006), and b) that CFS leaves numerical but not semantic processing selectively intact, the specific pattern of response times (RTs) reported by Bahrami and colleagues (2010) merits further investigation from a conceptual point of view. In a series of three priming experiments (and two control experiments to estimate prime visibility), the Bahrami and colleagues showed that unconsciously presented non-symbolic and symbolic primes (sets of Gabor patches and Arabic digits, respectively) induced a priming effect for non-symbolic numerosity targets which was dependent on the numerical distance between target (t) and prime (p). Specifically, the priming effect signaled “interference” (i.e., slower RTs relative to a prime-absent baseline) for negative t-p distances (e.g. target: 4, prime: 5, t-p distance: -1), and “facilitation” for positive t-p distances. Facilitatory priming was relatively small and less robust for zero t-p distance (i.e., numerically congruent trials). While indeed intriguing and robust across experiments, this pattern of results is in fact difficult to reconcile with previous findings from priming studies with visible and invisible numerosity primes. For example, it has usually been found that when a target is preceded by a prime number, participants respond more quickly when the absolute t-p distance is smaller (Dehaene et al., 1998; Koechlin, Naccache, Block, & Dehaene, 1999). This well-established feature of numerosity priming is generally explained by representational overlap between the prime and the target (Van Opstal, Gevers, De Moor, & Verguts, 2008). Furthermore, the pattern of priming across t-p distances has been shown to depend on the notation of the prime: While V-shaped priming (centered on zero t-p distance) was found for symbolic digit primes, a step-like priming function resulted from trials with non-symbolic dot primes (Roggeman, Verguts, & Fias, 2007). Bahrami and colleagues (2010), however, reported roughly identical linear priming functions for both non-symbolic and symbolic primes (Exp. 2 and 3, respectively).

In three behavioral experiments, we (Hesselmann & Knops, 2014) have explored the boundary conditions of distance-dependent numerical priming effects when primes are rendered invisible by interocular suppression (CFS). In the first experiment, we found a large priming effect following a linear priming function, as reported previously (Bahrami et al., 2010), but concluded that this effect was confounded with an effect of target numerosity due to a flaw in the original experimental design and data analysis adopted by Bahrami and colleagues. We proposed an alternative and well-established data analysis to overcome this confound. The second experiment supported our hypothesis of a pervasive confound because a reduction of the effect of target numerosity resulted in a reduction of the priming effect. Using an optimized experimental design, the third experiment showed a V-shaped priming effect when the target was within the subitizing range, but this effect was primarily based on identity priming (e.g. target: 3, prime: 3, t-p distance: 0), and therefore did not provide conclusive evidence for high-level numerical processing under CFS. To sum up, when tested properly, no conclusive evidence for non-conscious priming from non-symbolic numerosities was observed (Hesselmann, Darcy, Sterzer, & Knops, 2015). Rather - in line with the proposition of Kunde and colleagues (Kiesel, Kunde, & Hoffmann, 2008) - the observed priming effects (Naccache & Dehaene, 2001a, 2001b) were due to a “match with pre-specified cognitive action-trigger conditions” (p. 311; Kiesel et al., 2008). Importantly, while being based on semantic categorization the assumed action-trigger conditions do not necessarily require a semantic elaboration of the stimulus at hand.

These results cast some doubt on the automatic processing of non-symbolic numerical information.

Subitizing

An important limit to the ANS is the number range. While the discrimination performance for numerosities remains constant over a large numerical range and can readily be described by the Weber-Fechner law, very low numerosities diverge systematically from this behavior. Humans are extremely precise in enumerating sets that comprise between one three or four objects only. In this numerical range, accuracy is close to 100% and reaction times do not vary as a function of number of items. This capacity to arrive at rapid and exact numerical judgments for sets with few items (usually ≤ 4) is referred to as subitizing (Kaufman, Lord, Reese, & Volkmann, 1949). Subitizing was initially thought to invoke the same mechanisms as numerosity estimation (Ross, 2003) or based on a pre-attentive mechanism that assigns a Finger of INSTantiation (i.e. a sort of index or pointer) to individual (or grouped) items until the limited capacity of FINSTs is depleted (Trick & Pylyshyn, 1994). Recent evidence, however, suggests that subitizing does not rely on the same mechanisms as estimation but rather is

associated with a flexible, attention-based individuation mechanism. Revkin and colleagues (Revkin, Piazza, Izard, Cohen, & Dehaene, 2008) tested naming performance for all numerosities between 1 and 8 on one hand and for decade numerosities between 10 and 80 (10, 20, 30, etc.) on the other. Results showed that the variation of naming performance (reaction times, error rates, variation coefficient) differed dramatically between these number ranges, despite comparable discrimination difficulty within sets (i.e. the ratio between items 1 through 8 is equal to the ratio between items between 10 and 80). This speaks against the idea that a single set of cognitive mechanisms guides performance in both number ranges. A second line of research shows that subitizing and estimation are differentially influenced by the modulation of available processing resources. While subitizing was largely unaffected by a rivalling concurrent task, estimation performance declined under dual-task conditions (Anobile, Turi, Cicchini, & Burr, 2012; Burr, Turi, & Anobile, 2010). A series of recent studies suggests that subitizing capacity (i.e. the number of items that can immediately and effortlessly enumerated) is functionally associated with working memory limits (Melcher & Piazza, 2011; Piazza et al., 2011). In contrast to the FINST approach that assumes a fixed limit of available pointers (Trick & Pylyshyn, 1994), Melcher and Piazza assume that the visual system can flexibly assign a variable amount of limited resources to individual items – depending on task demands and the saliency of the items (Franconeri et al., 2013; Todd & Marois, 2004). In line with this assumption, individual visuo-spatial working memory capacity was positively correlated with subitizing range (Piazza et al., 2011). By manipulating task demands, Melcher and Piazza were able to influence the amount of resources that were assigned to individual items. In tasks that required a precise and rich representation, items were assigned more resources compared to situations in which coding the mere presence or absence of an items sufficed. Subitizing was reduced when little resources were left compared to situations where task demands were low, hence leaving more resources for subitizing. Using fMRI, we were able to demonstrate that posterior superior parietal cortex contains topographic information of individuated objects that the visual system can flexibly exploit for the current goals (Knops, Piazza, Sengupta, Eger, & Melcher, 2014). This attention-grabbing mechanism provides relevant information for visual short term memory, object tracking, grasping, or the enumeration of sets of objects. Depending on the level of detail needed for the current task, the amount of attentional resources that are dedicated to individual items varies and gives rise to different neural response profiles with varying numerosity. We measured brain activation of adult subjects performing either a visual short-term memory (vSTM) task consisting of holding in mind precise information about the orientation and position of a variable number of items, or an enumeration task consisting of assessing the number of items in those sets. We show that task-specific capacity limits (three to four items in enumeration and two to three in vSTM) are neurally reflected in the activity of the posterior parietal cortex (PPC): an identical set of voxels

in this region, commonly activated during the two tasks, changed its overall response profile reflecting task-specific capacity limitations. For the enumeration task, we were looking for voxels with a response profile that would parallel the behavioral results, that is, voxels that did not exhibit an increase of activation for low numerosities (n≤3) but a parametric increase in activation for higher numerosities (n >3), equivalent to an exponential function. For the vSTM task, we traced voxels that showed a complementary response profile with an increase of activation for lower numerosities (n≤3), reaching a plateau for higher numerosities (n≥3), equivalent to the inverse of an exponential function. Figure 7 shows the resulting activated networks projected onto an inflated brain template using the Human PALS (population-average landmark and surface-based)-B12 Atlas (Van Essen, 2005; Van Essen & Dierker, 2007) implemented in Caret software (Van Essen et al., 2001).

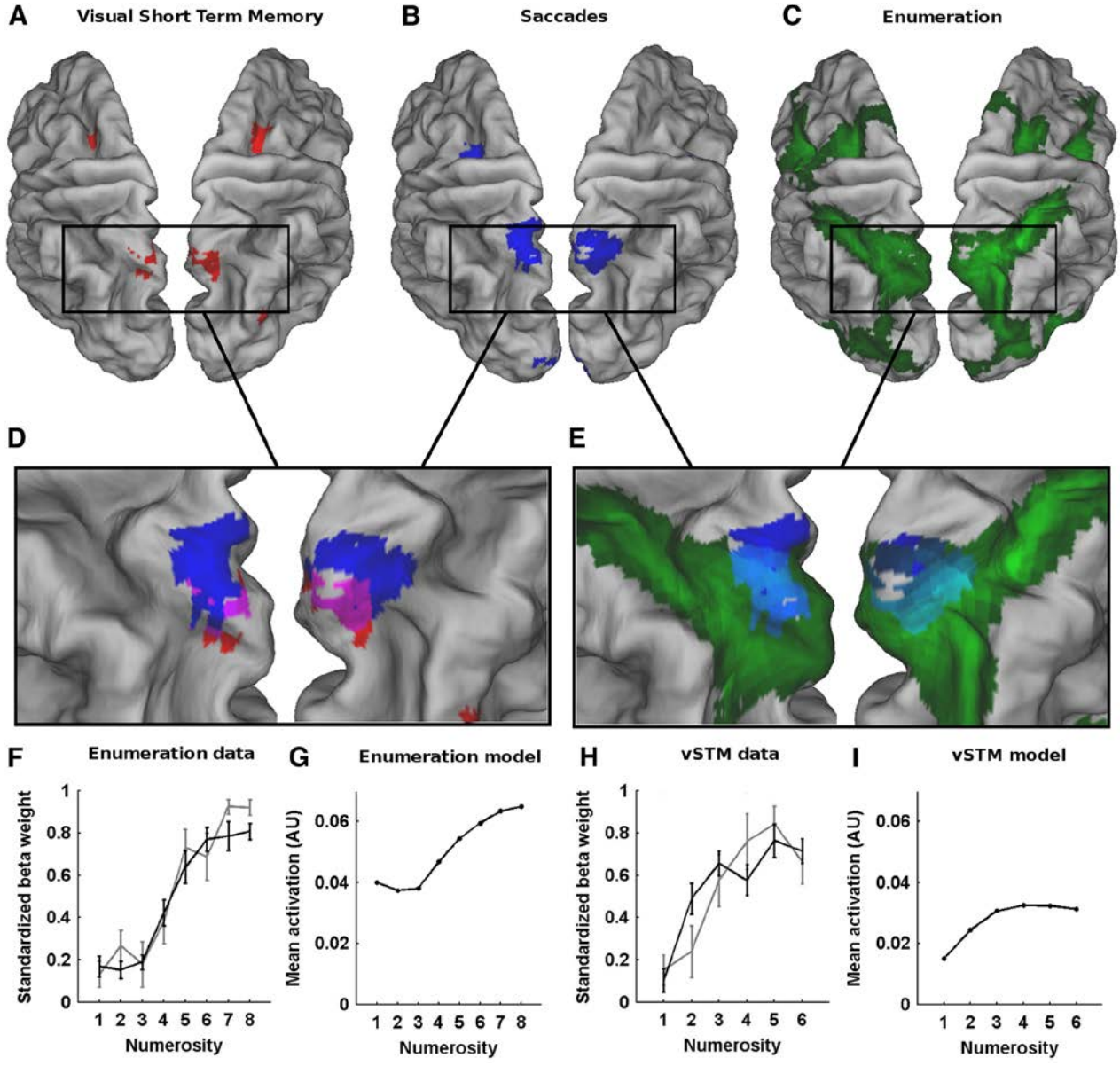

Figure 7: Brain activation results. A, Brain regions exhibiting a vSTM profile. B, Results of the saccades localizer. C, Brain regions with a subitizing profile. All random-effects contrasts

projected onto the top view of left and right hemispheres of an inflated brain template, thresholded at p<0.05 (FDR-corrected) except saccades localizer (p0.005, uncorrected). D, E, Enlarged view of the overlapping activation in PPC. D, Overlap (purple) between saccades localizer (blue) and vSTM (red) activation. E, Overlap (turquoise) between saccades localizer (blue) and enumeration (green) activation. F–I, Empirically observed and computational model activation profiles in the enumeration task (F, G, respectively) and the vSTM task (H, I, respectively) expressed as standardized ß weights (data) and arbitrary units (model). Empirical activation profile for enumeration is based on voxels that have been identified by the overlap between vSTM and saccades. The profile for vSTM is based on voxels that have been identified by the overlap between enumeration and saccades. Results from the main and control experiments are shown in black and gray, respectively. Error bars depict SEM.

vSTM (Figure 7A) activated bilateral precentral regions (frontal eye fields), superior parietal cortex and occipital cortex. Figure 7B shows the activations elicited by the saccades localizer task, consisting mainly of superior parietal and occipital regions. Enumeration (Figure 7C) activated a large network of frontal, precentral, and parietal regions extending into the occipital cortex. Virtually identical brain regions were obtained when using regressor profiles with lower inflection points of 3 and 2, better matching the empirically observed profiles for enumeration and vSTM, respectively.

To demonstrate the flexible change of response profile due to the representational precision the task at hand requires, we identified voxels in superior parietal cortex that were activated by the saccades localizer and whose response profile would match the presumed profile for vSTM and enumeration, respectively. Then, in a second step, we plotted the response profile of those voxels in the other task. That is, for vSTM voxels, we plotted the response in the enumeration task. For enumeration voxels, we plotted the response in the vSTM task. The response profile changed completely as a function of the specific task at hand. Voxels that paralleled the behavioral profile in the enumeration task changed their profile in the context of the vSTM task and vice versa. In both cases, the ß values varied significantly with numerosity. To statistically validate that voxels in PPC changed their response profile with task requirements, we analyzed the data points that were common to both tasks (i.e., numerosities 1–6) by fitting a log-linear function (Anobile, Cicchini, & Burr, 2012) to the ß weights in the context of both tasks. The function of the form

$$Y = a\left((1-\lambda)N + \lambda\,\frac{N_{max}}{\ln\ N_{max}}\ln N\right)$$

where Y denotes the beta weight and N the number of items in the display comprises a linear (a) and a logarithmic component (λ) that will be positive if the response profile is (logarithmically) compressed (as should be the case in the vSTM task) and negative if the

response profile takes an exponential form (as should be the case in the enumeration task for numerosities 1–6). We fitted this function to the individual mean ß weights of the voxels that were identified with the above described approach in both tasks (*n*=35 voxels from the vSTM task and *n*=51 voxels from the enumeration task) and compared the λ parameters obtained for both tasks. Indeed, we observed positive λ parameters for vSTM (mean λ = 1.05) and negative λ parameters (mean λ = -0.88) for enumeration (both significantly different from zero).

This behavior speaks for the presence of a flexible saliency/priority map where a small number of items can be presented with high precision with minimal noise to allow for rich encoding of stimulus features, such as orientation and spatial position, that were required in our vSTM context. With lower precision, more items would be represented at the cost of lower feature resolution, albeit sufficient for mere enumeration of items in a given set (Melcher & Piazza, 2011). Such map architecture has a number of advantages. First, it provides a way to account for both evidence of discrete representations and also the fact that capacity limits change across context and task (Franconeri et al., 2013; Melcher & Piazza, 2011), providing a way forward from debates about slots versus resources (Franconeri et al., 2013). Second, these maps are biologically plausible models of the well-studied behavior of neurons in PPC areas, such as the lateral intraparietal sulcus (Gottlieb, 2007).

Results from electrophysiological studies converge on the idea that subitizing crucially depends on “an attention-based individuation mechanism that binds specific features to locations and provides a stable representation of a limited set of relevant objects (Mazza & Caramazza, 2015).”

In sum, subitizing needs to be separated from estimation because it is characterized by different behavioral performance profiles and invokes different cognitive mechanisms. Subitizing recruits parietal circuits whose flexible behavior is in line with the idea of a saliency map architecture.

Implications

The findings reported here challenge several crucial assumptions that have been formulated for the ANS.

First, these findings challenge a crucial assumption of the previously described computational model by Dehaene and Changeux (1993). Dehaene and Changeux assume that the output of the location maps is normalized, meaning that only the mere presence or absence of an item is coded – irrespective of its size, luminance, saliency or priority. This output is passed on to a summation instance where activity increases monotonically with the number of items. Our results call for a change of this model to implement the salience map

instance. Based on my findings, numerosity information follows a saliency-modulated summation coding at this instance. A possible hierarchy of processing steps is illustrated in Figure 8.

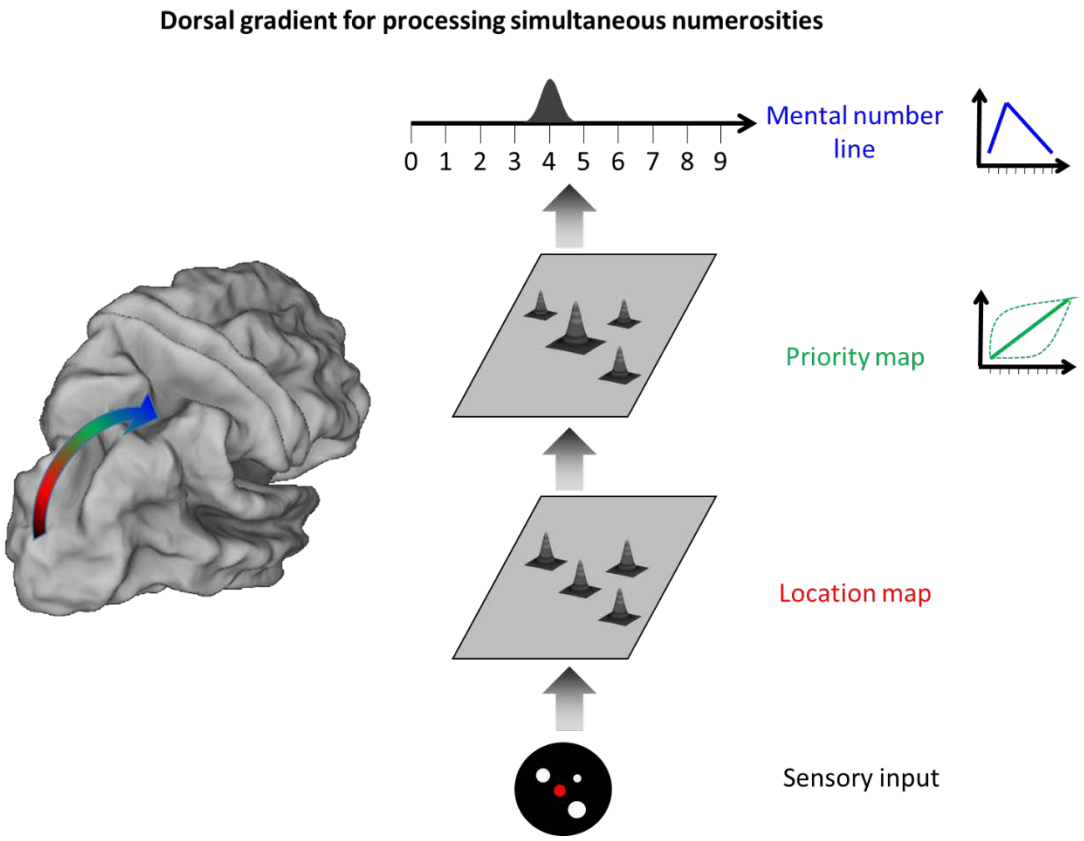

Figure 8: Schematic model of dorsal brain regions involved in processing simultaneous numerosities. Sensory input is fed forward to regions in occipital cortex that are characterized by a retinotopic organization and represent the locations of objects on the visual scene. At the next level along the gradient, the output of the location map is integrated with bottom-up saliency (e.g. brightness, contrast, etc.) and top-down priority (e.g. relevance in the current task) resulting in a priority map. In contrast to previous models, the activation of these areas is not a linear function of number (summation coding) but flexibly modulated by priority. This information is fed forward to an instance that represents numerosity information in a place coding scheme and corresponds to the neural instantiation the number line.

A second implication of my findings concerns the assumption that parietal cortex automatically represents numerosity irrespective of input modality or mode. Instead, our findings suggest two modifications. First, numerosity information does not appear to be automatically coded under all circumstances. Non-consciously presented numerosity information does not undergo semantic elaboration – at least not to a degree that would influence behavior in a priming study. Second, the necessity to utilize the result of a semantic

elaboration of consciously perceived numerosity appears to be an often-overlooked prerequisite for decodable numerosity information in parietal cortex. Again, we may potentially conceive of response selection and preparation as a process that merely amplifies numerosity codes in parietal cortex that remain undetectable for fMRI otherwise. Further studies are required to further specify the role of response-related processes for detecting numerosity information in parietal cortex.

Third, the notion of mode-independency of the parietal numerosity representation needs to be reconsidered. Sequential numerosity stimuli did not evoke parietal activity. This is in striking contrast with findings from neurophysiology and demonstrates the importance of conducting carefully planned experiments to empirically validate findings in other species or from other techniques in humans. Beyond this general notion, our finding may be explained by the overall architecture and functional role of parietal cortex. That is, one may also wonder whether the parietal functional architecture is suited to sequential stimuli. Parietal regions exhibit a map-like architecture and guide visual perception by integrating bottom-up salience information, top-down strategic goals, and reward experiences into a single map that indicates the relative priority of individual items in a visual scene (Bisley & Goldberg, 2003; Franconeri et al., 2013; Roggeman, Fias, & Verguts, 2010). Inter-item competition for resources limit performance across cognitive domains such as enumeration and visual short term memory (Knops, Piazza, et al., 2014). Therefore, repeatedly presenting stimuli at one location might not sufficiently activate a visual system built to integrate item information over space rather than over time. The exact mechanisms of the neural coding of sequential numerosities remains elusive for the moment.

Core Processes Underlying Mental Arithmetic

[Publications on this topic:

- Glaser, M. & Knops, A. (2020). When adding is right – temporal order judgments reveal spatial attention shifts during two-digit mental arithmetic. *Quarterly Journal of Experimental Psychology, 73*(7), 1115–1132.
- Didino, D., Pinheiro-Chagas, P., Wood, G., & Knops, A. (2019). Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect. *Frontiers in Psychology 6*;10:160. doi: 10.3389/fpsyg.2019.00160.
- Pinheiro-Chagas, P., Didino, D. Haase, V. G., Wood, G. & Knops, A. (2018). The developmental trajectory of the operational momentum effect, *Frontiers in Psychology – Developmental Psychology, 9*:1062. doi: 10.3389/fpsyg.2018.01062.
- Katz, C. Hoesterey, H. & Knops, A. (2017). A role for attentional reorienting during approximate multiplication and division. *Journal of Numerical Cognition, 3*(2), 246 – 269.
- Pinheiro-Chargas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Startling-Alves, I., Willmes, K., & Haase, V. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One. 9*(11): e111155. doi:10.1371/journal.pone.0111155
- Katz, C. & Knops, A. (2014). Operational momentum in approximate multiplication and division? *PLoS One*; 9(8):e104777. doi: 10.1371/journal.pone.0104777.
- Knops, A., Thirion, B., Hubbard, E.M., Michel, V. & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science, 324*(5934):1583-5.
- Knops, A., Viarouge A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and non-symbolic calculation: Evidence from the operational momentum effect. *Attention, Perception & Psychophysics, 71*(4), 803-821.]

Mental arithmetic is a summary term describing cognitively very distinct processes. Staring with basic skills such as solving simple addition, subtraction, multiplication and division problems, we acquire more abstract mathematical skills such as solving equations with one or more unknowns (2x + 3y = 23), for example. Here, I will concentrate on basic mental arithmetic processes.

I will argue that abstract mathematical competencies are tightly linked to the aforementioned ANS. I will argue that abstract arithmetic competencies are grounded in the ANS. Both, the ANS and arithmetic processes rely on overlapping brain circuits (Arsalidou & Taylor, 2011). Parietal cortex (PC) appears to be of particular importance. I will demonstrate that this notion goes beyond the report of mere overlap between activations in functional imaging studies. Instead, the spatial pattern of activation that is elicited by different tasks in parietal and prefrontal areas manifests a high degree of ordered similarity and possesses predictive value across cognitive domains. Low-level perceptual processes such as saccades lead to spatial

patterns of activation in posterior superior parietal lobe (PSPL) that are predictive of patterns during abstract approximate calculation processes (Knops, Thirion, Hubbard, Michel, & Dehaene, 2009). This is interpreted in terms of cultural recycling of cortical maps for cognitive purposes that go well beyond the evolutionary scope of a given region. The proposal is that human mathematics builds from foundational concepts (space, time, and number) by progressively co-opting cortical areas whose prior organization fits with the cultural need.

Mental arithmetic recruits spatial coordinate transformations

Building on the notion of a spatial magnitude representation, mental arithmetic has been conceived of as an interaction between verbally mediated fact retrieval and spatial displacements on the MNL (Hubbard, Piazza, Pinel, & Dehaene, 2005), mediated by PPC. Indeed, the population code representing the head centered position of an object in the visual field is an optimal probabilistic additive combination of the eye-centered coordinates and the eye position coordinates (Beck et al., 2011). Since these coordinates are primarily coded in PPC, these areas thus host basic perceptual neural circuits that are capable of integrating the incoming output of two separate neuronal populations in an additive way. Consequently, neuronal recycling may co-opt these circuits for mental arithmetic, where two numerical pieces of information need to be integrated and may correspond to the position on an internal spatial magnitude representation (MNL). Indeed, we recently used a dual-task paradigm to demonstrate a symmetric interference between mental arithmetic (subtraction and multiplication) and both visuo-spatial and phonological working memory (Cavdaroglu & Knops, 2015). Further evidence for the recruitment of spatial processes comes from the operational momentum (OM) effect. OM describes a systematic bias in estimating the outcomes of simple addition and subtraction problems. Outcomes of addition problems are overestimated while outcomes of subtraction problems are underestimated (Knops, Viarouge, et al., 2009; McCrink et al., 2007). The OM has been observed with addition and subtraction, both in symbolic (i.e. Arabic digits) and non-symbolic notation (Knops, Viarouge, et al., 2009; Pinhas & Fischer, 2008). OM has been interpreted as a prime example of neuronal recycling since it is considered a consequence of the reuse of cortical circuits that have evolved for spatial processing during the course of mental arithmetic (Fischer & Shaki, 2014). In line with this we demonstrated that brain activity elicited by left- and rightward saccades, which are accompanied by respective attentional shifts, predict whether participants were performing centrally-presented addition or subtraction problems (Knops, Thirion, et al., 2009). We found that addition problems corresponded to the neural activity associated with right-ward saccades, presumably since participants oriented towards larger numbers on the right side of the MNL. More evidence for attentional contribution to mental arithmetic comes from a recent report that found single-digit addition and subtraction problems to be solved faster when the second operand was presented

on the right or the left side of a central fixation mark, respectively (Mathieu, Gourjon, Couderc, Thevenot, & Prado, 2016). Interestingly, no such bias was observed with multiplication (Katz & Knops, 2014), which is consistent with the idea that unlike procedural mental calculation arithmetic fact retrieval is mediated by recall from long term memory (Verguts & Fias, 2005). Recent results from my lab more specifically point to the reorienting of attention as the major component driving the OM (Knops, Hoesterey, & Katz, 2017), again implying a major contribution of inhibitory processes to mental arithmetic. Since in this task we used multiplication and division, this may point to the inhibition of concurrently selected targets (i.e. related results) that need to be inhibited in favor of the correct target (i.e. the correct solution).

While these results imply that attention contributes to mental arithmetic, only establishing the hall mark effects from the attention domain would provide unanimous evidence for this notion. Attention has been found to modulate the perception at very low levels of the visual processing hierarchy. For example, Carrasco and colleagues repeatedly observed increased contrast sensitivity at attended locations at the expense of decreased contrast sensitivity at unattended locations (Carrasco, 2011). Similar results have been observed for other core dimensions such as spatial frequency (Abrams, Barbot, & Carrasco, 2010), brightness (Tse, 2005), speed (Turatto, Vescovi, & Valsecchi, 2007) and object size (Anton-Erxleben, Henrich, & Treue, 2007). The underlying neural mechanism has been a topic of lively debate. Several, not necessarily mutually exclusive mechanisms have been proposed. First, a multiplicative increase of neuronal firing rates has been proposed (McAdams & Maunsell, 1999). Second, a contrast gain has been discussed, whereby attention increases the effective contrast of stimuli by a scale factor (X. Li & Basso, 2008; Martínez-Trujillo & Treue, 2002). Finally, electrophysiological as well as human neuroimaging studies imply that attention yields a shift in baseline activity irrespective of contrast (Buracas & Boynton, 2007; Reynolds, Pasternak, & Desimone, 2000).

In a recent study, we used a temporal order judgment (TOJ) task to measure deflections of spatial attention (Glaser & Knops, 2020). This paradigm has already been used in the field of numerical cognition with single digits (Casarotti, Michielin, Zorzi, & Umiltà, 2007) and has the advantage of assessing spatial attention free from motor processes because it requires un-speeded verbal responses. It entailed the visual presentation of two lateralized targets with a varying stimulus onset asynchrony (SOAs; see Figure 13). Participants had to indicate which stimulus was presented first (right first vs. left first). The response to the TOJ task had to be given verbally to avoid effects of stimulus-response compatibility that come with button press reactions. Based on these responses, the point of subjective simultaneity (PSS) could be calculated. It constitutes the time lag between the lateralized stimuli that is needed for them to be perceived as being presented simultaneously. It illustrates the estimated point in time where

the probability of saying that the right stimulus appeared first equals the probability of saying that the left stimulus appeared first (i.e. 50 % each).

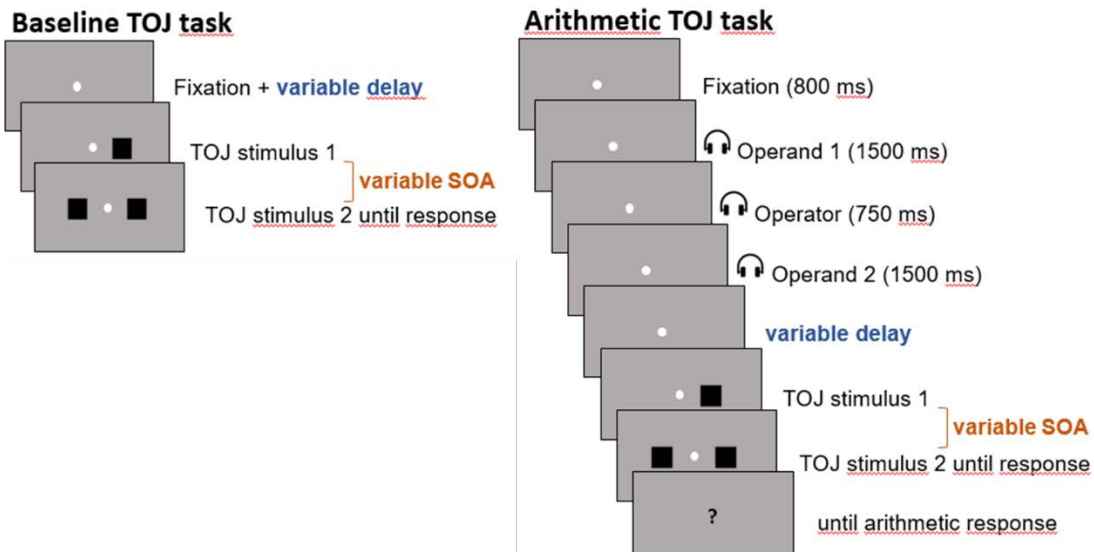

Figure 9: Task sequence and timing of a single trial of the baseline TOJ task (left) and the arithmetic TOJ task (right)

In two experiments, we used a TOJ paradigm to investigate whether we can observe attentional shifts to the right and left during the calculation phase of complex (two-digit) addition and subtraction processing, respectively. While Experiment 1 included both addition and subtraction problems with carrying/borrowing, this stimulus feature was not meant to serve as an experimental factor, resulting in too few trials to analyze its impact on performance. In Experiment 2, we intentionally varied the carry property of the arithmetic task to assess how task difficulty modulates the effects (carry ≙ difficult vs. non-carry ≙ easy). Finally, in both experiments we varied the delay between the arithmetic problem presentation and the TOJ task to investigate when arithmetically induced attentional shifts occur and how long they would persist.

Crucially, both experiments revealed (Figure 14) that addition processing shifts attention more rightward compared to the baseline task which corroborates earlier findings (S. X. Li & Cai, 2014; Liu et al., 2017; Masson et al., 2018). The present study adds to the existing literature in two important ways. First, we observed attentional shifts during the calculation phase of arithmetic processing which indicates that attention shifts are part of the calculation process. Second, the present study could show that complex, two-digit addition problems are sufficient to induce attentional shifts to the right during addition processing indicating that visuo-spatial attention mechanisms are recruited during the processing of complex arithmetic problems.

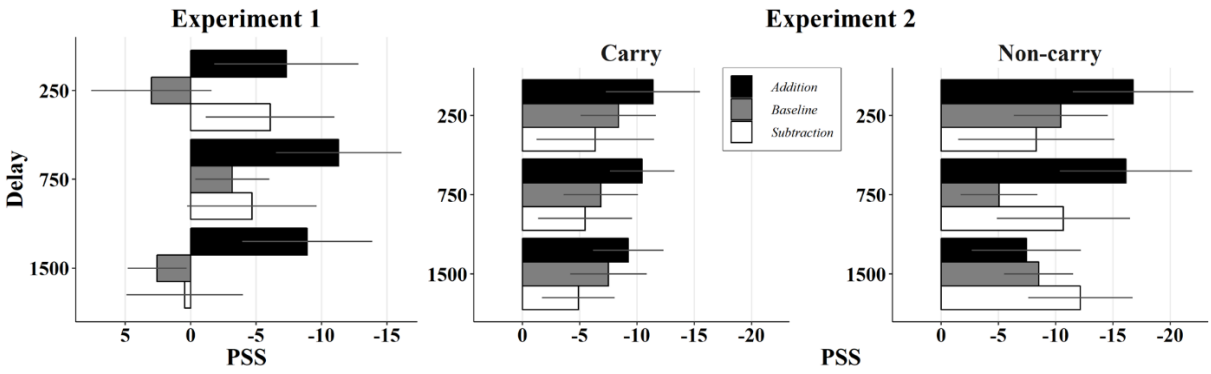

Figure 10: Mean PSSs (points of subjective simultaneity) for each operation and delay condition of Experiment 1 (left panel) and Experiment 2 (right panel). Error bars represent the standard error of the mean (SEM). PSS values constitute the time lags between the lateralized TOJ stimuli that are needed for them to be perceived as being presented simultaneously. Positive PSSs indicate attention in the left visual field and negative PSSs indicate attention in the right visual field.

The second main observation of the present study is the absence of spatial attention shifts to the left for subtraction (compared to the baseline). This finding is reminiscent of a study by Masson and colleagues (Masson et al., 2017) who found that subtraction processing was unaffected by leftward or rightward optokinetic stimulation. Furthermore, the fMRI classifier trained on rightward and leftward saccades (accompanied by left- and rightward shifts of attention) of Knops and colleagues (Knops, Thirion, et al., 2009) was unable to correctly classify subtraction trials. This finding issues a challenge to all existing theories as it cannot be explained by spatial or non-spatial accounts. How can we explain this absence of an effect for subtraction?

One possibility is that subtraction tasks might be solved via addition strategies as it has already been suggested by some (Masson et al., 2017). This might have led to attentional shifts to the right which masked operation-congruent shifts to the left. Another possible explanation for the attenuated spatial bias in subtraction is that the increased working memory load (i.e. cognitive load) in subtraction processing is associated with reduced spatial processing which might have led to less spatial biases.

Finally, this study also set out to investigate the time course of the attentional shifts. In this regard, the analysis of the arithmetic subset (only addition and subtraction task) in Experiment 1 revealed that independent of operation, attention shifts more rightward in the second delay of 750 ms than in the other delays of 250 and 1,500 ms. In Experiment 2, it revealed a tendency for decreasing attention shifts over time and significant differences between the slopes of these

linear trajectories over delay of the two arithmetic operations. Hence, results were not conclusive regarding the time course of attentional shifts. If, as proposed above, approximate processes drive the attentional shift to the right for addition, then one would expect biases early during arithmetic processing and a decrease during later stages of the calculation process. However, our study does not provide conclusive evidence for (or against) this hypothesis.

Further research is needed to clarify whether this is due to the involvement of addition strategies or due to an increased working memory load that attenuates spatial processing in subtraction.

In sum, the existing results point to the contribution of basic inhibitory, perceptual and attentional mechanisms that have evolved for spatial coding of visual space and guidance of eye movements to basic mental arithmetic.

Developmental trajectory of the Operational Momentum Effect

Three mutually not exclusive mechanisms have been proposed to explain the OM effect: attentional shift account, heuristic account, and compression account. However, none of them aimed to describe how this effect changes over development. Evidence shows that the neural network that support mental calculation undergoes substantial functional changes during development and reaches an adult-like configuration only during adolescence (Arsalidou, Pawliw-Levac, Sadeghi, & Pascual-Leone, 2018; Peters & De Smedt, 2018; Rosenberg-Lee, Barth, & Menon, 2011; Soltanlou et al., 2017; Soltanlou et al., 2018). Therefore, in order to fully understand the cognitive mechanisms lying at the core of the OM effect it is important to measure its developmental dynamics and to evaluate whether the current accounts are able to explain these age-related changes. In what follows, we introduce these accounts of the OM effect and discuss the developmental trajectories predicted by each of them.

It has been proposed that mental calculation is grounded in neural circuits that originally evolved for processing visuospatial information (Anderson, 2010; Dehaene & Cohen, 2007; Knops, Thirion, et al., 2009). Moreover, various evidence supports the existence of a functional relationship between visuospatial attention (i.e., shift of spatial attention) and mental calculation (Masson & Pesenti, 2014, 2016; Masson, Pesenti, & Dormal, 2013; Masson et al., 2017; Mathieu et al., 2016). In line with these studies, the *attentional shift account* proposes that the OM effect is the result of this functional relationship (Knops, Thirion, et al., 2009; Knops, Viarouge, et al., 2009; McCrink et al., 2007). The central assumption of the attentional shift account hypothesizes that non-symbolic addition and subtraction are implemented by shifting spatial attention on a spatially oriented MNL. During approximate calculation, the first

operand is mapped on the MNL, then the attentional focus shifts from the current position (i.e., the point corresponding to the magnitude of the first operand) to a new position (i.e., the point corresponding to the magnitude of the result) by a distance corresponding to the magnitude of the second operand. The OM effect is produced by a bias in the attentional shift, that is the attentional focus moves too far along the MNL in the direction of the operation, generating an overestimation and an underestimation of the result of addition and subtraction, respectively. Strong evidence for the hypothesis that visuospatial attention is co-opted during mental calculation is provided by the overlap in the posterior superior parietal lobule (PSPL) of the neural activity associated with left/right saccades (i.e., visuospatial orientation) and mental calculation.

McCrink and Wynn (McCrink & Wynn, 2009) proposed the *heuristic account* to explain the finding that the OM effect also affects performance in 9 months old infants. This account assumes that infants adopted a simple heuristic to solve the problems: “if adding, accept larger outcomes”, “if subtracting, accept smaller outcomes”. For addition, this heuristic approach might encourage infants to perceive larger outcomes as more plausible compared smaller ones, and vice versa for subtraction. Recently, McCrink and Hubbard (McCrink & Hubbard, 2017) interpreted the finding that the OM effect increased in adults when available attentional resources were limited by dividing attention between two concurrent tasks as further evidence for the heuristics account. However, the heuristic account and the attentional shift account are deeply intertwined and can be considered as a single mechanism (i.e., heuristics-via-spatial-shifts account), that is the heuristic decision results from the visuospatial system (McCrink & Hubbard, 2017). Therefore, we will only focus on the attentional shift account, assuming that the two accounts provide equivalent predictions.

The attentional shift account has been developed to explain the OM effect in adults. Therefore, no predictions or hypotheses were proposed regarding how the attentional shifts on the MNL that accompany addition and subtraction emerge and whether they undergo substantial changes during development. Here, we propose that formal schooling (i.e., acquiring arithmetical skills) could reinforce (or even contribute to develop) the idea that addition is related with shifts towards larger numbers and subtraction towards smaller numbers. Namely, although mental calculation might be implemented as an attentional shift on the MNL before formal schooling, repeated exposition to spatial-numerical associations (e.g. the number line) might consolidate a systematic movement direction during the acquisition of arithmetical skills. Moreover, the systematic association between operations and results (i.e., when adding, the result is always larger than both operands; when subtracting, the result is always smaller than the first operand) that children are exposed to could boost the attentional shift on the MNL. The influence of the attentional shift in the estimation of the result might increase with age and in turn a larger and more systematic bias would emerge.

Therefore, one may predict an increasing OM effect during childhood. Moreover, it is worth noting that the co-opting of visuospatial attention during mental calculation seems to increase with age. In fact, significant functional changes associated with the neural activity elicited by symbolic arithmetic problem-solving have been found between 2nd and 3rd graders, that is 7-9 years old children (Rosenberg-Lee et al., 2011). During the processing of symbolic arithmetic problems, 3rd grade children showed greater activity in brain regions related to visuospatial attentional processes (posterior parietal cortex: intraparietal sulcus; superior parietal lobule; angular gyrus) and high-order visual processing (ventral visual areas: lingual gyrus, right lateral occipital cortex, and right parahippocampal gyrus), compared to 2nd grade children.

The *compression account* has been proposed by McCrink and colleagues (McCrink et al., 2007) and deploys the logarithmic compression of the MNL to explain the OM effect. This compressed metric would generate a systematic operational bias in the direction of the operation due to the implementation of a linear arithmetic operation (i.e., addition or subtraction) on a logarithmically scaled mental representation. This mechanism acts in three steps. First, the operands are encoded as logarithmically compressed magnitudes on the MNL. Second, the logarithmic transformation is undone, which means that the operands are uncompressed to a linear scale. Third, the two uncompressed operands are added or subtracted. The OM effect results from the inaccuracy of the uncompression process. If the uncompression is ineffective the arithmetic operation is performed on logarithmic values and thus the generated outcome corresponds to an extreme overestimation or underestimation for addition and subtraction, respectively. If the uncompression is highly accurate the operation is performed on the linear scale, in which case the generated outcome corresponds (approximately) to the arithmetically correct result. A more plausible scenario is to assume that the actual degree of uncompression lies between these two extreme possibilities. An example can help describe this idea. If uncompression fails, adding two operands (e.g., 26 and 14) corresponds to adding their logarithmically compressed internal representation, that is $\log(26) \approx 3.26$ and $\log(14) \approx 2.64$, respectively. Since adding the logarithm of two numbers is equivalent to multiplying their linear values, the system generates an extreme overestimation of the correct result: $\log(26) + \log(14) \approx 5.9$, which in linear scale corresponds to $e^{5.9} \approx 26 \times 14 \approx 364$. However, the actual approximate addition performed by the system is much more accurate (see for example (McCrink et al., 2007)), and thus the uncompression is to some extent carried out and the generated outcome is much closer to the correct result. The same reasoning is valid to explain the mechanisms underpinning the underestimation of subtraction outcomes.

What developmental trajectory of the OM effect is expected according to the compression account? This account focuses on the logarithmic compression of the MNL. A large body of evidence suggests that the representational metric of the MNL metric shifts from a logarithmic

to a linear scale during childhood (Booth & Siegler, 2006, 2008; Opfer & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003) but for a different interpretation see (Barth & Paladino, 2011). The logarithmic-to-linear shift of the MNL implies that the compression of this magnitude representation decreases with age and probably with accumulation of experience in formal mathematics teaching. Therefore, the uncompression of the operands, performed before the approximate mental calculation, starts from a highly logarithmic scale in young children and from a more linear scale in adults. The degree of uncompression required to generate an accurate outcome is thus greater in young children and this in turn could lead to a stronger OM effect. The compression account therefore predicts that the size of the OM effect is higher in young children and decreases with age to reach an adult-like pattern in older children. It is worth noting that, as discussed below, the inverse OM effect (i.e., overestimation of subtraction problems) found in 6/7 years old children (Knops, Zitzmann, & McCrink, 2013) already provides evidence against this account.

We investigated the development of the OM effect in a group of 162 children from 8 to 12 years old. Participants had to select among five response alternatives the correct result of approximate addition and subtraction problems. Response alternatives were simultaneously presented on the screen at different locations.

While no response bias was observed for the youngest age group, children aged 9 and older showed a clear response bias, that is they showed an OM effect (see Figure 15). Interestingly, the OM effect monotonically increased with age. The increase of the OM effect was accompanied by an increase in overall accuracy. That is, while younger children made more and non-systematic errors, older children made less but systematic errors.

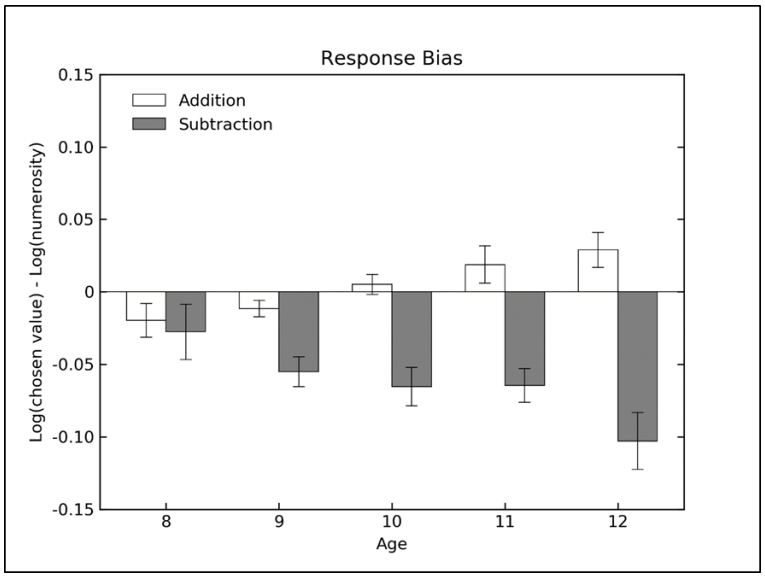

Figure 11: Mean response bias (i.e., difference between the logarithm of the chosen response and the logarithm of the correct outcome) as a function of age and operation (addition in white and subtraction in grey). Error bars represent the standard error of the mean.

This monotonous increase of approximate calculation bias with age is not predicted by the compression account (i.e., linear calculation performed on a compressed code). The attentional shift account however provides a possible explanation of these results based on the functional relationship between visuospatial attention and mental calculation and on the influence of formal schooling. We propose that the acquisition of arithmetical skills could reinforce the systematic reliance on the spatial mental number line and attentional mechanisms that mediate the displacement along this metric. Our results provide a step in the understanding of the mechanisms underlying approximate calculation and an important empirical constraint for current accounts on the origin of the OM effect.

Our results were criticized with respect to theoretical and methodological considerations by Fischer, Miklashevsky and Shaki (Fischer, Miklashevsky, & Shaki, 2018, henceforth FM&S). We rebutted these considerations in a commentary (Didino, Pinheiro-Chagas, Wood, & Knops, 2019).

First, FM&S allege that the compression account of the OM served as a strawman hypothesis because it had been invalidated by their previous experiments, where FM&S observed a stronger bias for zero problems (e.g. 6+0) compared to non-zero problems (Pinhas & Fischer, 2008). FM&S argue that this suffices to invalidate the compression account “because the logarithm of zero is not defined”. This argumentation is flawed because FM&S mix up logarithm as a mathematical function (not defined for zero, indeed) with logarithm as a model (coding scheme) to describe the compressed internal scale of the representation of magnitudes (Harvey et al., 2013; Nieder & Miller, 2003). In the latter case, the logarithmic

function is used as mathematical approximation of the relation between external physical magnitude and its internal representation. However, it makes no sense to assume that cortical circuits actually compute the faithful “mathematical log transformation” of a given sensory information. The intensity of external physical stimuli is internally represented via non-linear spatio-temporal neural codes (e.g. rate code, population code). Basing their criticism on the restriction of the mathematical definition of the logarithm to positive real numbers, FM&S conflate the mathematical definition with the neural and cognitive representation of magnitudes. Moreover, even assuming that the cognitive system would actually be bound to this particular mathematical formulation of the relation between physical stimulus magnitude and sensation, another framework has been put forward that does define a mathematical solution of zero magnitudes. Stevens's power function (with positive real exponents smaller than 1) can provide identical predictions and is defined for zero. In sum, the fact that “the logarithm of zero is not defined” does not invalidate the compression account.

Second, FM&S ask “how does [the attentional shift] account explain larger OM with zero problems?” In Pinhas and Fischer’s task (2008), zero problems only required to map a number (the first operand) onto a labelled line, since these problems are solved by means of rules (i.e., N+0=N, N−0=N) rather than mental calculation (Butterworth, Zorzi, Girelli, & Jonckheere, 2001; Campbell & Metcalfe, 2007). Therefore, FM&S’s question is not valid because its premise (i.e., zero-problems produce a OM) is not valid. Without providing behavioural and neurophysiological evidence that zero and non-zero problems are solved with the same strategies, FM&S cannot assume that their biases arise from the same underlying mechanisms. The attentional shift account aims to describe how the mental manipulation of (at least) two numerosities produces a mis-estimation of the outcome, rather than biases in zero problems, which reflect a different phenomenon.

Third, we argued that the attentional shift account and the heuristic account provide equivalent predictions. Fischer and colleagues criticise this by stating that it is in conflict with results from McCrink and Hubbard and cite: “… the use of heuristics is generally increased when attention is decreased” (McCrink & Hubbard, 2017). Our interpretation of McCrink and Hubbard’s manuscript was based on the idea that these two accounts “are actually so deeply intertwined that they are indistinguishable” (p. 240) and on the fact that McCrink and Hubbard’s findings “can be best described with a heuristics-via-spatial-shifts account” (p. 241).

Four, FM&S criticise that the downward (upward) movement of addends (subtrahends) would be inconsistent with “the vertical MNL” and ask “why […] operations along a horizontal MNL [were] primed with vertical movements?”. We argue that these movements actually mimic our daily experience: adding objects from the top into a box (downward movement) and subtracting them from inside a box to the top (upward movements). Any effect of this supposed

inconsistency between physical vertical movements of the operands and attentional movement on the MNL should have weakened, eliminated or even reversed the OM. Yet, we did not observe such interference. They also reasoned that the centre-to-top movement of the subtrahends "removed attention from the place of mentally simulating the outcome, thus impeding subtraction." First, this conclusion is inconsistent with findings from previous studies (McCrink et al., 2007; McCrink & Hubbard, 2017), where OM was observed despite subtrahends moving to the right (i.e., inconsistently with the horizontal MNL). Second, FM&S conflate mental simulation of addition and subtraction with attentional focus in external space. After all, the outcomes are estimated in the participants' minds – not in external space where no numerical information is present at that point in time.

Finally, the idea that in our previous studies "the normal ingredients of OM are dis-ordered or diluted" originates from the divergent definition of the OM. As originally defined by McCrink and colleagues (McCrink et al., 2007), the OM reflects a misestimation of the outcome of an arithmetic problem. In number-to-line mapping tasks, participants have to locate addition and subtraction outcomes on a labelled line (McCrink et al., 2007; Pinhas & Fischer, 2008) or to modify the length of a line proportionally to addition and subtraction outcomes (Shaki, Pinhas, & Fischer, 2018; Shaki, Sery, & Fischer, 2015). These paradigms do not measure outcome deviations, but rather they require an additional transformation process where the outcome is converted into another physical dimension (number to position or length). Both tasks can be subject to strategical (e.g., use of reference points; (Barth & Paladino, 2011; Slusser, Santiago, & Barth, 2013)) or procedural biases (e.g., perceptual hysteresis). Therefore, any observed biases may arise from the additional transformation process rather than the calculation process itself. Results from procedures that analyse only the final location on a labelled line (Pinhas & Fischer, 2008) or the length of a segment (Shaki et al., 2018; Shaki et al., 2015) must be interpreted cautiously because they are not measuring OM but biases that may well take place after the calculation process and have their origin in the transformation algorithm.

In sum, the operational momentum is joining other experimental findings on the idea that spatial attention contributes to mental arithmetic. This may be interpreted as neural recycling, reusing parietal circuits that evolved for eye movement planning and uses vector addition in the context of mental arithmetic.

Research plans

The studies that I am envisaging can be grouped in three integrated work packages (WP).

The first work package addresses the question how numerosities are perceived, processed and neurally represented. The second work package aims at elucidating the core processes of mental arithmetic and their reliance on different lower-level sensory-motor systems. The third work package is dedicated to the question how domain-general and domain-specific processes are interrelated and contribute to the development of numerical competencies.

WP 1 - Numerosity perception

Work package 1 aims at specifying the characteristics of spatial and temporal numerosity extraction.

Spatial and temporal enumeration

Experiment 1 aims at delineating the cortical and cognitive instances that are involved in the extraction of numerosity information from space. As introduced above, object-based approaches vie with indirect approaches. While the former postulate that numerosity extraction operates over individuated objects and stipulates the emergence of number-selective neuronal populations, the latter approach considers number a secondary dimension that is indirectly inferred, for example, by computing the ratio between low and high spatial frequencies.

Kanizsa-type illusory contours (ICs, see *Figure 16*) provide the spectator with the illusion of perceiving a contour defining a white bar which is physically absent. ICs will be used to orthogonally vary the number of items and the non-numerical visual features in a scene. ICs provide the experimenter with a tool to disentangle the number of objects in a scene from the number of occupied positions that give rise to summary statistics such as convex hull, density etc.

Participants will be presented with a prime and a target display. The prime informs about which object will be relevant in the target display and therefore modulates the amount of attentional resources dedicated to either ICs (i.e. the white bar) or the constituting but separate “pacman” objects. The task will be to enumerate the number of relevant items. A central notion of the object-based approaches of numerosity perception is the individuation of objects that recruits priority maps in PPC where items receive resources in accordance with their current priority. Hence, our task allows to flexibly define, within the same display, what counts as an object. As can be seen in *Figure 16*a and *Figure 16*b, by turning the aperture of the pacman objects we can manipulate the number of ICs without changing the number of occupied locations. That is, although the non-numerical features remain constant in those displays

object-based approach predicts different numerosity percepts accompanied by respective changes in BOLD signal. If one accepts that ICs count as objects, the density in *Figure 16*b is higher than in the *Figure 16*a. Hence, the indirect approach would predict higher number of perceived items than the object-based approach and higher activation in parietal cortex. The separate manipulation of the number of targets and the number of occupied positions allows us to independently track which brain areas adhere to which feature.

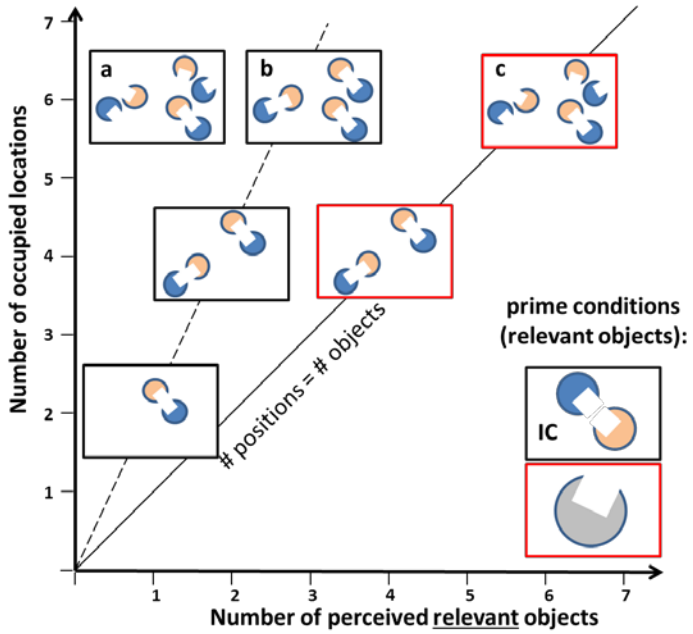

Figure 12: Relation between number of physically occupied positions in visual field and number of relevant objects. Object relevance is primed by presenting participants with symbols as in shown lower right corner, indicated here by colored frame. Red frame: enumerate ICs (c). Black frame: enumerate "pacmen" (a & b). Difference between iso-numeric line (bisecting line) and position-location ratio (dotted lines) can be manipulated by angle of aperture (cf. a & b vs. c).

Further experiments will use simultaneously and sequentially presented numerosities to further investigate the distinct processes underlying temporal and sequential enumeration. In a previous study from my lab, sequential numerosities were presented via repeatedly flashing centrally located disks on screen (Figure 5A-b). This conflates object location (centred in sequential vs. scattered for simultaneous) with the mode of enumeration (simultaneous vs. sequential). To disentangle these dimensions, participants will be presented with sequential numerosities that appear at different locations (Figure 5A-c). Contrasting this with a simultaneous dot patterns (Figure 5A-a) allows isolating the brain areas associated with accumulation of numerical information over time. MVPA will be used to probe (a) number sensitivity in different PPC ROIs and (b) generalisation between modes. Contrasting spatially

scattered (Figure 5A-c) and centred (Figure 5A-b) sequential numerosities allows isolating the brain areas dedicated to spatial integration. The hypothesis is that PPC is centrally engaged during spatial enumeration, while sequential enumeration hinges on primary sensory buffers in medial occipital and middle temporal areas and insular cortex, whose characteristics remain to be identified.

WP 2 - Core processes underlying mental arithmetic

Directionality and specificity of neuronal recycling

WP 2.1 aims at further delineating directionality of the number-space association which may have its origin in grounded, embodied, or situated factors.

Beyond the left-right oriented mental magnitude representation a bottom-to-top organization has been brought forward to explain a number of experimental findings (Winter, Matlock, Shaki, & Fischer, 2015). This is thought to be grounded in the physical law of gravity by which the ground level represents a natural reference point (Winter et al., 2015). Adding objects will lead to an increase in vertical direction and hence up is associated with more, down is associated with less. Consequently, in addition one should expect an association of up-ward and down-ward actions with addition and subtraction, respectively.

WP 2 will test this notion by extending into the vertical dimension previous experiments that only tested left-right associations with mental arithmetic. In line with previous experiments (Knops, Thirion, et al., 2009), participants will be presented with two tasks while lying in the MR scanner. First, participants execute saccadic eye movements along well-defined trajectories along eight axes (see Figure 17). The activity in PPC will be used to train a classifier to distinguish the axis along which participants executed saccadic eye movements. Second, participants will solve approximate additions and subtractions in Arabic notation. Without further training the classifier initially trained to differentiate saccade direction will be tested with activation patterns from the calculation task. Will the classifier identify subtractions as upward or rightward saccades?

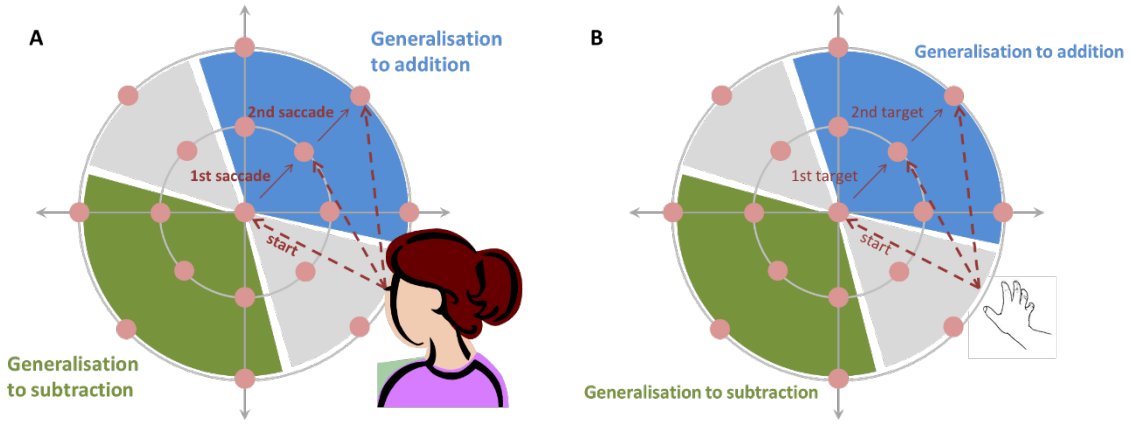

Figure 13: Cartoon depiction of the hypotheses for the generalisation from saccadic eye movements (A) or manual actions (B) space to arithmetic operations. A - Saccadic targets are shown as reddish dots. Participants fixate in predetermined order (first: inner circle, second: outer circle) the target positions, allowing for separate modelling of BOLD activity to saccades along the eight cardinal axes. First, a classifier will be trained to "predict" the saccadic eye movements. Then, without further training, we will test if the classifier generalises from any of the eight axes to the arithmetic operation (i.e. addition and subtraction). Blue and green shaded areas depict the hypothetical successful generalisation. For grey areas no prediction can be derived from the literature. B - An analogue experiment with reaching and pointing movements will test the specificity of the oculomotor system contribution to mental arithmetic.

Voxels with directional selectivity show contralateral preference (Leoné, Toni, & Medendorp, 2014). Voxels in left PPC show preference to saccades to the right while homolog voxels in the left hemisphere show preference for saccadic targets in the left visual field. If the number line is operating on an eye-centred reference frame, we should observe hemispheric modulation of the generalisation from saccades to arithmetic operations, because decoding accuracy improves with overall BOLD amplitude (Tong, Harrison, Dewey, & Kamitani, 2012). Consequently, generalisation from rightward saccades to addition should be best in left hemisphere while generalisation from leftward saccades to subtraction should be best in right hemisphere.

Is mental arithmetic limited to the re-use of oculomotor system?

WP 2.2 will test the association of hand-related parameter computation with mental arithmetic. Similar mechanisms as those involved in eye movement guide hand movements such as reaching and grasping, which rely on parietal circuits located at the posterior and anterior end of the intraparietal sulcus, respectively. Notably, the parietal reach region (PRR)

in posterior parietal cortex has been demonstrated to represent reaching targets in gaze-centered coordinates and to reflect plans for reaching movements (Andersen & Buneo, 2002; Buneo & Andersen, 2006). In contrast, neurons located in the anterior bank of the IPS extending into postcentral sulcus preferentially encode grasping movements (Konen, Mruczek, Montoya, & Kastner, 2013) in hand-centered coordinates (Iwamura, Iriki, & Tanaka, 1994; Taira, Mine, Georgopoulos, Murata, & Sakata, 1990). These neurons encode shape, size, and orientation of to-be-grasped objects (Murata, Gallese, Luppino, Kaseda, & Sakata, 2000). The planning of both saccadic eye movements and hand movements requires the representation and dynamic updating of spatial position information from several sources in different reference frames lends itself to implement basic mathematical operations on a neuronal basis. Hence, it has been proposed that hand-related cortical circuits, too, may be involved in the interaction between space, number and attention (Hubbard et al., 2005).

The finding that finger gnosis in children predicts later mathematical skills in elementary school (Noël, 2005) implies a functional link between hand motor system and arithmetic. This functional association persists into adulthood. In adults, the spatial pattern of BOLD activity in posterior parietal cortex and IPS during multiplication and subtraction correlated significantly with the activation pattern during a finger discrimination task (Andres, Michaux, & Pesenti, 2012), implying shared functional circuits during both tasks. While it has been shown that this association is specific to finger movements and does not generalize to foot movements, for example (Michaux, Masson, Pesenti, & Andres, 2013), it remains unknown exactly what shared process fingers and arithmetic commonly hinge on.

The Gerstmann syndrome (Gerstmann, 1924), a tetrad of impaired finger gnosis, left-right confusion, agraphia and acalculia after left-hemispheric lesions affecting left angular gyrus (AG) and supramarginal gyrus (SMG) is sometimes interpreted as clinical evidence for an association between finger sense and arithmetic. In line with this idea, TMS over the left AG has been shown to lead to disruption of finger sense and arithmetic in healthy adults (Roux, Boetto, Sacko, Chollet, & Trémoulet, 2003; Rusconi, Walsh, & Butterworth, 2005).

Critically, mathematics education research suggests abandoning finger-based number representations in order to encourage children to form abstract magnitude representations. This is mainly based on the (true or false) inference that the use of finger-based arithmetic strategies is the reason for poor math abilities. Yet the above results imply that finger-based number representations linger on into adulthood and the observation of children with poor math abilities relying more on finger counting than normally developing children does not imply any causality. It may be that children rely on finger counting because they are poor in math. Another possibility is that those children relying on finger-based strategies show this behavior because

they lack the ability to flexibly shift between different representations which. This shifting, in turn, includes the inhibition of no-longer-needed representations.

In an fMRI experiment, MVPA will be used to classify the activation in hand-related areas during reaching and pointing. Participants will be asked to reach for or point to objects, following predefined trajectories that vary from trial to trial. LEGO® bricks will be used due their non-magnetic properties and easy customizability. An example is depicted in Figure 17. A classifier will be trained to recognize the executed movement direction based on activity in anterior (pointing) or posterior (reaching) aspects of the IPS that are determined by means of independent localizer scans. In a second step, the classifier is then presented with activation from addition and subtraction tasks from the same participant. If the planning of reaching movements is functionally associated to mental arithmetic, one would expect similar generalizations from right-ward reaching movements to additions and left-ward reaching movements to subtraction. If, however, the association between hands and numbers is due to unconscious counting or a mere co-localized activation no systematic pattern of generalization from movement to arithmetic operation is expected.

WP 3 - Interaction between domain-general and domain-specific factors during development

Mathematical thinking encompasses a broad variety of cognitive and neural processes related to the perception, understanding and manipulation of numerical content. This includes domain-specific factors such as the core system that enables us to approximately perceive and process numerical information, the *approximate number system (ANS)*. However, when investigating the development of mathematical cognition, we are not looking at an encapsulated cognitive module, supported by a single neurocognitive system but rather at a wide-spread network of interrelated cognitive processes with complex neural underpinnings. In the domain of numerical cognition, the focus has recently broadened towards the question to what degree domain-general factors contribute to the development of arithmetic competencies, such as inhibitory capacities (Allan, Hume, Allan, Farrington, & Lonigan, 2014). In a developmental framework of domain-relevance, Karmiloff-Smith (Karmiloff-Smith, 2015) suggests how the continuum of domain-general to domain-specific processes might develop: the infant brain starts out with a number of basic level processing tendencies or heuristics. Each of these tendencies might be more relevant to the processing of certain different kinds of input over others, i.e. more relevant to a particular domain, and thus can become more domain-specific over time. Dehaene and Cohen's (Dehaene & Cohen, 2007) neuronal recycling hypothesis could be seen as a potential neural explanation as to why some factors are more domain-relevant than others for a particular domain. As specified previously, the idea rests on the notion that numerical magnitude is mentally represented along a number line with smaller numbers to the left of larger numbers. When engaging in mental arithmetic, addition is

associated with attentional shifts to the right and subtraction with attentional shifts to the left. However, on top of this approximate heuristic (larger -> right; smaller -> left) that is based on the core number system, children acquire more formal symbolic number skills during schooling (counting, place value structure of Arabic numbers, written arithmetic, commutativity, etc.). Together, this may culminate in a cognitive system that some perceive as comprising at least three independent but highly interactive systems (Borst, Aite, & Houde, 2015; Houde & Borst, 2015). First, an intuitive system that uses heuristics and is able to operate on a fast gist of a given situation. Second, a more analytical, algorithmic system that operates on conscious and planned strategies. The third system is orthogonal to the first two systems and is used to regulate their respective involvement in a given situation (e.g. inhibit premature responses).

Despite recent evidence suggesting the involvement of inhibitory control functions and spatial attention to mental arithmetic (Szucs, Devine, Soltesz, Nobes, & Gabriel, 2014; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011), neither inhibition nor spatial attention represent unitary constructs. To the contrary, they comprise different facets (see below). What is still missing, are (a) a detailed description of the longitudinal trajectories and inter-relation of these factors and their facets, as well as (b) their respective neuro-functional correlates. We need to understand the key processes and how they (a) interact and (b) develop over lifetime.

Inhibitory control in mental arithmetic – Inhibition is a cognitive process that serves to suppress existent or forthcoming engagement in either perceptual or response-related processes. No consensus has been reached concerning the factorial structure of inhibition. One the one hand, inhibition has been divided into two main facets, response inhibition (go/no go task, stop signal task) and distractor resistance (as in number Stroop or letter flanker tasks) (Rey-Mermet, Gade, & Oberauer, 2017). An alternative factorization (Zhang, Geng, & Lee, 2017) of inhibition differentiates between interference resolution, action withholding (as in go/no go paradigms), and action cancellation (as in stop-signal task). At the neural level, inhibition relies on a set of regions including the inferior frontal gyrus, the right median cingulate, the paracingulate gyri, and the right superior parietal gyrus (Zhang et al., 2017) (Crone & Steinbeis, 2017)

Inhibition plays a pivotal role in mental arithmetic. During numerosity perception, inhibition is important for allowing the formation of an abstract representation over and above non-numerical stimulus features (Gilmore et al., 2013). During mental arithmetic, inhibition serves in excluding non-adapted strategies (Vanbinst & De Smedt, 2016) or inconclusive response alternatives (Cho et al., 2012). However, evidence is inconclusive (Bellon, Fias, & De Smedt, 2016; Keller & Libertus, 2015). This may in part be due to unspecific way the term inhibition has been used, equally referring to the inhibition of motor responses and irrelevant visual or

auditory information. A detailed description of how different facets of inhibition contribute to the development of numerical competencies remains elusive.

An ongoing project that received funding from the French National Agency for Research (ANR-18-CE28-0003) aims at disentangling the different components that drive the association of mental arithmetic with domain-specific and domain-general functions (*Figure 18*). In particular, we are interested in the developmental trajectory of this association which gives rise to seemingly non-linear trends (e.g. OM effect in babies and adults but not in children). Yet, this may merely reflect the differential impact of different heuristics, core capacities and emergence of controlled, top-down strategies.

Hypotheses – The emergence of numerical competencies is subject to different influences during development. During early years, intuitive behavior prevails. The behavior jointly reflects the characteristics of the most important neuro-cognitive core capacity – the approximate number system – and core intuitions about the physical world (e.g. object permanence). With increasing age, children acquire more elaborated and explicit numerical competencies such as counting, for example. These often precede the understanding of the underlying concepts, such as cardinality, for example. During this period, domain-general competencies such as inhibition start to emerge and influence behavior. As introduced above, inhibition and certain attentional facets continue developing throughout childhood and reach maturity at the age of emerging adulthood only. As children “develop increasing intentional control over their behavior and cognition (Bjorklund, 2013)”, they are increasingly capable of (a) inhibiting inappropriate (or planned) motor responses and (b) resisting to distractors. This is reflected in the increasing ability to reorient attention from invalidly cued to unattended locations in a Posner paradigm, for example.

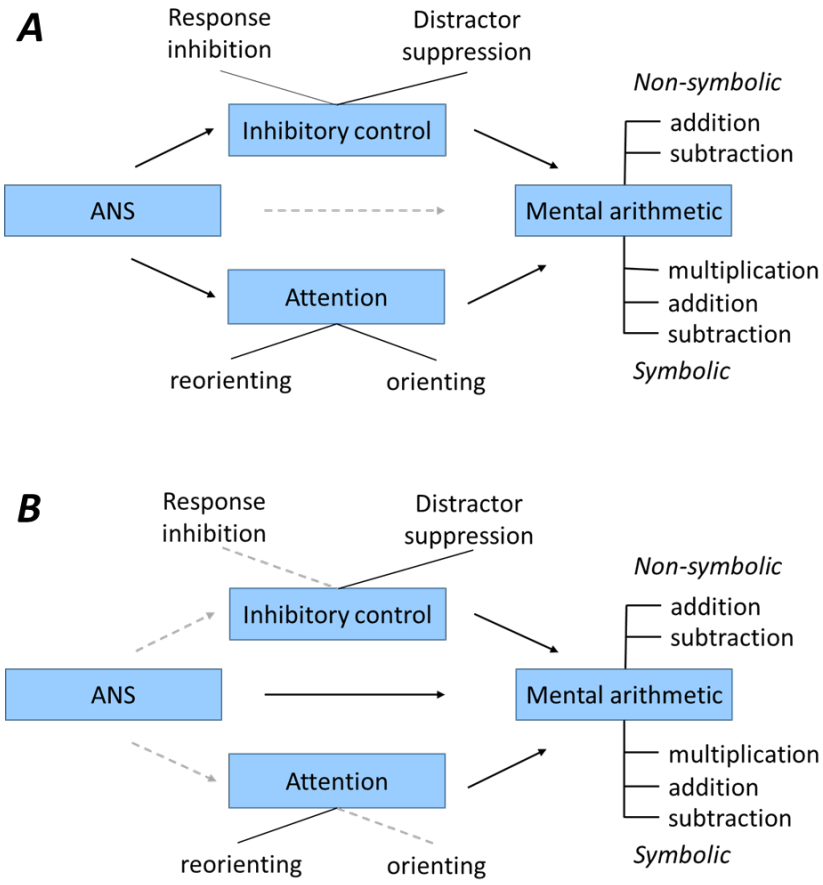

Figure 14: Schematic depiction of the complex interrelation between domain-specific factor ANS and domain-general factors (inhibition, attention). Here, only two of many possible relationships are shown. A. The ANS influences performance in numerical tasks only via domain-general factors (mediation model). B. ANS, distractor suppression (but not response inhibition), and reorienting (but not orienting) exert independent influence on mental arithmetic.

From this complex interaction and based on the above findings, the following hypotheses can be derived.

1. The performance in numerical tasks relies on the joint contributions of three cognitive systems (core capacities/heuristics, explicit strategies, executive control).
    a. During development, the impact of core capacities/heuristics decreases while the impact of the other systems increases.
    b. This should be reflected in the correlational pattern within a given cohort across testing sessions. For example, the correlation between numerical performance and ANS acuity at t1 for a given cohort should be higher compared to the correlation at t3.

c. At the neural level, we should not only observe a shift from parietal to frontal areas but also an increase in functional connectivity between these structures as children become older.

2. Inhibitory control differentially influences performance in numerical contexts.

   a. The association between distractor resistance and numerical performance is based on an improved functional connectivity between prefrontal and parietal areas that may (or may not) be reflected in structural connectivity. Activation patterns observed during inhibition tasks correlate with those during mental arithmetic if the stimulus yields contradictory information that need to be inhibited.

   b. In task contexts with contradictory stimulus-response associations, children's capability of resisting inappropriate responses will help improving performance in numerical tasks such as non-symbolic, approximate arithmetic if reliable 'default-associations' had been established beforehand. Otherwise, this facet of inhibitory control should not have a strong impact on performance.

   c. Training of distractor resistance should generalize to numerical contexts with contradictory stimulus dimensions and improve children's performance in incongruent conditions. No improvement should be observed in congruent conditions.

3. Attentional capacities differentially contribute to performance in numerical contexts:

   a. If the correlation between OM and re-orienting is due to the inhibitory component of the re-orienting, partialing out inhibition should have a detrimental effect on the correlation – even in simplified OM paradigms (see below).

   b. If the OM results from shifting attention along the mental number line, we should observe a positive correlation between benefit and OM effect. Symbolic math competencies correlate with orienting capacity.

   c. If OM-like effects in early pre-schoolers (1-3 years) rely on intuitive heuristics, they should neither be predictive of later school achievements nor OM effect.

   d. At the neural level, the association between attention and mental arithmetic should be reflected in correlated patterns of activation in parietal areas.

We will address the above questions in an accelerated longitudinal design (Galbraith, Bowden, & Mander, 2017) that goes beyond existent cross-sectional data (Figure 19). Four cohorts of children will be recruited during the project and tested once a year during three

consecutive years. At start, cohorts will be split by 2 years of age, starting at the age of 12 months (cohort 1) and ranging until the age of 7 years (cohort 4). Overall, this allows following the development of numerical skills and the relation with domain-specific and domain-general factors from the age of 1 until the age of 9. Cohorts (2-4) will consist of 100 children who will be tested during the last 20 weeks of each school year in partnering schools in the Paris area and Caen (i.e. on average ~12 children/week). To examine the early onsets of attentional and inhibitory contributions to basic mental arithmetic, Maria Dolores de Hevia will investigate and follow up on a group of 12-months-olds over the period of three years.

Data will be analyzed using latent growth models (c.f. (Watt, 2008)) and more simple mediation and correlational analyses.

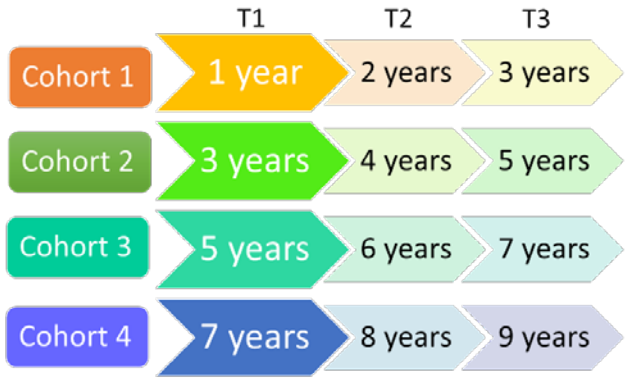

Figure 15: Schematic depiction of the longitudinal design for project years 1 to 3. Columns show the chronological age of the different cohorts (rows) at time of testing 1 (T1) to 3 (T3).

Cognitive measures

A bundle of specific cognitive parameters (standardized and experimental measures) will be measured, covering arithmetic and numerical performance, attentional parameters, inhibitory functions, and the most important control variables known to influence scholar achievement (e.g. general intelligence, socio-economic status, reading skills). By measuring different facets of inhibitory control (distractor suppression & response inhibition), we aim at delineating their differential contributions to the development of numerical competencies.

Table 2 lists the cognitive constructs we will assess in this project and an established test example. These refer to cohorts 2 – 4, since no standardized test are available for infants in cohort 1.

The methods used with 12-month-old infants will be modelled from our previous studies with infants (Macchi Cassia, Bulf, McCrink, & De Hevia, 2017; Macchi Cassia, McCrink, de Hevia, Gariboldi, & Bulf, 2016). In particular, we will use the habituation paradigm and our dependent variable will be the looking time during test trials. Infants will be habituated to a series of numerical sequences arranged in either increasing or decreasing order, and will be subsequently presented with three pair of test trials presenting new sequences in which the same ordinal direction will be composed of larger and smaller numerosities. If infants show an

OM effect while ordering number, then they should look longer at test to the series that violate the momentum created during habituation: those habituated to increasing number should look longer to smaller numerosities at test, and those habituated to decreasing number should look longer to larger numerosities at test.

Table 1: List of tests for the cognitive constructs relevant in the current project

Cohort	**Cognitive construct**	**Test**
Cohort 2 (3-5) & Cohort 3 (5-7)	distractor suppression	child-friendly version of flanker tasks
	response Inhibition	Animal Stroop; fruit & vegetable Stroop, Infant-friendly version of Stop-Signal task
	Attention (orienting & reorienting)	child-friendly version of Posner task
	ANS acuity	Panamath (www.panamath.org)
	Numerical competencies	Subtests of TEDI Math (counting forward & backward, object enumeration, numerical ordering objects & digits, addition, subtraction)
	IQ-facets	Subtests (Matrix Reasoning, Vocabulary, Picture Naming, Picture Memory) of WPPSI-IV
Cohort 4 (7-9)	distractor suppression	Flanker task Number Stroop
	response Inhibition	Stop-Signal task, fruit & vegetable Stroop
	Attention (orienting & reorienting)	Posner task
	ANS acuity	Panamath
	Numerical competencies	TEDI Math/TEDI Math Grands (numerical ordering objects & digits, addition, subtraction, multiplication, conceptual knowledge, base-10 system, Arabic number comparison) (Noël & Grégoire, 2015)
	IQ-facets	Subtests of WISC-V (Vocabulary, Matrix Reasoning, Picture Span) (Wechsler, 2014)

Methods used at 24 and 36 months of age will be modelled after our previous research with young children (de Hevia, Vanderslice, & Spelke, 2012). In particular, an experimenter will present cards showing either an increasing or a decreasing series of numerical

sequences and will then be shown two options to choose from, one that accords and one that violates the momentum created during the demonstration. We expect children to manifest their OM by choosing the series that accords with the momentum. A computerized version of the task will be also envisioned.

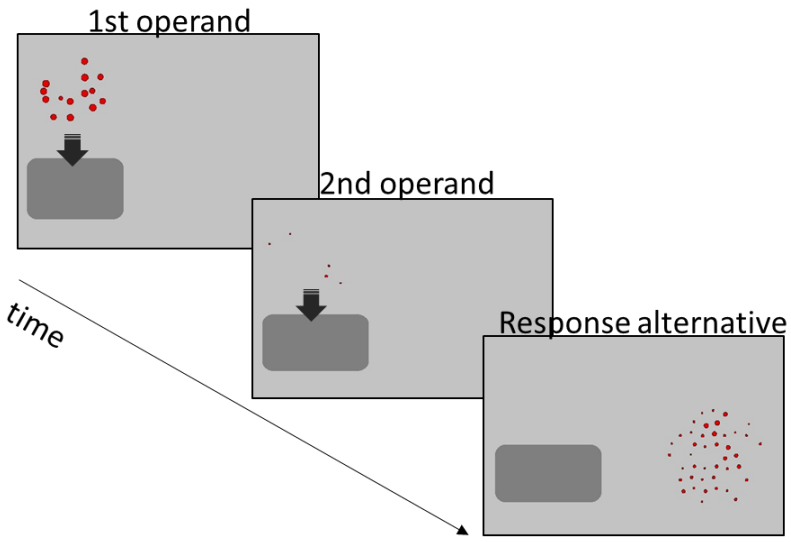

Figure 16: Screen shots from the paradigm programmed for the non-symbolic calculation task. After seeing one number of dots vanishing behind an occluder, a second number of dots vanishes behind the occluder (addition) or will be removed from behind it (subtraction). In the final step, a response alternative is presented next to the occluder. The task is to decide whether the quantity behind the occluder (the outcome) or the response alternative is numerically larger.

fMRI task

We go beyond previous studies by not only measuring behavioral parameters but also the neuro-functional correlates of mental arithmetic and how these coincide with more domain-general brain functions. In a longitudinal design, twenty children from cohort 3 and 4 will be tested once a year over the period of three years. During scanning, we will administrate inhibition tasks (distractor suppression & response inhibition), an attentional task (Posner task), and a *non-symbolic* addition and subtraction task while recording whole brain BOLD response. For cohort 4 (age: 7 – 9), we will also present *symbolic* addition, subtraction, and multiplication problems.

Applying state-of-the-art multivariate analysis methods (e.g. representational similarity analysis) allow us to go beyond the description of merely overlapped activity and specifically parametrize the degree to which numerical performance recruits domain-general capacities. High resolution structural MRI data and DTI will allow tracing morphological changes and relate them to functional parameters (behavioral & neural).

References

Abrams, J., Barbot, A., & Carrasco, M. (2010). Voluntary attention increases perceived spatial frequency. *Attention, perception & psychophysics, 72*(6), 1510–1521. doi: 10.3758/app.72.6.1510

Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: a meta-analysis. *Dev Psychol, 50*(10), 2368-2379. doi: 10.1037/a0037493

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual review of neuroscience, 25*, 189–220. doi: 10.1146/annurev.neuro.25.112701.142922

Anderson, M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behav Brain Sci, 33*(4), 245-266; discussion 266-313. doi: 10.1017/S0140525X10000853

Andres, M., Michaux, N., & Pesenti, M. (2012). Common substrate for mental arithmetic and finger representation in the parietal cortex. *NeuroImage, 62*(3), 1520–1528. doi: 10.1016/j.neuroimage.2012.05.047

Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition, 122*(3), 454-459.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2012). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vision Res, 74*, 102-109. doi: 10.1016/j.visres.2012.06.005

Anton-Erxleben, K., Henrich, C., & Treue, S. (2007). Attention changes perceived size of moving visual patterns. *Journal of vision, 7*(11), 5.1–9. doi: 10.1167/7.11.5

Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., & Pascual-Leone, J. (2018). Brain areas associated with numbers and calculations in children: Meta-analyses of fMRI studies. *Dev Cogn Neurosci, 30*, 239-250. doi: 10.1016/j.dcn.2017.08.002

Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science, 21*, 224-233.

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev Sci, 14*(1), 125-135. doi: 10.1111/j.1467-7687.2010.00962.x

Beck, J. M., Latham, P. E., & Pouget, A. (2011). Marginalization in neural circuits with divisive normalization. *Journal of Neuroscience, 31*(43), 15310-15319. doi: 10.1523/JNEUROSCI.1706-11.2011

Bellon, E., Fias, W., & De Smedt, B. (2016). Are Individual Differences in Arithmetic Fact Retrieval in Children Related to Inhibition? *Front Psychol, 7*, 825. doi: 10.3389/fpsyg.2016.00825

Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science, 299*(5603), 81-86. doi: 10.1126/science.1077395

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev Psychol, 42*(1), 189-201. doi: 10.1037/0012-1649.41.6.189

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev, 79*(4), 1016-1031. doi: 10.1111/j.1467-8624.2008.01173.x

Borst, G., Aite, A., & Houde, O. (2015). Inhibition of misleading heuristics as a core mechanism for typical cognitive development: evidence from behavioural and brain-imaging studies. *Dev Med Child Neurol, 57 Suppl 2*, 21-25. doi: 10.1111/dmcn.12688

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philos Trans R Soc Lond B Biol Sci, 364*(1525), 1831-1840. doi: 10.1098/rstb.2009.0028

Bulthé, J., Smedt, B. d., & Op de Beeck, H. P. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. *NeuroImage, 87*, 311–322. doi: 10.1016/j.neuroimage.2013.10.049

Buneo, C. A., & Andersen, R. A. (2006). The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia, 44*(13), 2594–2606. doi: 10.1016/j.neuropsychologia.2005.10.011

Buracas, G. T., & Boynton, G. M. (2007). The effect of spatial attention on contrast response functions in human visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 27*(1), 93–97. doi: 10.1523/jneurosci.3162-06.2007

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *J Vis, 10*(6), 20. doi: 10.1167/10.6.20

Butterworth, B., & Kovas, Y. (2013). Understanding neurocognitive developmental disorders can improve education for all. *Science, 340*(6130), 300-305. doi: 10.1126/science.1231022

Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A. R. (2001). Storage and retrieval of addition facts: the role of number comparison. *Q J Exp Psychol A, 54*(4), 1005-1029. doi: 10.1080/713756007

Campbell, J. I., & Metcalfe, A. W. (2007). Arithmetic rules and numeral format. *Eur. J. Cogn. Psychol., 19*(335-355). doi: 10.1080/09541440600717610

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res, 51*(13), 1484-1525. doi: 10.1016/j.visres.2011.04.012

Casarotti, M., Michielin, M., Zorzi, M., & Umiltà, C. (2007). Temporal order judgment reveals how number magnitude affects visuospatial attention. *Cognition, 102*(1), 101-117. doi: 10.1016/j.cognition.2006.09.001

Castelli, F., Glaser, D. E., & Butterworth, B. (2006). Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. *Proc Natl Acad Sci U S A, 103*(12), 4693-4698. doi: 10.1073/pnas.0600444103

Cavdaroglu, S., Katz, C., & Knops, A. (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception. *NeuroImage, 116*, 135–148. doi: 10.1016/j.neuroimage.2015.04.019

Cavdaroglu, S., & Knops, A. (2015). Mental subtraction and multiplication recruit both phonological and visuospatial resources: evidence from a symmetric dual-task design. *Psychological research*. doi: 10.1007/s00426-015-0667-8

Cavdaroglu, S., & Knops, A. (2018). Evidence for a Posterior Parietal Cortex Contribution to Spatial but not Temporal Numerosity Perception. *Cereb Cortex, 29*(7), 2965 - 2977. doi: 10.1093/cercor/bhy163

Cave, C. B., Blake, R., & McNamara, T. P. (1998). Binocular rivalry disrupts visual priming. *Psychological Science, 9*, 299-302.

Cho, S., Metcalfe, A. W., Young, C. B., Ryali, S., Geary, D. C., & Menon, V. (2012). Hippocampal-prefrontal engagement and dynamic causal interactions in the maturation of children's fact retrieval. *J Cogn Neurosci, 24*(9), 1849-1866. doi: 10.1162/jocn_a_00246

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. *Neuron, 53*(2), 307–314. doi: 10.1016/j.neuron.2006.12.025

Crone, E. A., & Steinbeis, N. (2017). Neural Perspectives on Cognitive Control Development during Childhood and Adolescence. *Trends Cogn Sci, 21*(3), 205-215. doi: 10.1016/j.tics.2017.01.003

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proc Natl Acad Sci U S A, 108*(49), 19552-19557. doi: 10.1073/pnas.1113195108

de Hevia, M. D., Vanderslice, M., & Spelke, E. S. (2012). Cross-dimensional mapping of number, length and brightness by preschool children. *PLoS One, 7*(4), e35530. doi: 10.1371/journal.pone.0035530

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: a neuronal model. *Journal of cognitive neuroscience, 5*(4), 390–407. doi: 10.1162/jocn.1993.5.4.390

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model in number processing. *Mathematical Cognition, 1*, 83 - 120.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384-398. doi: 10.1016/j.neuron.2007.10.004

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of experimental psychology. Human perception and performance, 16*(3), 626–641.

Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., . . . Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature, 395*(6702), 597-600.

Didino, D., Knops, A., Vespignani, F., & Kornpetpanee, S. (2015). Asymmetric activation spreading in the multiplication associative network due to asymmetric overlap between numerosities semantic representations? *Cognition, 141*, 1–8. doi: 10.1016/j.cognition.2015.04.002

Didino, D., Pinheiro-Chagas, P., Wood, G., & Knops, A. (2019). Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect. *Front Psychol, 10*, 160. doi: 10.3389/fpsyg.2019.00160

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Dev Psychol, 43*(6), 1428-1446. doi: 10.1037/0012-1649.43.6.1428

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009a). Deciphering cortical number coding from human brain activity patterns. *Curr.Biol., 19*(19), 1608-1615.

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009b). Deciphering cortical number coding from human brain activity patterns. *Current biology : CB, 19*(19), 1608–1615. doi: 10.1016/j.cub.2009.08.047

Fischer, M. H., Miklashevsky, A., & Shaki, S. (2018). Commentary : The Developmental Trajectory of the Operational Momentum Effect. *Front Psychol, 9*, 2259. doi: 10.3389/fpsyg.2018.02259

Fischer, M. H., & Shaki, S. (2014). Spatial associations in numerical cognition--from single digits to arithmetic. *Quarterly journal of experimental psychology (2006), 67*(8), 1461–1483. doi: 10.1080/17470218.2014.927515

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn Sci, 17*(3), 134-141. doi: 10.1016/j.tics.2013.01.010

Galbraith, S., Bowden, J., & Mander, A. (2017). Accelerated longitudinal designs: An overview of modelling, power, costs and handling missing data. *Stat Methods Med Res, 26*(1), 374-398. doi: 10.1177/0962280214547150

Galfano, G., Penolazzi, B., Vervaeck, I., Angrilli, A., & Umilta, C. (2009). Event-related brain potentials uncover activation dynamics in the lexicon of multiplication facts. *Cortex, 45*(10), 1167-1177. doi: 10.1016/j.cortex.2008.09.003

Galfano, G., Rusconi, E., & Umilta, C. (2003). Automatic activation of multiplication facts: evidence from the nodes adjacent to the product. *Q J Exp Psychol A, 56*(1), 31-61. doi: 10.1080/02724980244000332

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychol (Amst), 171*, 17-35. doi: 10.1016/j.actpsy.2016.09.003

Gebuis, T., & Reynvoet, B. (2014). The neural mechanism underlying ordinal numerosity processing. *J Cogn Neurosci, 26*(5), 1013-1020. doi: 10.1162/jocn_a_00541

Gerardi, K., Goette, L., & Meier, S. (2013). Numerical ability predicts mortgage default. *Proc Natl Acad Sci U S A, 110*(28), 11267-11271. doi: 10.1073/pnas.1220568110

Gerstmann, J. (1924). *Fingeragnosie: eine umschriebene Störung der Orientierung am eigenen Körper*: Springer.

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., . . . Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One, 8*(6), e67374. doi: 10.1371/journal.pone.0067374

Glaser, M., & Knops, A. (2020). When adding is right: Temporal order judgements reveal spatial attention shifts during two-digit mental arithmetic. *Q J Exp Psychol (Hove), 73*(7), 1115-1132. doi: 10.1177/1747021820902917
Gottlieb, J. (2007). From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron, 53*(1), 9-16. doi: 10.1016/j.neuron.2006.12.009
Gross, J., Hudson, J., & Price, D. (2009). *The long term costs of numeracy difficulties: Every Child a Chance trust*. London: KMPG.
Guillaume, M., Mejias, S., Rossion, B., Dzhelyova, M., & Schiltz, C. (2018). A rapid, objective and implicit measure of visual quantity discrimination. *Neuropsychologia, 111*, 180-189. doi: 10.1016/j.neuropsychologia.2018.01.044
Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Dev Psychol, 44*(5), 1457-1465. doi: 10.1037/a0012682
Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc Natl Acad Sci U S A, 109*(28), 11116-11120. doi: 10.1073/pnas.1200196109
Hannagan, T., Nieder, A., Viswanathan, P., & Dehaene, S. (2017). A random-matrix theory of the number sense. *Philos Trans R Soc Lond B Biol Sci, 373*(1740). doi: 10.1098/rstb.2017.0253
Harvey, B. M., & Dumoulin, S. O. (2017). Can responses to basic non-numerical visual features explain neural numerosity responses? *NeuroImage, 149*, 200-209. doi: 10.1016/j.neuroimage.2017.02.012
Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science, 341*(6150), 1123-1126. doi: 10.1126/science.1239052
Hesselmann, G. (2013). Dissecting visual awareness with FMRI. *Neuroscientist, 19*(5), 495-508. doi: 10.1177/1073858413485988
Hesselmann, G., Darcy, N., Sterzer, P., & Knops, A. (2015). Exploring the boundary conditions of unconscious numerical priming effects with continuous flash suppression. *Conscious Cogn, 31*, 60-72. doi: 10.1016/j.concog.2014.10.009
Hesselmann, G., & Knops, A. (2014). No conclusive evidence for numerical priming under interocular suppression. *Psychol Sci, 25*(11), 2116-2119. doi: 10.1177/0956797614548876
Houde, O., & Borst, G. (2015). Evidence for an inhibitory-control theory of the reasoning brain. *Front Hum Neurosci, 9*, 148. doi: 10.3389/fnhum.2015.00148
Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature reviews. Neuroscience, 6*(6), 435–448. doi: 10.1038/nrn1684
Huckauf, A., Knops, A., Nuerk, H. C., & Willmes, K. (2008). Semantic processing of crowded stimuli? *Psychol Res, 72*(6), 648-656. doi: 10.1007/s00426-008-0171-5
Iwamura, Y., Iriki, A., & Tanaka, M. (1994). Bilateral hand representation in the postcentral somatosensory cortex. *Nature, 369*(6481), 554–556. doi: 10.1038/369554a0
Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *J Neurosci, 31*(38), 13535-13545. doi: 10.1523/JNEUROSCI.1691-11.2011
Karmiloff-Smith, A. (2015). An alternative to domain-general or domain-specific frameworks for theorizing about human evolution and ontogenesis. *AIMS Neurosci, 2*(2), 91-104. doi: 10.3934/Neuroscience.2015.2.91
Katz, C., & Knops, A. (2014). Operational momentum in multiplication and division? *PLoS One, 9*(8), e104777. doi: 10.1371/journal.pone.0104777
Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The Discrimination of Visual Number. *The American Journal of Psychology, 62*(4), 498-525. doi: 10.2307/1418556

Keller, L., & Libertus, M. (2015). Inhibitory control may not explain the link between approximation and math abilities in kindergarteners from middle class families. *Front Psychol, 6*, 685. doi: 10.3389/fpsyg.2015.00685
Kiesel, A., Kunde, W., & Hoffmann, J. (2008). Mechanisms of subliminal response priming. *Adv Cogn Psychol, 3*(1-2), 307-315. doi: 10.2478/v10053-008-0032-1
Knops, A., Dehaene, S., Berteletti, I., & Zorzi, M. (2014). Can approximate mental calculation account for operational momentum in addition and subtraction? *Q J Exp Psychol (Hove), 67*(8), 1541-1556. doi: 10.1080/17470218.2014.890234
Knops, A., Hoesterey, H., & Katz, C. (2017). A role for attentional reorienting during approximate multiplication and division. *Journal of Numerical Cognition, 3*(2), 246-269.
Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(30), 9857–9866. doi: 10.1523/jneurosci.2758-13.2014
Knops, A., Thirion, B., Hubbard, E. M., Michel, V., & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science (New York, N.Y.), 324*(5934), 1583–1585. doi: 10.1126/science.1171599
Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. *Attention, perception & psychophysics, 71*(4), 803–821. doi: 10.3758/app.71.4.803
Knops, A., Zitzmann, S., & McCrink, K. (2013). Examining the presence and determinants of operational momentum in childhood. *Front Psychol, 4*, 325. doi: 10.3389/fpsyg.2013.00325
Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 1882–1905.
Konen, C. S., Mruczek, R. E. B., Montoya, J. L., & Kastner, S. (2013). Functional organization of human posterior parietal cortex: grasping- and reaching-related activations relative to topographically organized cortex. *Journal of neurophysiology, 109*(12), 2897–2908. doi: 10.1152/jn.00657.2012
Kosillo, P., & Smith, A. T. (2010). The role of the human anterior insular cortex in time processing. *Brain Struct Funct, 214*(5-6), 623-628. doi: 10.1007/s00429-010-0267-8
Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88*(2), 223-242.
Lasne, G., Piazza, M., Dehaene, S., Kleinschmidt, A., & Eger, E. (2019). Discriminability of numerosity-evoked fMRI activity patterns in human intra-parietal cortex reflects behavioral numerical acuity. *Cortex, 114*, 90-101. doi: 10.1016/j.cortex.2018.03.008
Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2016). From “sense of number” to “sense of magnitude”: The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences, 40*, e164. doi: 10.1017/S0140525X16000960
Leoné, F. T. M., Toni, I., & Medendorp, W. P. (2014). Two-dimensional spatial tuning for saccades in human parieto-frontal cortex. *NeuroImage, 87*, 476–489. doi: 10.1016/j.neuroimage.2013.09.067
Li, S. X., & Cai, Y. C. (2014). The effect of numerical magnitude on the perceptual processing speed of a digit. *J Vis, 14*(12). doi: 10.1167/14.12.18
Li, X., & Basso, M. A. (2008). Preparing to move increases the sensitivity of superior colliculus neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 28*(17), 4561–4577. doi: 10.1523/jneurosci.5683-07.2008
Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense. Large-number discrimination in human infants. *Psychol Sci, 14*(5), 396-401.
Liu, D., Cai, D., Verguts, T., & Chen, Q. (2017). The Time Course of Spatial Attention Shifts in Elementary Arithmetic. *Sci Rep, 7*(1), 921. doi: 10.1038/s41598-017-01037-3

Lyons, I. M., Ansari, D., & Beilock, S. L. (2015). Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain. *Human brain mapping, 36*(2), 475–488. doi: 10.1002/hbm.22641
Macchi Cassia, V., Bulf, H., McCrink, K., & De Hevia, M. D. (2017). Operational Momentum during ordering operations for size and number in 4-month-old infants. *Journal of numerical cognition, 3*(2), 270-287.
Macchi Cassia, V., McCrink, K., de Hevia, M. D., Gariboldi, V., & Bulf, H. (2016). Operational momentum and size ordering in preverbal infants. *Psychol Res, 80*(3), 360-367. doi: 10.1007/s00426-016-0750-9
Martínez-Trujillo, J., & Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron, 35*(2), 365–370.
Masson, N., Letesson, C., & Pesenti, M. (2018). Time course of overt attentional shifts in mental arithmetic: Evidence from gaze metrics. *Q J Exp Psychol (Hove), 71*(4), 1009-1019. doi: 10.1080/17470218.2017.1318931
Masson, N., & Pesenti, M. (2014). Attentional bias induced by solving simple and complex addition and subtraction problems. *Q J Exp Psychol (Hove), 67*(8), 1514-1526. doi: 10.1080/17470218.2014.903985
Masson, N., & Pesenti, M. (2016). Interference of lateralized distractors on arithmetic problem solving: a functional role for attention shifts in mental calculation. *Psychol Res, 80*(4), 640-651. doi: 10.1007/s00426-015-0668-7
Masson, N., Pesenti, M., & Dormal, V. (2013). Spatial bias in symbolic and non-symbolic numerical comparison in neglect. *Neuropsychologia, 51*(10), 1925-1932. doi: 10.1016/j.neuropsychologia.2013.06.004
Masson, N., Pesenti, M., & Dormal, V. (2017). Impact of optokinetic stimulation on mental arithmetic. *Psychol Res, 81*(4), 840-849. doi: 10.1007/s00426-016-0784-z
Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., & Prado, J. (2016). Running the number line: Rapid shifts of attention in single-digit arithmetic. *Cognition, 146*, 229–239. doi: 10.1016/j.cognition.2015.10.002
Mazza, V., & Caramazza, A. (2015). Multiple object individuation and subitizing in enumeration: a view from electrophysiology. *Frontiers in Human Neuroscience, 9*(162). doi: 10.3389/fnhum.2015.00162
McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 19*(1), 431–441.
McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Perception & psychophysics, 69*(8), 1324–1333.
McCrink, K., & Hubbard, T. L. (2017). Dividing Attention Increases Operational Momentum. *J Num Cogn, 3*(2), 230-245.
McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *J Exp Child Psychol, 103*(4), 400-408. doi: 10.1016/j.jecp.2009.01.013
Melcher, D., & Piazza, M. (2011). The role of attentional priority and saliency in determining capacity limits in enumeration and visual working memory. *PLoS One, 6*(12), e29296. doi: 10.1371/journal.pone.0029296
Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct, 214*(5-6), 655-667. doi: 10.1007/s00429-010-0262-0
Michaux, N., Masson, N., Pesenti, M., & Andres, M. (2013). Selective interference of finger movements on basic addition and subtraction problem solving. *Exp Psychol, 60*(3), 197–205. doi: 10.1027/1618-3169/a000188
Miller, P. (2013). Stimulus number, duration and intensity encoding in randomly connected attractor networks with synaptic depression. *Front Comput Neurosci, 7*, 59. doi: 10.3389/fncom.2013.00059
Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*(5109), 1519–1520.

Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of neurophysiology, 83*(5), 2580–2601.
Naccache, L., & Dehaene, S. (2001a). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cereb Cortex, 11*(10), 966-974. doi: 10.1093/cercor/11.10.966
Naccache, L., & Dehaene, S. (2001b). Unconscious semantic priming extends to novel unseen stimuli. *Cognition, 80*(3), 215-229.
Niedeggen, M., Rosler, F., & Jost, K. (1999). Processing of incongruous mental calculation problems: evidence for an arithmetic N400 effect. *Psychophysiology, 36*(3), 307-324.
Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual review of neuroscience, 32*, 185–208. doi: 10.1146/annurev.neuro.051508.135550
Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science (New York, N.Y.), 313*(5792), 1431–1435. doi: 10.1126/science.1130308
Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297*(5587), 1708-1711.
Nieder, A., & Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *Journal of Neuroscience, 27*(22), 5986-5993. doi: 10.1523/JNEUROSCI.1056-07.2007
Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron, 37*(1), 149-157.
Nieder, A., & Miller, E. K. (2004). Analog numerical representations in rhesus monkeys: evidence for parallel processing. *J.Cogn Neurosci., 16*(5), 889-901.
Noël, M. P. (2005). Finger gnosia: a predictor of numerical abilities in children? *Child Neuropsychol, 11*(5), 413-430. doi: 10.1080/09297040590951550
Noël, M. P., & Grégoire, J. (2015). *TediMath Grands, Test diagnostique des compétences de base en mathématiques du CE2 à la 5e.* Montreuil, France.
Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cogn Psychol, 55*(3), 169-195. doi: 10.1016/j.cogpsych.2006.09.002
Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* London: NRDC.
Peters, L., & De Smedt, B. (2018). Arithmetic in the developing brain: A review of brain imaging studies. *Developmental Cogn Neurosci, 30*, 265 - 279. doi: http://dx.doi.org/10.1016/j.dcn.2017.05.002
Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends Cogn Sci, 14*(12), 542-551. doi: 10.1016/j.tics.2010.09.008
Piazza, M., & Eger, E. (2015). Neural foundations and functional specificity of number representations. *Neuropsychologia*. doi: 10.1016/j.neuropsychologia.2015.09.025
Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition, 121*(1), 147-153. doi: 10.1016/j.cognition.2011.05.007
Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron, 53*(2), 293–305. doi: 10.1016/j.neuron.2006.11.022
Pinel, P., Piazza, M., Le, B. D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron, 41*(6), 983-993.
Pinhas, M., & Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition, 109*(3), 408-415. doi: 10.1016/j.cognition.2008.09.003
Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., . . . Haase, V. G. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One, 9*(11), e111155. doi: 10.1371/journal.pone.0111155

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychol Sci, 19*(6), 607-614. doi: 10.1111/j.1467-9280.2008.02130.x
Rey-Mermet, A., Gade, M., & Oberauer, K. (2017). Should We Stop Thinking About Inhibition? Searching for Individual and Age Differences in Inhibition Ability. *J Exp Psychol Learn Mem Cogn*. doi: 10.1037/xlm0000450
Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron, 26*(3), 703–714.
Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychol Sci, 24*(7), 1301-1308. doi: 10.1177/0956797612466268
Roggeman, C., Fias, W., & Verguts, T. (2010). Salience maps in parietal cortex: imaging and computational modeling. *NeuroImage, 52*(3), 1005-1014. doi: 10.1016/j.neuroimage.2010.01.060
Roggeman, C., Santens, S., Fias, W., & Verguts, T. (2011). Stages of nonsymbolic number processing in occipitoparietal cortex disentangled by fMRI adaptation. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 31*(19), 7168–7173. doi: 10.1523/jneurosci.4503-10.2011
Roggeman, C., Verguts, T., & Fias, W. (2007). Priming reveals differential coding of symbolic and non-symbolic quantities. *Cognition, 105*(2), 380-394. doi: 10.1016/j.cognition.2006.10.004
Rosenberg-Lee, M., Barth, M., & Menon, V. (2011). What difference does a year of schooling make? Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *Neuroimage., 57*(3), 796-808.
Ross, J. (2003). Visual discrimination of number without counting. *Perception, 32*(7), 867-870. doi: 10.1068/p5029
Roux, F.-E., Boetto, S., Sacko, O., Chollet, F., & Trémoulet, M. (2003). Writing, calculating, and finger recognition in the region of the angular gyrus: a cortical stimulation study of Gerstmann syndrome. *Journal of neurosurgery, 99*(4), 716–727. doi: 10.3171/jns.2003.99.4.0716
Rusconi, E., Galfano, G., Rebonato, E., & Umilta, C. (2006). Bidirectional links in the network of multiplication facts. *Psychol Res, 70*(1), 32-42. doi: 10.1007/s00426-004-0187-4
Rusconi, E., Galfano, G., Speriani, V., & Umilta, C. (2004). Capacity and contextual constraints on product activation: evidence from task-irrelevant fact retrieval. *Q J Exp Psychol A, 57*(8), 1485-1511. doi: 10.1080/02724980343000873
Rusconi, E., Walsh, V., & Butterworth, B. (2005). Dexterity with numbers: rTMS over left angular gyrus disrupts finger gnosis and number processing. *Neuropsychologia, 43*(11), 1609–1624. doi: 10.1016/j.neuropsychologia.2005.01.009
Shaki, S., Pinhas, M., & Fischer, M. H. (2018). Heuristics and biases in mental arithmetic: revisiting and reversing operational momentum. *Think. Reason., 24*, 138-156. doi: 10.1080/13546783.2017.1348987
Shaki, S., Sery, N., & Fischer, M. H. (2015). 1+ 2 is more than 2 + 1: violations of commutativity and identity axioms in mental arithmetic. *J. Cogn. Psychol., 27*, 471-477. doi: 10.1080/20445911.2014.973414
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev, 75*(2), 428-444. doi: 10.1111/j.1467-8624.2004.00684.x
Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol Sci, 14*(3), 237-243. doi: 10.1111/1467-9280.02438
Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *J Exp Psychol Gen, 142*(1), 193-208. doi: 10.1037/a0028560
Smets, K., Sasanguie, D., Szücs, D., & Reynvoet, B. (2015). The effect of different methods to construct non-symbolic stimuli in numerosity estimation and comparison. *Journal of Cognitive Psychology, 27*(3), 310-325. doi: 10.1080/20445911.2014.996568
Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A. C., & Nuerk, H. C. (2017). Increased arithmetic complexity is associated with domain-

general but not domain-specific magnitude processing in children: A simultaneous fNIRS-EEG study. *Cogn Affect Behav Neurosci, 17*(4), 724-736. doi: 10.3758/s13415-017-0508-x

Soltanlou, M., Artemenko, C., Ehlis, A. C., Huber, S., Fallgatter, A. J., Dresler, T., & Nuerk, H. C. (2018). Reduction but no shift in brain activation after arithmetic learning in children: A simultaneous fNIRS-EEG study. *Sci Rep, 8*(1), 1707. doi: 10.1038/s41598-018-20007-x

Sterzer, P., Stein, T., Ludwig, K., Rothkirch, M., & Hesselmann, G. (2014). Neural processing of visual information under interocular suppression: a critical review. *Front Psychol, 5*, 453. doi: 10.3389/fpsyg.2014.00453

Stoianov, I., & Zorzi, M. (2012). Emergence of a 'visual number sense' in hierarchical generative models. *Nat Neurosci, 15*(2), 194-196. doi: 10.1038/nn.2996

Stoianov, I. P., & Zorzi, M. (2017). Computational foundations of the visual number sense. *Behav Brain Sci, 40*, e191. doi: 10.1017/S0140525X16002326

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2014). Cognitive components of a mathematical processing network in 9-year-old children. *Dev Sci, 17*(4), 506-524.

Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental brain research, 83*(1), 29–36.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428*(6984), 751-754. doi: 10.1038/nature02466

Tokita, M., & Ishiguchi, A. (2010). How might the discrepancy in the effects of perceptual variables on numerosity judgment be reconciled? *Atten Percept Psychophys, 72*(7), 1839-1853. doi: 10.3758/APP.72.7.1839

Tokita, M., & Ishiguchi, A. (2012). Behavioral evidence for format-dependent processes in approximate numerosity representation. *Psychon Bull Rev, 19*(2), 285-293. doi: 10.3758/s13423-011-0206-6

Toll, S. W., Van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2011). Executive functions as predictors of math learning disabilities. *J Learn Disabil, 44*(6), 521-532. doi: 10.1177/0022219410387302

Tong, F., Harrison, S. A., Dewey, J. A., & Kamitani, Y. (2012). Relationship between BOLD amplitude and pattern classification of orientation-selective activity in the human visual cortex. *NeuroImage, 63*(3), 1212–1222. doi: 10.1016/j.neuroimage.2012.08.005

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol Rev, 101*(1), 80-102.

Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent surfaces. *Vision Res, 45*(9), 1095–1098. doi: 10.1016/j.visres.2004.11.001

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nat Neurosci, 8*(8), 1096-1101. doi: 10.1038/nn1500

Tsuchiya, N., Koch, C., Gilroy, L. A., & Blake, R. (2006). Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision, 6*, 1068-1078.

Turatto, M., Vescovi, M., & Valsecchi, M. (2007). Attention makes moving objects be perceived to move faster. *Vision Res, 47*(2), 166–178. doi: 10.1016/j.visres.2006.10.002

Van Essen, D. C. (2005). A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *NeuroImage, 28*(3), 635-662. doi: 10.1016/j.neuroimage.2005.06.058

Van Essen, D. C., & Dierker, D. L. (2007). Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron, 56*(2), 209-225. doi: 10.1016/j.neuron.2007.10.015

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association, 8*(5), 443-459.

Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: comparison and priming effects in numerical and nonnumerical orders. *Psychon Bull Rev, 15*(2), 419-425.
Vanbinst, K., & De Smedt, B. (2016). Individual differences in children's mathematics achievement: The roles of symbolic numerical magnitude processing and domain-general cognitive functions. *Prog Brain Res, 227*, 105-130. doi: 10.1016/bs.pbr.2016.04.001
Verguts, T., & Fias, W. (2004a). Representation of number in animals and humans: a neural model. *Journal of cognitive neuroscience, 16*(9), 1493–1504. doi: 10.1162/0898929042568497
Verguts, T., & Fias, W. (2004b). Representation of number in animals and humans: a neural model. *J.Cogn Neurosci., 16*(9), 1493-1504.
Verguts, T., & Fias, W. (2005). Interacting neighbors: A connectionist model of retrieval in single-digit multiplication. *Memory & Cognition, 33*(1), 1-16. doi: 10.3758/bf03195293
Watt, H. M. G. (2008). A latent growth curve modeling approach using an accelerated longitudinal design: the ontogeny of boys' and girls' talent perceptions and intrinsic values through adolescence. *Educational Research and Evaluation, 14*(4), 287-304.
Wechsler, D. (2014). *Wechsler intelligence scale for children-fifth edition*. Bloomington, MN: Pearson.
Winter, B., Matlock, T., Shaki, S., & Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci Biobehav Rev, 57*, 209-219. doi: 10.1016/j.neubiorev.2015.09.005
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1-B11.
Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Dev Sci, 8*(1), 88-101. doi: 10.1111/j.1467-7687.2005.00395.x
Zhang, R., Geng, X., & Lee, T. M. C. (2017). Large-scale functional neural network correlates of response inhibition: an fMRI meta-analysis. *Brain Struct Funct, 222*(9), 3973-3990. doi: 10.1007/s00429-017-1443-x
Zimba, L. D., & Blake, R. (1983). Binocular rivalry and semantic processing: out of sight, out of mind. *J Exp Psychol Hum Percept Perform, 9*(5), 807-815.

Research Activities and Supervision

I received funding from 4 different funding organizations (DFG [Germany], ANR [France], ERC [Europe]) for 6 different projects and one international workshop with an overall volume of € 1 700 000. Among these grants, the funding from the German Research Foundation (KN 959/2) allowed me to set up my own research team at the Humboldt-Universität Berlin, where I supervised 2 PhD students and 1 Post-Doctoral Student. The ANR funding allows me to supervise 1 PhD student and 1 post-doctoral student. Before accepting the CNRS position in France, I stood in for Prof. Peter Frensch and headed the unit 'Allgemeine Psychology' [Experimental Psychology] of the Humboldt-Universität for one year. This entailed ca. 270 hours of teaching and the supervision and administration of 3 post-doctoral students, 6 PhD students and 4 student assistants.

Grants [as PI]

2018 - 2022	How Domain-general Functions Contribute to the Development of Numerical Competencies (Agence National de la Recherche, ANR; DFCDNC) [PI]	433,000 €
2018 - 2020	On the structure of arithmetic facts in memory and its interaction with numerical magnitude representation (main applicant: Daniele Didino) [PI]	150,000 €
2015	Symposium ‚Contributions of dorsal stream to visual perception'. DFG (German Research Foundation) HE 6244/3-1 with Guido Hesselmann	10,000 €
2012 - 2016	The impact of the number sense and spatial processing on mental arithmetic. DFG KN 959/2-1 [PI]	840,000 €
2009 - 2010	Neural Dynamics of mental arithmetic – An fMRI guided TMS-Study. DFG KN 959/1-1	113,000 €
2010	*Attentional shifts in mental arithmetic – nature and relevance. Autonomous Region Trentino (ERC Marie Curie Action Cofund) – declined*	*131,000 €*
2006 - 2007	Emotional Control of Cognition – On the Influence of Mood on Executive Functions in Patients with Biploar Mood Disorder. Nachwuchsprogram der Medizinischen Fakultät der RWTH Aachen, AZ 34/06	36,500 €

Publications with supervised post-doctoral students (+), PhD students (*) and undergraduate students (#):

1. **Glaser*, M.** & **Knops, A.** (2020). When adding is right – temporal order judgments reveal spatial attention shifts during two-digit mental arithmetic. *Quarterly Journal of Experimental Psychology*, 73(7), 1115–1132.
2. **Didino[+], D., Breil[#], C.** & **Knops, A.** (2019). The influence of semantic processing and response latency on the SNARC effect. *Acta Psychologica, 196*: 75-86.
3. **Didino[+], D.**, Pinheiro-Chagas, P., Wood, G. & **Knops, A.** (2019) Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect. *Front. Psychol. 10*:160. doi: 10.3389/fpsyg.2019.00160
4. **Cavdaroglu*, S.** & **Knops, A.** (2018). Evidence for a posterior parietal cortex contribution to spatial but not temporal numerosity perception, *Cerebral Cortex, 29(7)*, 2965–2977, https://doi.org/10.1093/cercor/bhy163
5. Pinheiro-Chagas, P., **Didino[+], D.** Haase, V. G., Wood, G. & **Knops, A.** (2018). The developmental trajectory of the operational momentum effect, *Front. Psychol. 9*:1062. doi: 10.3389/fpsyg.2018.01062
6. **Katz*, C. Hoesterey[#]**, H. & **Knops, A.** (2017). A role for attentional reorienting during approximate multiplication and division. *Journal of Numerical Cognition, 3(2),* 246 - 269*.*
7. **Katz*, C.** & **Knops, A.** (2016). Decreased cerebellar-cerebral connectivity contributes to complex task performance. *Journal of Neurophysiology, 116,* 1434–1448.
8. **Cavdaroglu*, S.** & **Knops, A.** (2016). Mental subtraction and multiplication recruit both phonological and visuospatial resources: evidence from a symmetric dual-task design. *Psychological Research*, *80(4)*: 608-624.
9. **Cavdaroglu*, S., Katz*, C.** & **Knops, A.** (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception, *Neuroimage*, *116*: 135-148.
10. **Didino[+], D.**, **Knops, A.**, Vespignani, F., & Kornpetpanee, S. (2015). Asymmetric activation spreading in the multiplication associative network due to asymmetric overlap between numerosities semantic representations? *Cognition, 141*, 1-8*.*
11. **Katz*, C.** & **Knops, A.** (2014). Operational momentum in approximate multiplication and division? *PLoS One; 9(8):e104777. doi: 10.1371/journal.pone.0104777.*

Original Article

QJEP

Quarterly Journal of Experimental Psychology
2020, Vol. 73(7) 1115–1132

DOI: 10.1177/1747021820902917
qjep.sagepub.com
SAGE

When adding is right: Temporal order judgements reveal spatial attention shifts during two-digit mental arithmetic

Maria Glaser[1] and André Knops[1,2]

Abstract

Recent research suggests that addition and subtraction induce horizontal shifts of attention. Previous studies used single-digit (1d) problems or verification paradigms that lend themselves to alternative solution strategies beyond mental arithmetic. To measure spatial attention during the active production of solutions to complex two-digit arithmetic problems (2d) without manual motor involvement, we used a temporal order judgement (TOJ) paradigm in which two lateralised targets were sequentially presented on screen with a varying stimulus onset asynchrony (SOA). Participants verbally indicated which target appeared first. By varying the delay between the arithmetic problem presentation and the TOJ task, we investigated how arithmetically induced attention shifts develop over time (Experiment 1, $n = 31$ and Experiment 2, $n = 58$). In Experiment 2, we additionally varied the carry property of the arithmetic task to examine how task difficulty modulates the effects. In the arithmetic task, participants were first presented with the arithmetic problem via headphones and performed the TOJ task after the delay before responding to the arithmetic task. To account for spontaneous attentional biases, a baseline TOJ was run without arithmetic processing. Both experiments revealed that addition induces shifts of spatial attention to the right suggesting that visuospatial attention mechanisms are recruited during complex arithmetic. We observed no difference in spatial attention between the carry and noncarry condition (Experiment 2). No shifts were observed for subtraction problems. No common and conclusive influence of delay was observed across experiments. Qualitative differences between addition and subtraction and the role of task difficulty are discussed.

Keywords

Spatial biases; numerical cognition; MNL; numbers and space; SNA; mathematical cognition

Received: 7 June 2019; revised: 26 November 2019; accepted: 5 December 2019

A large amount of research points into the direction of a functional association between numbers and space (Knops, 2018). For example, several studies demonstrated that the processing of small and large numbers can induce shifts of visuospatial attention to the left and right side, respectively (Casarotti et al., 2007; Fischer et al., 2003; Galfano et al., 2006; Nicholls et al., 2008). Indicating that this association is bidirectional, other studies showed that the manipulation of spatial attention can affect numerical processing (Ranzini et al., 2015, 2016; Stoianov et al., 2008). These spatial-numerical associations imply that internal spatial representations activated by number magnitude and external representations of space recruit similar attentional mechanisms. In this context, the most common metaphor for numerical representation is the mental number line (MNL; Dehaene, 1992; Restle, 1970). The central idea behind the MNL is that numerical magnitude is represented on an analogue number line in ascending order with smaller numbers on the left and larger numbers on the right.

The spatial organisation of the numerical representations has implications for their manipulation during mental arithmetic. McCrink et al. (2007) observed an overestimation when participants were solving nonsymbolic addition problems and underestimation when they were solving nonsymbolic subtraction problems (termed *operational momentum* effect, OM). Knops, Viarouge, and Dehaene

[1]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany
[2]LaPsyDÉ, UMR CNRS 8240, Université de Paris, Paris, France

Corresponding author:
Maria Glaser, Department of Psychology, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany.
Email: maria.glaser@hu-berlin.de

(2009) extended this observation to symbolic (Arabic digits) numerosities but found that the effect was stronger for nonsymbolic arithmetic. The origins of the OM are still under debate. One prominent account recognises this effect as evidence for an impact of the spatial-numerical representation on arithmetic. Specifically, the *spatial attention account* (Hubbard et al., 2005; McCrink et al., 2007) proposes that addition and subtraction processing involves covert movements along the MNL to the right and left, respectively. The OM is supposedly caused by an "overshoot" during this motion on the MNL. This account received support from a neuroimaging study that showed that parietal activation patterns induced by addition processing resembled activations produced by rightward saccades (Knops, Thirion, et al., 2009), suggesting a shared neural circuitry between arithmetic processing and visuospatial processing.

However, alternative accounts exist that attribute the OM effect to nonspatial mechanisms or at least not exclusively to spatial processes: According to the *compression account* (Chen & Verguts, 2012; McCrink et al., 2007), the OM is caused by a systematically inaccurate decompression of presumably logarithmically compressed magnitude representations. The calculation on these faulty magnitudes will lead to an overestimation or underestimation of the result. Furthermore, according to the "if adding, accept more" and "if subtracting, accept less" *heuristics account* (McCrink et al., 2007; McCrink & Wynn, 2009), the OM is caused by the utilisation of the general principle that for addition (subtraction) outcomes are accepted as long as they are larger (smaller) than the initial operand. McCrink and Hubbard (2017) recently proposed that the heuristics account and the attentional shift account might even belong to one single mechanism (*heuristics-via-spatial shifts account*). They suggested a greater reliance on a heuristic where information from the visuospatial system is fed into the decision when attentional load is high. Finally, the *arithmetic heuristics and biases* model (AHAB, Shaki et al., 2018) postulates three competing mechanisms: the more-or-less heuristic, a sign-space association, and an anchoring bias. These biases concurrently contribute to the performance and produce differential patterns of performance depending on their relative contribution weight in the given task.

Strong support for a role of attentional biases in arithmetic processing comes from studies that show how arithmetic processing influences spatial attention. First, Knops, Viarouge, and Dehaene (2009) reported that participants preferably selected response options on the right-hand side for addition and on the left-hand side for subtraction. Termed *spatial-operation association of responses* (SOAR), this effect already suggested a link between mental arithmetic and spatial attention shifts. Using mainly symbolic arithmetic problems, some have started to study that link systematically and showed that the processing of addition and subtraction problems can induce spatial biases to the right and left, respectively (M. Li et al., 2018; Liu et al., 2017; Masson et al., 2018; Masson & Pesenti, 2014). The influence of arithmetic processing on attention has also been observed via gaze behaviour (Klein et al., 2014; Zhu et al., 2018). Further studies demonstrating how spatial attention (Masson & Pesenti, 2016; Mathieu et al., 2016) and eye movements (Masson et al., 2017) affect mental arithmetic highlight the bidirectional association between attentional shifts and arithmetic problem solving.

The research reported in this article aims at further elucidating the spatial attention shifts induced by symbolic mental arithmetic. Taking a more systematic look at existing studies on that issue, it becomes obvious that they varied with regard to two important aspects: the complexity of the arithmetic task (*simple, moderate, and complex*[1]) and the point in time at which the spatial measurement occurred. With regard to the complexity, the majority of studies investigating spatial biases induced by arithmetic processing used *simple* or *moderate* arithmetic problems (Masson et al., 2018; Masson & Pesenti, 2014; Zhu et al., 2018). Evidence from studies using *complex* arithmetic is scarce (Klein et al., 2014; M. Li et al., 2018; Liu et al., 2017) and inconclusive. For example, Liu and colleagues (2017) systematically varied the onset of a target detection task with respect to the presentation of an arithmetic verification task. Yet, there remains a doubt concerning the actual nature of the mental processing during the course of a verification task. In particular, verification tasks lend themselves to alternative strategies of checking the correctness of the presented response alternative (e.g., parity check, approximation, and guessing) that do not necessarily involve the performance of a basic arithmetic operation. Hence, the first question that arises is whether spatial biases can also be induced by complex arithmetic stimuli. While traditionally procedural solution strategies have been distinguished from memory retrieval (Ashcraft & Guillaume, 2009; Barrouillet & Fayol, 1998; Zbrodoff & Logan, 2005), recent studies suggest that the importance of procedural solution strategies increases from simple to complex arithmetic problems (Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012; Uittenhove et al., 2016). Hence, with complex arithmetic stimuli in the context of an active production task the activation of numerical representations would be arguably increased due to the heightened use of procedural strategies leading to even stronger spatial biases. Yet, to date few studies have used complex problems in an active task context.

Second, regarding the stage at which the spatial measurement is introduced, the attentional shift hypothesis predicts strong spatial biases after the presentation of the second operand and before the result is given, that is, during calculation—because this is when the activation of the numerical representation is the most likely. However, only one study measured spatial attention within that interval and used an attentional paradigm (Liu et al., 2017, Experiments 2, 3). Hence, introducing a measurement of spatial attention within the calculation phase of an arithmetic production task may further elucidate the contribution of attentional

processes during calculation rather than response-related processing stages and alternative strategies.

To address these knowledge gaps, this study used a temporal order judgement (TOJ) task to measure deflections of spatial attention. This paradigm has already been used in the field of numerical cognition with single digits (Casarotti et al., 2007) and has the advantage of assessing spatial attention free from motor processes because it requires unspeeded verbal responses. It entailed the visual presentation of two lateralised targets with varying SOAs. Participants had to indicate which stimulus was presented first (right first vs. left first). The response to the TOJ task had to be given verbally to avoid effects of stimulus-response compatibility that come with button press reactions. Based on these responses, the point of subjective simultaneity (PSS) could be calculated. It constitutes the time lag between the lateralised stimuli that is needed for them to be perceived as being presented simultaneously. It illustrates the estimated point in time where the probability of saying that the right stimulus appeared first equals the probability of saying that the left stimulus appeared first (i.e., 50 % each). A negative PSS indicates that the left stimulus needs to be presented earlier (implying attention in right visual field) and a positive PSS indicates that the right stimulus needs to be presented earlier (implying attention in left visual field).

The key objectives of this study can be summarised as follows: We aimed at investigating whether the active production of solutions to addition (subtraction) problems induces attentional shifts to the right (left). Therefore, we presented a TOJ task as a spatial measurement within the calculation stage, that is, before the result had to be given. We used two-digit addition and subtraction problems to examine whether attentional shifts can be induced by complex arithmetic tasks. In addition, we decided to vary the delay between the second operand and the TOJ task, similarly to Liu et al. (2017), because there is still some uncertainty as to when arithmetically induced attentional shifts occur, how long these shifts persist and whether addition and subtraction differ regarding the time course of attentional shifts. We predicted that addition problems shifted attention more rightward compared to a baseline (see Method section) and compared to subtraction problems. For subtraction problems, on the contrary, we expected a more leftward shift of attention compared to a baseline and to addition problems. No directional hypotheses were made regarding the delay or its interaction with the operation during the arithmetic task.

Experiment 1

Method

Participants first performed a baseline TOJ task and then continued with the arithmetic TOJ task, which contained a mixed set of addition and subtraction tasks. The baseline was included to control for naturally occurring attentional biases (e.g., pseudoneglect, Bowers & Heilman, 1980) and consisted of the TOJ paradigm without any arithmetic processing. During the arithmetic TOJ task, participants were first presented with the arithmetic problem via headphones and after a varying delay performed the TOJ task, before verbally providing the solution of the arithmetic problem. This response ordering was chosen because the probability of calculation processes going on and, therefore, being able to measure arithmetically induced spatial biases is largest when its measurement is put in the time window between the arithmetic task presentation and the arithmetic response.

The arithmetic problems were presented auditorily to control for implicit influences of reading direction and thereby shifts of attention that are associated with the visual task presentation (even if the task components are presented sequentially). Another advantage of this modality was that if any observations were made on a visuospatial level (i.e., in the TOJ task), these could not originate from simple short-term associational processes, but would have to be due to a central processing stage. This would indicate that the numerical representation in itself is spatial.

To account for the fact that the "auditory presentation of numerical information may be responsible for a relatively slow mapping function compared to visual number presentation" (Myachykov et al., 2016, pp. 385–386), we chose a wide spectrum of delays (i.e., 250, 750, and 1,500 ms) between the arithmetic task presentation and the TOJ task.

Finally, while Liu et al. (2017) asked their participants to judge whether a given proposal to the arithmetic problem was correct, the task of this study was to give the result of the arithmetic problem to ensure that participants fully completed the adequate calculation process and did not make use of heuristics or approximate calculation processes.

Participants. Thirty-nine German-speaking students from the Humboldt-Universität zu Berlin took part in the experiment in exchange for course credit. Each of the participants had normal or corrected-to-normal vision and hearing. Six of the participants were excluded from further analysis because their average arithmetic accuracy was below 50%. Data from two further participants was removed due to experimental errors. The remaining 31 participants (19 female; 28 right-handed) were aged between 18 and 33 years ($M = 23.26$, $SD = 4.20$). The experiment was noninvasive and all procedures were carried out in accordance with the ethical standards established by the Declaration of Helsinki and approved by the Ethics Committee of the Humboldt-Universität zu Berlin. Informed consent was obtained in written format from all individual participants included in the study.

Design. For the purpose of the study a 3 (operation: baseline, addition, subtraction) × 3 (delay in ms: bins around 250, 750, 1,500) × 8 (SOA in ms: −110, −70, −40, −20, 20,

Table 1. Stimulus characteristics of arithmetic problems of Experiment 1.

Operation	Problemsize		Operand 1		Operand 2		Result	
	M	*SD*	Min	Max	Min	Max	Min	Max
Addition	159.83	25.85	31	76	21	45	53	98
Subtraction	160.20	24.41	52	98	21	49	31	73

40, 70, 110, negative values indicate that the left target appeared first) within-subject design was employed.[2]

Stimuli. The two-digit arithmetic problems were constructed using the following criteria: the first operand (O1) was always larger than the second operand (O2) to prevent participants from inferring the operation from the magnitude of the first operand (see also Table 1). The mean problem size (defined as O1 + O2 + result) was matched between addition (M = 159.8) and subtraction problems (M = 160.2) to eliminate the confounds between the magnitude of the problems and shifts of spatial attention. We used an equal amount of carry and noncarry problems,[3] as well as an equal amount of odd and even results for each operation. Furthermore, we excluded arithmetic problems with repeated operands/result (e.g., 43 + 43 = 86, 98 – 49 = 49), as well as problems with operands with identical unit or decade values (e.g., 42 + 32 = 74), due to better memory accessibility for this kind of tasks compared to problems with nonidentical elements (Campbell & Gunter, 2002). Finally, we eliminated problems that contained zero digits in their operands or result (e.g., 20, 80) or multiples of 11 (e.g., 33, 66).[4] The item-selection process resulted in a list of 192 addition and 192 subtraction problems (8 SOA levels × 3 delay levels × 8 repetitions = 192), so that during the experimental block of the experiment, no arithmetic problem was presented twice. For practise trials, a random set of 12 stimuli was selected out of the full list of arithmetic problems (see S1 in the online Supplementary Material for the complete list of arithmetic problems of Experiment 1).

The audio (.wav) files of the number words for the auditory task presentation were recordings in stereo from a female native speaker of German done by a professional media service. They were edited via Audacity (version: 2.1.3) as follows: All audio files of operands were brought to the same length of 1,500 ms (SD < 1 ms) by reducing silence within number words and by consistently allowing a silence of around 150 ms before voice onset (M_{Onset} = 151 ms, SD_{Onset} = 22 ms) and a silence of around 100 ms after voice offset (M_{Offset} = 1,409 ms, SD_{Offset} = 56 ms). Likewise, the audio files of the operators were brought to the same length of 750 ms with similar voice onset ($Plus_{Onset}$ = 150 ms, $Minus_{Onset}$ = 140 ms) and offset ($Plus_{Offset}$ = 594 ms, $Minus_{Offset}$ = 626 ms).

Apparatus. The experiment was presented using Matlab software (R2015a) and the *Psychtoolbox* package (version 3.0.14; Brainard, 1997) on a PC and a 24-inch LCD monitor (resolution 1,080 × 1,920; refresh rate 100 Hz). Auditory stimuli were delivered binaurally via headphones with equal intensity. Responses were recorded with a microphone and noted down online by the experimenter who sat in the same room as the participant.

Task and procedure. Participants were seated at a desk facing the screen at a distance of approximately 60 cm and were asked to keep their gaze fixated on the centre of the screen throughout the experiment. All visual stimuli were presented on a grey background (RGB: [127.5 127.5 127.5], 80 cd/m^2) to minimise visual after-effects. At the beginning of each trial, a white fixation dot (radius: .48°) appeared in the centre of the screen. The visual TOJ targets consisted of two black squares (1.98° × 1.98°) which were presented symmetrically in the left and right visual field, with their centre 5.09° from fixation. This resulted in a distance from fixation to size of stimulus ratio of about 2.5 which is in line with other studies making use of the TOJ paradigm (e.g., Eimer & Grubert, 2015; S.-X. Li & Cai, 2014; Liddle et al., 2009; Poncelet & Giersch, 2015).

In the baseline TOJ task, participants completed 12 practise trials before the start of the 192 experimental trials, which were presented in four blocks of 48 trials each. SOAs and delays for the practise trials were chosen randomly from all existing conditions. During one individual baseline TOJ trial (Figure 1A), participants were first presented with the fixation dot. To ensure central fixation, the fixation dot was presented throughout the trial. After the duration of 500 ms plus a variable delay (250, 750, and 1,500 ms) as well as a jitter (−80, −40, +40, +80 ms), the first TOJ stimulus appeared right or left to fixation. After a varying SOA (−110, −70, −40, −20, 20, 40, 70, 110 ms; negative values indicate that the left target appeared first), the second TOJ stimulus appeared on screen (both targets visible). Then, the participant had to verbally indicate which target had appeared first (/bi/ = left and /be/ = right). The syllables "bi" and "be" were used to keep the initial consonant identical. Thereby, the detection of the voice onset could not be confounded by the initial consonant. Voice onset for the TOJ task was measured from the onset of the second TOJ stimulus and was only used for timing the subsequent trial. About 1,100 ms after the participant's voice onset, the screen turned grey for a 500 ms intertrial interval (ITI). If 2 s after the onset of the second stimulus no verbal response

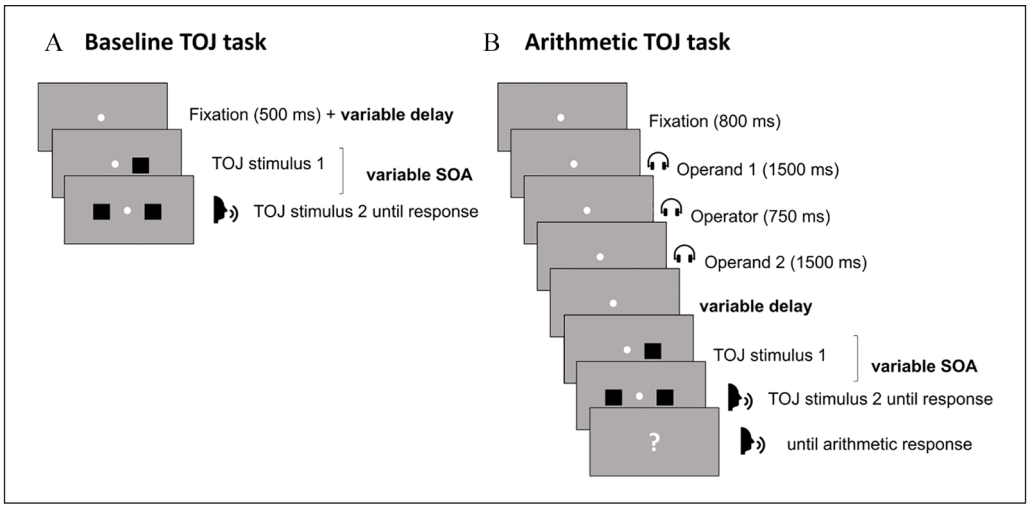

Figure 1. Task sequence and timing of a single trial of the baseline TOJ task (A) and the arithmetic TOJ task (B).

was detected, participants were prompted to respond faster or louder by a message on the screen for 1,600 ms before the ITI and the next trial was launched.

In the arithmetic TOJ task, participants completed 12 practise trials before the start of the 384 experimental trials, which were presented in 12 blocks of 32 trials each. Again, SOAs and delays for the practise trials were chosen randomly from all existing conditions. Within an individual arithmetic TOJ trial (Figure 1B) participants were first presented with the fixation dot. After 800 ms, the first operand (1,500 ms), the operator (750 ms), and the second operand (1,500 ms) were presented auditorily via headphones. Then, after a variable delay (250, 750, 1,500 ms after the offset of the second operand) and a jitter (−80, −40, +40, +80 ms) the two TOJ stimuli appeared right and left to fixation with a varying SOA (−110, −70, −40, −20, 20, 40, 70, 110 ms) and the participant again had to indicate verbally which target had appeared first (/bi/ = left and /be/ = right). About 1,000 ms after the participant's voice onset in response to the TOJ task, a question mark replaced the fixation dot in the centre of the screen to indicate that the verbal response for the arithmetic task could be given. Voice onset time for the arithmetic task was measured from the onset of the question mark. The screen turned grey for a 750 ms-ITI, 2,100 ms after voice onset had been detected. If 10 s after the onset of the question mark no verbal response was detected, participants were again prompted to respond faster or louder by a message on the screen for 1,600 ms before the ITI. This response deadline was used to encourage an onset of calculation before the TOJ task. For both the TOJ task and the arithmetic task, participants received no feedback on their judgements and erroneous trials were not repeated. The complete experimental session lasted 2.5 to 3 hr, and participants had the opportunity to take breaks between the experimental blocks.

Data analysis. Only data from experimental trials was analysed. Responses (TOJ task: left first vs. right first; arithmetic task: response to the arithmetic problem), as well as response times (voice onsets) were recorded. Relevant dependent variables for the analysis reported here were arithmetic accuracies, arithmetic voice onset times and TOJ responses. Because only TOJ responses of arithmetically accurate trials could be analysed further, data of six participants had to be removed as their arithmetic accuracy was below 50 %. For analysis of the arithmetic voice onset times we used only cases of correctly answered arithmetic problems and only cases within the interval of ± 2 *SD* of the individual's mean.

The crucial measurement of spatial attention was data from the TOJ task, both in the baseline condition and in the arithmetic task (addition and subtraction condition). These data allowed us to calculate the individual points of subjective simultaneity (PSS), which constitute the estimated time lag between the lateralised TOJ stimuli where the probability of saying that the right TOJ stimulus appeared first equals the probability of saying that the left stimulus appeared first. Negative PSSs indicate that for equal probabilities, the left stimulus would have to be presented earlier which implies attention in the right visual field. Positive PSSs indicate that for equal probabilities, the right stimulus would have to be presented earlier which implies attention in the left visual field. It was calculated by fitting logistic functions to the psychometric data, that is, the TOJ responses, using a maximum likelihood criterion. For each subject and each condition combination (operation × delay), two values were calculated per SOA: (a) *valid trials*: the absolute number of TOJ responses out of eight repetitions, where a response to the TOJ task was given (i.e., no missings) and in case of the arithmetic task, where the arithmetic task was responded to accurately and (b) the absolute number of *right-first responses* of the TOJ

task within these valid trials. The resulting number of valid trials and corresponding right-first responses (baseline task: 99.73%, addition: 80.29%, subtraction: 74.73% of all trials) were then used for a curve-fitting procedure that was implemented in the *Palamedes* toolbox for Matlab (version 1.8.2; Prins, 2018). Guess rates γ and lapse rates λ were fixed at 0 and at .01, respectively. The remaining parameters (threshold α and slope β) were fitted separately for each subject and operation × delay combination. Goodness of fit was estimated for each function. In case of the present parametrisation, the PSS is equivalent to the threshold parameter α of the fitted curve. Hence, the curve-fitting procedure resulted in individual PSSs for each subject and each condition combination (operation × delay). PSSs of curves with an unacceptably poor goodness of fit *p* value below .05 were discarded (i.e., 12.5% of the whole PSS data set, cf. Kingdom & Prins, 2016). No PSS fell outside the interval between the two extreme SOAs (−110, 110).

The final PSS values were analysed using linear mixed-effects models (LMMs) using the *lme4* package for R (Bates et al., 2015). This approach was favoured over an analysis of variance (ANOVA) because the final PSS data set contained missing values (due to the exclusion of PSSs of poorly fitted curves), and mixed-effects models are more robust with respect to missing data (Baayen et al., 2008). As a first step of the modelling procedure, the ideal random-effects structure had to be established. To this end, we tested whether the inclusion of a random effect significantly improved the fit of the model. Then, the fixed-effects structure was analysed by comparing the goodness of fit of models with and without the fixed effects while keeping the random-effects structure constant. For all model comparisons, likelihood-ratio tests were used. To test specific hypotheses, we eventually inspected the parameter statistics of the final model. Whether slopes differed significantly from zero was assessed via *t*-tests using the Satterthwaite approximation of the degrees of freedom (lmerTest package for R, Kuznetsova et al., 2017). All statistical analyses were carried out using the R software (version 3.3.1, Team, 2016).

Results

Arithmetic accuracy. Mean accuracy of the reduced sample (see above) was moderate ($M = .80$, $SD = .12$). Operation-wise, mean accuracies were .82 ($SD = .11$) for the addition condition and .77 ($SD = .12$) for the subtraction condition. A *t*-test on the arcsine-transformed accuracy rates revealed that participants made significantly more errors in subtraction trials compared to addition trials, $t(30) = 5.07$, $p < .001$.

Arithmetic voice onset times. Mean voice onset to the arithmetic task was 2.03 s ($SD = 1.58$). For addition, the mean voice onset time was 1.83 s ($SD = 1.83$) and for subtraction 2.11 s ($SD = 2.01$). A paired *t*-test on the voice onset times revealed that participants took significantly longer to respond to subtraction problems compared to addition problems, $t(30) = -5.57$, $p < .001$.

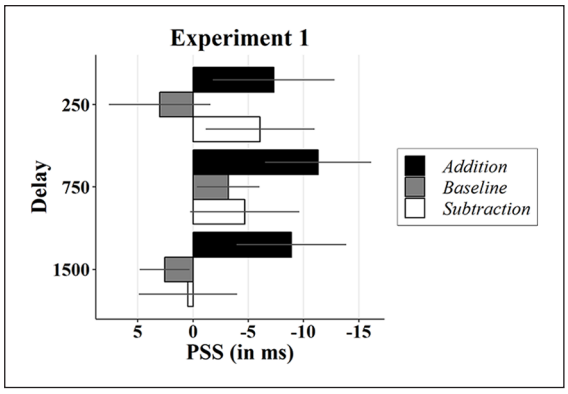

Figure 2. Mean PSSs (points of subjective simultaneity) for each operation and delay condition of Experiment 1.
Error bars represent the standard error of the mean (SEM).
PSS values constitute the time lags between the lateralised TOJ stimuli that are needed for them to be perceived as being presented simultaneously. Positive PSSs indicate attention in the left visual field and negative PSSs indicate attention in the right visual field.

PSS. Averaged over all delays and participants, the mean PSS was 1.15 ms ($SD = 15.07$) for the baseline TOJ task, −8.93 ms ($SD = 24.42$) within the addition condition and −4.09 ms ($SD = 23.62$) within the subtraction condition (See Figure 2 for a detailed depiction of mean PSSs).

To assess whether operation and delay or their interaction had a significant impact on the PSS, we ran LMMs.[5] First, the random-effects structure had to be determined. To this end, we compared the baseline model that contained all fixed effects (operation, delay, and their interaction) as well as a random effect of subjects (i.e., an adjustment to the intercept grouped by subjects; 1|subject) to a second model which additionally contained random slopes for operation (i.e., allowing the effect of operation to vary across subjects; operation|subject). This significantly improved the fit of the model, $\chi^2(5) = 44.05$, $p < .001$. However, the further inclusion of random slopes for delay (i.e., allowing the effect of delay to vary over subjects; delay|subject) did not improve the fit of the model, $\chi^2(6) = 4.74$, $p = .577$, and the model that additionally contained random slopes for the interaction between operation and delay did not converge implying that it might have been too complex for the amount of data. Hence, the final random-effects structure included random intercepts over subjects and random slopes for operation (operation|subject).

Second, the fixed-effects structure had to be assessed. For that purpose, the random-effects structure was kept constant and the fixed effects were added incrementally. Here, adding the effect of operation significantly improved the fit, $\chi^2(2) = 6.68$, $p < .05$. However, further including the

Table 2. Estimated parameters and statistics of mixed-effects modelling of PSS data from Experiment 1.

Predictors	PSS				
	Estimates	*SE*	*CI* (95%)	*T*	*p*
(Intercept)	1.08	2.62	[−4.06, 6.22]	.41	.683
Addition	−10.18	3.77	[−17.57, −2.79]	−2.70	**.011***
Subtraction	−5.05	3.08	[−11.08, .98]	−1.64	.111

PSS: point of subjective simultaneity; CI: confidence interval.
*$p < .05$.

Table 3. Estimated parameters and statistics of mixed-effects modelling of PSS data of the arithmetic task subset from Experiment 1.

Predictors	PSS				
	Estimates	*SE*	*CI* (95%)	*T*	*p*
(Intercept)	−9.06	4.20	[−17.30,−.83]	−2.16	**.038***
Subtraction	5.11	2.21	[.78, 9.45]	2.31	**.022***
Delay (linear)	1.75	2.76	[−3.65, 7.15]	.63	.527
Delay (quadratic)	4.49	2.65	[−.70, 9.69]	1.70	.092
Subtraction × Delay (linear)	−2.66	3.94	[−10.37, 5.06]	−.67	.501
Subtraction × Delay (quadratic)	−3.23	3.74	[−10.55, 4.09]	−.86	.389

PSS: point of subjective simultaneity; CI: confidence interval.
*$p < .05$.

effect of delay only marginally improved the model fit, $\chi^2(2)=5.96$, $p=.051$, and the inclusion of the interaction term (operation*delay) did not improve the fit of the model at all, $\chi^2(4)=1.93$, $p=.748$. This indicates that the type of operation had a significant effect on the individual PSSs. For delay, there was only a trend, and the interaction was not significant at all.

Looking at the final model (PSS ~ operation + operation|subject) that included only operation as fixed factor (contrast coded with the TOJ baseline condition set as baseline), it became obvious that the PSS values were significantly more negative when the participants were solving addition problems compared to the baseline condition, $t(30.64)=-2.70$, $p<.05$, (Table 2). This indicates that in the addition condition, the left target had to be presented earlier than the right stimulus to be perceived as being presented simultaneously.

Table 2 also reveals that the intercept does not significantly differ from zero, $t(29.82)=.41$, $p=.683$. Because operation is the only predictor of the final model, this indicates that the PSSs of the baseline condition did not differ significantly from zero (because the intercept is the expected mean when all predictors are zero or set to their reference level).

Additional analysis of the arithmetic task subset. The arithmetic task consisted of a dual-task paradigm and was thereby inherently different from the baseline task which was a single-task paradigm. Therefore, we decided to explore the time course of attentional shifts in the arithmetic task, that is, in the addition and subtraction conditions, without the baseline task. The determination of the random-effects structure resulted in random intercepts over subjects (1|subject). The fixed-effects analysis revealed a significant improvement of the model fit when the fixed effect of operation was entered, $\chi^2(1)=5.32$, $p<.05$, and no improvement of the fit neither with the fixed effect of delay, $\chi^2(2)=2.35$, $p=.309$, nor with the interaction of delay and operation, $\chi^2(2)=1.20$, $p=.55$. Based on the research question of the time course, we analysed the larger model that contained the effects of operation, delay, and interaction term ($PSS_{arithmetic}$ ~ operation + delay + operation*delay + 1|subject). The factor delay was contrast coded as polynomial to check for trends in the PSS over delay. The analysis showed that PSS values were less negative when the participants were solving subtraction problems compared to when they were solving addition problems, $t(131.36)=2.31$, $p<.05$ (see Table 3). Furthermore, PSSs tended to change in a quadratic fashion over time with most negative PSSs during the delay of 750 ms, $t(131.24)=1.70$, $p=.092$.

Discussion

In the first experiment, we set out to investigate whether and (if so) when complex arithmetic processing induces shifts of spatial attention. As the attentional shift hypothesis predicts strong spatial biases during the calculation phase, we introduced the TOJ task at exactly this point in time. We found that, compared to the baseline which involved no arithmetic processing, in the addition condition the left TOJ

target had to be presented earlier than the right target to be perceived as being presented simultaneously. This is interpreted as a more rightward directed attention during addition processing than during the baseline. Hence, our results suggest that two-digit addition problems can induce attentional shifts to the right. Moreover, the observation that auditory input was sufficient to induce these shifts provides evidence for a magnitude representation which is inherently spatial, that is, in the absence of external spatial stimuli. Furthermore, we found no interaction between operation and delay, indicating that the existing attentional shifts for addition occur during early and later stages of the arithmetic process. The additional analysis of the arithmetic subset revealed a potential quadratic pattern of PSS over delay with most negative PSSs during the 750 ms—delay but most importantly, again, no interaction with the factor operation. Note however, that these results of the arithmetic subset need to be treated with caution because they are based on a mixed model that showed no significant improvement of the goodness of fit when the fixed effects of delay and the delay × operation interaction were added to the model (NULL-effect).

Interestingly, in the full data set, no shift of attention to the left was observed for subtraction problems which is reminiscent of the finding that an fMRI classifier trained to distinguish rightward and leftward saccades could generalise to addition tasks but not to subtraction tasks (Knops, Thirion, et al., 2009). The observation of significant differences in the accuracy of addition and subtraction problems points towards at least quantitative processing differences between the operations. Considering the evidence of attentional shifts to the left when much simpler subtraction problems had to be solved (Masson & Pesenti, 2014; Zhu et al., 2018), it is possible that we failed to observe effects for subtraction because the arithmetic problems we used were too difficult. In addition, from a methodological point of view, the difference in performance between addition and subtraction might have also impeded data analysis because only TOJ responses of arithmetically correct trials were used for the curve fitting procedure. Therefore, for subtraction, less TOJ responses could be analysed compared to addition, conceivably causing less informative results.

In order to address these issues, we conducted a second experiment to further manipulate task difficulty by the carry/noncarry property of the arithmetic problems. (Note, that in Experiment 1, half of the arithmetic problems were carry problems and the other half were noncarry problems, but this property was never manipulated in the way that it became a separate experimental factor.) Problems that involve carrying (e.g., 67 + 25 = 92) are more difficult than problems that do not involve carrying (e.g., 62 + 25 = 87), probably due to additional subroutines which involve higher cognitive control (e.g., Deschuyteneer et al., 2005; Fürst & Hitch, 2000). It is possible that these additional processing steps of carry problems lead to ambiguous movements on the MNL, thereby inducing smaller or even no attentional shifts to the right (left) for addition (subtraction). Noncarry problems, on the contrary, would induce unidirectional movements on the MNL and unequivocal attentional shifts to the right (left) for addition (subtraction) problems. Therefore, it is conceivable that in Experiment 1, putative effects for subtraction were masked because both carry (more difficult) and noncarry (less difficult) stimuli were used. Lindemann and Tira (2011) let participants match the amount of dots to the result of an addition or subtraction problem by turning a knob. In noncarry problems (but not in carry problems), they observed a general overproduction of dots, but this was larger for addition problems compared to subtraction problems which they interpreted as evidence for an OM effect. According to the authors, the absence of an effect in carry problems, however, was due to a greater reliance on verbal processing strategies and a decrease in activations of magnitude representations within this task type which would be in line with the hypothesis presented above. In Experiment 2, we expected significantly more arithmetic errors and longer voice onset times in the carry condition than in the noncarry condition due to the difference in difficulty between these task types. More importantly, within the attentional shift framework (Hubbard et al., 2005; McCrink et al., 2007), a main effect of carry on spatial biases is expected due to the additional subroutines in carry problems which lead to ambiguous movements on the MNL.

The heuristics account does not make spatial predictions (McCrink et al., 2007; McCrink & Wynn, 2009). Only the introduction of an additional association of numerical magnitude with space (large-right and small-left) would allow to deduce some predictions. But even under this assumption, it would predict no differences between the carry conditions within an operation, because addition would always imply "more"/right and subtraction would always imply "less"/left. The heuristics-via-spatial-shifts account (McCrink & Hubbard, 2017), however, would predict a stronger reliance on this heuristic when attentional load is high. If the additional processing steps in carry problems increase attentional load, this account would predict stronger biases within the carry condition compared to the noncarry condition.

Finally, because it is plausible that the use of carry and noncarry problems in Experiment 1 also masked the interaction of operation and delay, we continued to vary the delay between the arithmetic task and the TOJ task in Experiment 2.

Experiment 2

Method

Participants. A total of 76 German-speaking students from the Humboldt-Universität zu Berlin took part in the experiment in exchange for course credit or monetary compensation (8€/hr). For 13 participants, the experiment was

Table 4. Stimulus characteristics of arithmetic problems of Experiment 2.

Carry	Operation	Problemsize		Operand 1		Operand 2		Result	
		M	*SD*	Min	Max	Min	Max	Min	Max
Noncarry	Addition	156.4	30.71	31	85	13	45	45	98
	Subtraction	156.23	29.7	45	98	13	45	31	85
Carry	Addition	156.38	26.25	32	79	13	39	51	97
	Subtraction	155.88	25.82	51	97	13	45	32	79

terminated due to a termination criterion (see below), three subjects had to be excluded from the final sample due to experimental errors and two subjects were not included because their mean TOJ performance was 2 *SD* below the sample mean (cutoff = .72). The final sample consisted of 58 participants (36 female, 1 person who chose not to state their gender; M_{age} = 24.62 years, *SD* = 5.0 years; range, 18–34 years; 51 right-handed), 30 and 28 participants belonged to the carry and noncarry group, respectively. Each of the participants had normal or corrected-to-normal vision and hearing. The experiment was noninvasive and all procedures were carried out in accordance with the ethical standards established by the Declaration of Helsinki and approved by the Ethics Committee of the Humboldt-Universität zu Berlin. Informed consent was obtained in written format from all individual participants included in the study.

Design. For the purpose of the second experiment, the factors operation (baseline, addition, and subtraction), SOA (110, −70, −40, −20, 20, 40, 70, 110 ms) and delay (bins around 250, 750, and 1,500 ms) were varied within-subjects. The factor carry (carry and noncarry) varied between subjects.

Stimuli. For the stimuli of the second experiment, the same criteria as in Experiment 1 were employed. The mean problem size was matched across operation × carry conditions ($M \approx 156$). The range of the second operand changed (13–49, Table 4) in comparison to the first experiment (21–49, Table 1). A total of 192 trials (8 SOA levels × 3 delay levels × 8 repetitions = 192) per carry condition and operation were created resulting in a pool of 384 tasks for each carry condition, so that during the experimental block of the experiment, no arithmetic problem was presented twice. For practise trials, a random set of 24 stimuli was selected out of the full list of arithmetic problems (see S2 in the Supplementary Material for the complete list of arithmetic problems of Experiment 2).

Again, voice onsets (M_{Onset} = 157 ms, SD_{Onset} = 30 ms) and offsets (M_{Offset} = 1,366 ms, SD_{Offset} = 144 ms) of the number words sound files were kept similar between the operands used and again all sound files had the same length of 1,500 ms (*SD* < 1 ms).

Apparatus. In the second experiment, the same experimental setup was used as in Experiment 1.

Task and procedure. The procedure was almost completely identical to that reported in the first experiment except that within the arithmetic TOJ task each participant was either presented with 384 carry or noncarry arithmetic problems. Moreover, in the arithmetic TOJ task, participants were presented with 12 practise trials and then were asked whether they would like to start the actual experiment or practise further. In case of the latter, another 12 practise trials were presented. Furthermore, the response deadline for the arithmetic task was reduced to 7.5 s, that is, when 7.5 s after the onset of the question mark, no verbal response was detected, participants were prompted to respond faster or louder by a message on the screen for 1,600 ms before the ITI. Finally, in the second experiment, arithmetic accuracy was checked online via the experimenter. In Experiment 1, participants with arithmetic accuracies below 50% were excluded because only the TOJ responses of accurate trials can be analysed further. To save resources and to only obtain data that could eventually be used in the analysis, we introduced the following termination criterion in Experiment 2: If after four blocks (i.e., 128 trials) arithmetic accuracy was below 50%, the experiment was terminated and the participant was partially compensated. The participant only received feedback about their performance if the accuracy was below 50%. If this was not the case, the experiment simply continued.

Data analysis. Data was analysed in the same way as in Experiment 1. In case of the second experiment, 99.86% of the baseline in the carry part of the experiment and 99.68% of the baseline in the noncarry part of the experiment, 84.43% of the carry-addition trials, 85.92% of the noncarry-addition trials, 80.23% of carry-subtraction trials, and 81.64% of noncarry-subtraction trials were used for the curve-fitting procedure (i.e., trials, where a response to the TOJ task was given and in case of the arithmetic task, where the arithmetic task was responded to accurately). Nine percent of the whole PSS data set were discarded due to an unacceptably poor goodness of fit of the fitted curves where the PSS values were extracted from (cf. Kingdom & Prins, 2016).

Results

Arithmetic accuracy. Mean accuracy of the reduced sample (see above) was moderate (*M* = .84, *SD* = .09). For the carry

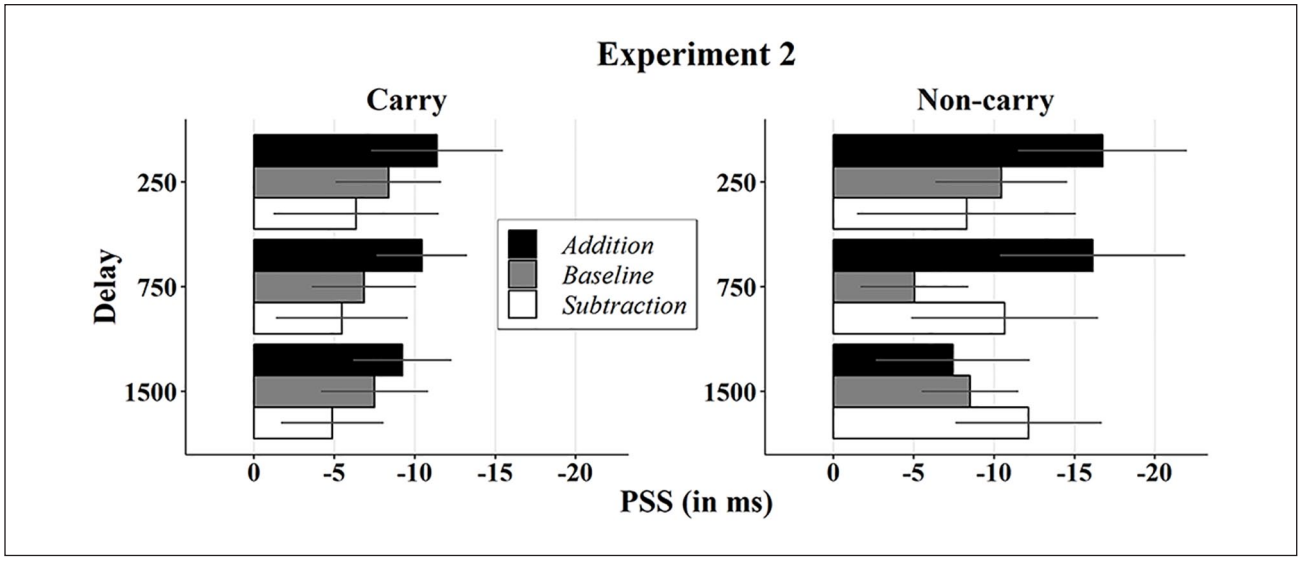

Figure 3. Mean PSSs (points of subjective simultaneity) for each operation, delay, and carry condition of Experiment 2.
Error bars represent the standard error of the mean (SEM).
PSS values constitute the time lags between the lateralised TOJ stimuli that are needed for them to be perceived as being presented simultaneously. Positive PSSs indicate attention in the left visual field and negative PSSs indicate attention in the right visual field.

condition, mean accuracy was .86 (SD=.09) for the addition condition and .8 (SD=.09) for the subtraction condition. For the noncarry condition, mean accuracy was .88 (SD=.09) for the addition condition and .83 (SD=.11) for the subtraction condition. A mixed ANOVA was applied to the arcsine-transformed accuracy rates. Operation was treated as within-subject factor and carry was treated as between-subject factor. The ANOVA revealed a main effect of operation, $F(1, 56)=36.16$, $p<.001$, but no main effect of carry, $F(1, 56)=2.01$, $p=.161$, and no interaction, $F(1, 56)=.001$, $p=.971$.

Arithmetic voice onset times. Mean voice onset to the arithmetic task was 1.48 s (SD=1.07). For the carry condition, mean voice onset time was 1.8 s (SD=1.07) for the addition condition and 2.09 s (SD=1.27) for the subtraction condition, respectively. For the noncarry condition, mean voice onset time was .96 s (SD=.65) for the addition condition and 1.04 s (SD=.76) for the subtraction condition, respectively. A mixed ANOVA was applied to the voice onset times with operation as within-subject factor and carry as between-subject factor. It revealed a main effect of operation, $F(1, 56)=21.44$, $p<.001$, a main effect of carry, $F(1, 56)=13.94$, $p<.001$, and a significant interaction, $F(1, 56)=6.28$, $p<.05$. Post hoc tests revealed that the difference of voice onset times between addition and subtraction was larger in the carry condition compared to the noncarry condition, $t(43.22)=2.56$, $p<.05$.

PSS. Averaged over all delays and carry-groups, the mean PSS was −7.68 ms (SD=14.91) for the baseline TOJ task, −11.91 ms (SD=20.33) for the addition condition, and −8.19 ms (SD=22.88) for the subtraction condition (See Figure 3 for a detailed depiction of mean PSSs).

We ran LMMs to assess the impact of the carry property, operation, or delay on the PSS.[6] To determine the random-effects structure, we compared a baseline model that contained all fixed effects and a random effect of subjects (1|subject) to a second model which additionally contained random slopes for operation (i.e., allowing the effect of operation to vary across subjects; operation|subject). This significantly improved the fit of the model, $\chi^2(5)=54.37$, $p<.001$. However, the models that also contained random slopes for delay (delay|subject) or the interaction between operation and delay by subjects (operation*delay|subject) did not converge. Hence, the final random-effects structure included random intercepts over subjects and random slopes for operation (operation|subject).

To determine the fixed-effects structure, the random-effects structure was kept constant and the fixed effects were added incrementally. Here, the inclusion of the effect of operation significantly improved the fit, $\chi^2(2)=6.54$, $p<.05$. However, further including the effect of delay, carry, or the interaction terms did not improve the fit of the model significantly (all p values $>.05$). This indicates that the type of operation had a significant effect on the individual PSSs.

Contrast-analysis of the final model (PSS ~ operation + operation|subject) that included only operation as fixed factor (contrast coded with the TOJ baseline condition set as baseline), revealed the trend that PSS values were more negative when the participants were solving addition

Table 5. Estimated parameters and statistics of mixed-effects modelling of PSS data from Experiment 2.

Predictors	PSS				
	Estimates	*SE*	*CI* (95%)	*T*	*p*
(Intercept)	−7.68	1.94	[−11.48,−3.87]	−3.96	**<.001*****
Addition	−4.14	2.09	[−8.25,−.04]	−1.98	.053
Subtraction	−.55	2.30	[−5.06, 3.97]	−.24	.813

PSS: point of subjective simultaneity; CI: confidence interval.
***$p < 0.001$.

Table 6. Estimated parameters and statistics of mixed-effects modelling of PSS data of the arithmetic task subset from Experiment 2.

Predictors	PSS				
	Estimates	*SE*	*CI* (95%)	*T*	*p*
(Intercept)	−9.64	3.79	[−17.07,−2.21]	−2.54	**.013***
Subtraction	3.54	1.52	[.56, 6.53]	2.33	**.021***
Delay (linear)	4.66	1.88	[.98, 8.34]	2.48	**.014***
Delay (quadratic)	1.00	1.87	[−2.66, 4.67]	.54	.592
Carry	−4.37	5.35	[−14.85, 6.11]	−.82	.417
Subtraction × delay (linear)	−5.10	2.64	[−10.27, .07]	−1.93	.054
Subtraction × delay (quadratic)	−1.25	2.64	[−6.43, 3.92]	−.47	.635

PSS: point of subjective simultaneity; CI: confidence interval.
*$p < .05$.

problems compared to the baseline condition, $t(57.29) = -1.98$, $p = .053$ (see also Table 5). This indicates that in the addition condition, the left target had to be presented earlier than the right stimulus to be perceived as being presented simultaneously. Table 5 also reveals that the intercept deviates significantly from zero, $t(58.07) = -3.96$, $p < .001$. As operation is the only predictor of the final model, this indicates that the PSSs of the baseline condition differed significantly from zero. As the estimate is negative, this hints at a slightly rightward preference in the baseline condition.

Just like in Experiment 1, we decided to analyse the time course of the PSS values in the arithmetic subset separately. Again, a random-effects structure was determined, resulting in random intercepts over subjects (1|subject). Then, the fixed-effects analysis revealed a significant improvement of the model fit when the fixed effect of operation was entered, $\chi^2(1) = 5.50$, $p < .05$, but no improvement of the fit neither when the fixed effects of delay, carry or the interactions were added to the model (all p values $> .05$). In the analysis of the larger model that also contained the interaction term of operation*delay ($PSS_{arithmetic} \sim$ carry + operation + delay + operation*delay + 1|subject), the factor delay was contrast coded as polynomial to check for trends in the PSS over delay. Analysis revealed that PSS values were less negative when the participants were solving subtraction problems compared to when they were solving addition problems, $t(259.68) = 2.33$, $p < .05$; see Table 6, corroborating the effect of operation found in the full PSS data set. On top of that, PSSs appeared to decrease with larger delays in a linear fashion, $t(260.04) = 2.48$, $p < .05$. Finally, the contrast analysis also revealed the trend that the (linear) slopes for the effect of delay on PSSs differed between the addition and the subtraction condition, $t(264.67) = -1.93$, $p = .054$.

Pooled analysis of Experiments 1 and 2. To ensure the pattern of results was similar in both experiments, we pooled the PSS data from the two experiments. The factor experiment consisted of the levels Experiment 1, Experiment 2 Carry, and Experiment 2 Noncarry. The random-effects structure analysis resulted in by-subject random slopes for operation (operation|subject). The fixed-effects analysis revealed a significant improvement of the fit when the fixed effect of operation was entered, $\chi^2(2) = 12.83$, $p < .01$, and a trend when the fixed effect of experiment was added to the model, $\chi^2(2) = 5.87$, $p = .053$. Contrast-analysis of the final model ($PSS_{Exp1Exp2} \sim$ operation + operation|subject; see Table 7) then again affirmed previous analyses by showing that PSS values were more negative when the participants were solving addition problems compared to the baseline condition, $t(87.91) = -3.26$, $p < .01$.

Discussion

In the second experiment, we manipulated task difficulty by the carry/noncarry property of the arithmetic problem. The aim was to investigate whether we could observe attentional shifts both to the right for addition (i.e., replicate findings of Experiment 1) and to the left for subtraction exclusively in noncarry problems or both in carry and noncarry problems.

Table 7. Estimated parameters and statistics of mixed-effects modelling of PSS data from Experiments 1 and 2.

Predictors	PSS				
	Estimates	*SE*	*CI* (95%)	*T*	*p*
(Intercept)	−4.66	1.62	[−7.84,−1.48]	−2.87	**.005****
Addition	−6.25	1.92	[−10.00,−2.50]	−3.26	**.002****
Subtraction	−2.08	1.86	[−5.73, 1.57]	−1.12	.267

PSS: point of subjective simultaneity; CI: confidence interval.
**$p < 0.01$.

Our analysis revealed a significant effect of operation on the PSSs which can be explained by the difference between the addition and the baseline condition with addition PSSs being more negative than the baseline PSSs. This can be interpreted as a rightward shift of attention during addition processing. Hence, when it comes to the significant effect of operation, we replicated the results of the first experiment; although in the second experiment, the difference between baseline and addition operation only reached borderline significance.

Experiment 2 was designed to vary difficulty of the arithmetic problems, especially to examine whether easier subtraction problems would induce an attentional shift to the left. However, like in Experiment 1, no bias was observed for subtraction problems. More interestingly, the analysis revealed no significant impact of the carry property on the PSSs and therefore the participants' attention. As we aimed at manipulating task difficulty via the carry property of the arithmetic task, we expected significantly more errors and longer voice onset times in the carry condition than in the noncarry condition. The main effect of carry condition on arithmetic voice onset times but not on arithmetic accuracy implies that the manipulation was at least partially successful.[7] It might even be the case, that the effect of carry mainly manifests itself as an effect on response times rather than accuracies. Therefore, the absence of an effect of carry cannot fully be attributed to a failed manipulation of carry.

Considering the potential outcomes regarding the carry property, the absence of an impact of carry on PSSs in the second experiment is in line with predictions from the heuristics account (McCrink et al., 2007; McCrink & Wynn, 2009) when additionally assuming an association of numerical magnitude with space (large-right and small-left). Independent of carry condition, addition would always imply "more"/right and subtraction would always imply "less"/left. The heuristics-via-spatial-shifts account (McCrink & Hubbard, 2017) predicts a stronger reliance on this heuristic when attentional load is high. Under the assumption that additional processing steps in carry problems increase attentional load, our data do not support this account. Alternatively, our results support the notion that arithmetic problem solving relies on two (potentially parallel) processes: a rough and approximate estimation of the result that hinges on the approximate number system and a second, verbally mediated processing step where operands are explicitly manipulated according to fixed procedures (cf. Dehaene & Cohen, 1995; Dehaene et al., 1999; Rousselle & Noël, 2008). The additional verbal processing load of carry problems might have affected the second, verbally mediated processing step, but not so much the first, implicit and approximate processing step. Observing shifts of attention to the right for carry and noncarry addition problems, therefore, implies that these spatial biases are part of the approximate processing step.

Furthermore, like in Experiment 1, we found no interaction between operation and delay. However, analysing only TOJ performance in the arithmetic task subset (i.e., without baseline) revealed a linear trend for the PSSs over delay and an interaction between the slopes of the linear PSS trajectories over time and operation. Although this analysis needs to be treated with caution because it is based on a model where the goodness of fit did not improve significantly when the fixed effect of the delay × operation interaction was added to the model, it indicates that the existing attentional shift for addition peaks in early stages of the calculation process and then declines in later stages. This adds to existing knowledge of Liu et al. (2017) who found spatial biases at their second-to-longest and longest delays of 300 and 500 ms, that is, compared to our range of delays at comparably early stages.

Finally, the combined analysis of the PSS data of Experiments 1 and 2 again revealed a main effect of operation and that PSS values were more negative when the participants were solving addition problems compared to the baseline condition. This can be interpreted as a rightward shift of attention during addition processing in both experiments of this study.

General Discussion

In two experiments, we used a TOJ paradigm to investigate whether we can observe attentional shifts to the right and left during the calculation phase of complex (two-digit) addition and subtraction processing, respectively. In Experiment 2, we additionally varied the carry property of the arithmetic task to assess how task difficulty modulates the effects (carry ≙ difficult vs. noncarry ≙ easy). Finally, in

both experiments, we varied the delay between the arithmetic problem presentation and the TOJ task to investigate when arithmetically induced attentional shifts occur and how long they would persist.

Crucially, both experiments revealed that addition processing shifts attention more rightward compared to the baseline task which corroborates earlier findings (e.g., M. Li et al., 2018; Liu et al., 2017; Masson et al., 2018; Masson & Pesenti, 2014). This study adds to the existing literature in two important ways. First, we observed attentional shifts during the calculation phase of arithmetic processing which indicates that attention shifts are part of the calculation process. To our knowledge, only one other study has investigated spatial biases in the calculation phase of arithmetic problem solving using an attentional paradigm. Liu et al. (2017, Experiments 2, 3) revealed spatial biases for addition (right) and subtraction (left) during the calculation process during a verification task. Hence, the fact that the results from the active production task of this study converge with findings from a verification task of Liu et al. (2017)—at least when it comes to addition processing—strongly supports the notion of an involvement of attentional mechanisms during mental calculation.

Second, this study could show that complex, two-digit addition problems are sufficient to induce attentional shifts to the right during addition processing indicating that visuospatial attention mechanisms are recruited during the processing of complex arithmetic problems. The majority of studies investigating spatial biases induced by arithmetic processing used simple or moderate arithmetic problems (Masson et al., 2018; Masson & Pesenti, 2014; Zhu et al., 2018) and only a minority used complex tasks (Klein et al., 2014; M. Li et al., 2018; Liu et al., 2017). Interestingly, the evidence for attentional shifts both in simple and complex arithmetic processing implies that both types of processing share common functional mechanisms. These mechanisms may include approximate arithmetic processes that are implicit, quick, and unaffected by the explicit, verbally mediated calculation processes going on (cf. Dehaene & Cohen, 1995; Dehaene et al., 1999). In line with that idea, this study revealed no significant differences in spatial biases between the carry and noncarry condition, despite the carry operation involving more (verbally mediated) computational steps than the noncarry operation. Although the findings of this study might be due to a partially failed manipulation (we found no differences in accuracies between carry and noncarry problems), results from Masson and Pesenti (2014) indicate that this might not have been the main contributor to the absence of an effect of the carry condition, as they found no differences in spatial biases between the carry conditions despite significant differences in accuracies and reaction times (i.e., a successful manipulation). Hence, these observations of no differences between carry conditions also point to an approximate processing stage that is an integral component of both carry and noncarry processing. Future research needs to investigate the nature of these kinds of mechanisms and whether the fast procedural strategies which have recently been observed during simple arithmetic are part of this approximate processing step or make up another independent mechanism (Barrouillet & Thevenot, 2013; Fayol & Thevenot, 2012; Uittenhove et al., 2016).

The fact that the visuospatial attention shifts were induced by arithmetic problems that were presented auditorily indicates an inherently spatial character of the numerical representation. Consequently, our study provides further evidence for the influence of the spatial organisation of mental magnitude representations on mental arithmetic. In line with the attentional shift hypothesis, addition processing might induce internal shifts of attention to the right on the numerical representation (i.e., the MNL). Specifically, the processing of the first operand might lead to a distribution of activation on the left-to-right oriented MNL. Then, attention might be moved towards the correct result, that is, in the direction of the particular operation (rightward in case of addition). This internal representational shift could then be observed externally as a shift of attention towards the right side of space.

The second main observation of this study is the absence of spatial attention shifts to the left for subtraction (compared to the baseline). This finding is reminiscent of a study by Masson et al. (2017) who found that subtraction processing was unaffected by leftward or rightward optokinetic stimulation. Furthermore, the fMRI classifier trained on rightward and leftward saccades (accompanied by leftward and rightward shifts of attention) of Knops, Thirion, et al. (2009) was unable to correctly classify subtraction trials. This finding issues a challenge to all existing theories, as it cannot be explained by spatial or nonspatial accounts. How can we explain this absence of an effect for subtraction?

One possibility is that subtraction tasks might be solved via addition strategies, as it has already been suggested by some (Masson et al., 2017). This might have led to attentional shifts to the right which masked operation-congruent shifts to the left. Several studies investigating self-reported strategies have identified the addition strategy as a feasible strategy for subtraction tasks (Geary et al., 1993; LeFevre et al., 2006; Seyler et al., 2003). In addition to that, Campbell (2008) demonstrated a latency advantage in problems that involve a two-digit minuend when they were presented in an addition format ($7 + \_ = 13$) compared to the standard format ($13 - 7 = \_$). This suggests that (especially) large subtraction problems are often solved by addition strategies. Shifts to the right induced by addition might then have masked operation-congruent shifts to the left by differences or mixed strategies within a task (i.e., both addition and subtraction strategies within a subtraction task), within a participant (i.e., strategy use depending on the problem) or even between participants

(i.e., interindividual strategy preferences). This means that addition and subtraction might essentially employ similar mechanisms but different strategies. Therefore, the notion of addition strategies used for subtraction tasks could still comply with the attentional shift account.

Another possible explanation for the attenuated spatial bias in subtraction is that the increased working memory load (i.e., cognitive load, Lavie, 1995) in subtraction processing is associated with reduced spatial processing which might have led to less spatial biases. When looking at the accuracy of addition and subtraction problems in both experiments of this study, it becomes obvious that they differ significantly in accuracy. This is in line with other studies investigating spatial biases in moderate to complex arithmetic which have found significant differences in accuracies and or response latencies between addition and subtraction (Liu et al., 2017; Masson et al., 2018; Masson & Pesenti, 2014). In that context, several studies have stressed the important role of working memory during arithmetic processing (Cavdaroglu & Knops, 2016; DeStefano & LeFevre, 2004; Fürst & Hitch, 2000) and have indicated that especially large subtraction problems put a high load on working memory (Seyler et al., 2003). Interestingly, a couple of recent studies could demonstrate a disruption of selective spatial attention under high working memory load (Naert et al., 2018; Pratt et al., 2011; Vossen et al., 2016). For example, Vossen et al. (2016) reported delayed spatial orienting in a spatial attention paradigm when verbal working memory load was high, and Naert et al. (2018) showed that high working memory load has an impact on spatial processing, especially by slowing down the detection of left-sided targets. Two further studies shed some light on the neural underpinnings of these multitasking effects: They demonstrated a decrease in parietal activation linked to spatial attention in driving (simulation) when concurrently performing a verbal sentence comprehension task (Choi et al., 2017; Just et al., 2008). Although it remains unclear whether these observations reflect either a capacity limit of a shared resource that is distributed across tasks (shared resources account, for axample, Just et al., 2008) or the allocation of a common resource to the tasks in an alternating fashion (task-switching account, for example, Vossen et al., 2016), these findings might explain the results of this study: The evidence presented above (including this study) demonstrated that the internal representation of numbers and external spatial representation recruit similar attentional mechanisms. Hence, high cognitive load during subtraction processing might have diminished spatial attention deployment on the internal numerical representation which might have led to reduced spatial biases—or as put by Masson et al. (2017, p. 847) "It is possible that the need to apply several computation steps when solving a subtraction might, in fact, weaken rather than strengthen the recruitment of visuospatial attentional mechanisms." A recent finding clearly supports this view by showing that there is a stronger link between the use of a spatial mapping and two-digit mental addition compared to subtraction and by associating two-digit subtraction to a higher working memory load (Montefinese et al., 2017).

In that context, it is worth mentioning that the carry property of Experiment 2 was initially varied to manipulate task difficulty. The goal was to investigate whether the additional processing strategies of carry problems have an impact on spatial biases and whether we would be able to observe attentional shifts to the left for the putatively easier noncarry problems. As pointed out above, the absence of effects of carry on spatial biases in this study indicates an approximate processing step that is inherent to carry and noncarry problems. This observation might present a challenge to the notion of reduced spatial processing with increased cognitive load because the additional processing steps in carry problems would imply an increased cognitive load and therefore less spatial processing. The analysis of the arithmetic accuracy in Experiment 2, however, revealed a main effect of operation but no main effect of carry, nor an interaction. This indicates that at least in regard to arithmetic accuracy, the carry property is less impactful when it comes to the role of task difficulty compared to the operation itself. This might explain why in this study, the increased cognitive load of subtraction had an impact on spatial biases (compared to addition) but the additional load of carry (compared to noncarry) did not. Future studies need to consider the role of task difficulty and differences between the operations when designing experiments using complex arithmetic problems.

Although the attentional shift framework is the most prominent explanation, the findings of this study should also be discussed in the light of alternative theories. First, the compression account (Chen & Verguts, 2012; McCrink et al., 2007) and the heuristics account (McCrink et al., 2007; McCrink & Wynn, 2009) do not assume spatial mechanisms, making it difficult for them to explain the results of this study. However, if one additionally assumes a spatial association, then the heuristics account can explain the observed spatial shifts to the right for addition processing (due to an addition-more-right association). This account could also explain the absence of an effect of the carry property because addition would always imply "more," independent of additional subroutines of the carry procedure. However, this account cannot explain the absence of shifts for subtraction. If we assume that subtraction is solved using various strategies that may involve reformulating the problem to an addition problem (see above), the heuristics account remains mute concerning the question whether this entails a reversal of the heuristic and hence does not provide any explanation for the absence of an attentional effect during subtraction. The heuristics-via-spatial-shifts account (McCrink & Hubbard, 2017) makes some spatial assumptions but would predict a stronger reliance on a (addition-more and subtraction-less)

heuristic when attentional load is high. Assuming that the additional processing steps in carry problems increase attentional load, this account would predict stronger biases in the carry condition compared to the noncarry condition, which contradicts the data of this study. Finally, as we cannot be sure about the weights of the contributing components of the AHAB model (Shaki et al., 2018), this findings cannot support nor rule out this account.

Finally, this study also set out to investigate the time course of the attentional shifts. In this regard, the analysis of the arithmetic subset (only addition and subtraction task) in Experiment 1 revealed that independent of operation, attention shifts more rightward in the second delay of 750 ms than in the other delays of 250 and 1,500 ms. In Experiment 2, it revealed a tendency for decreasing attention shifts over time and significant differences between the slopes of these linear trajectories over delay of the two arithmetic operations. Hence, results were not conclusive regarding the time course of attentional shifts. If, as proposed above, approximate processes drive the attentional shift to the right for addition, then one would expect biases early during arithmetic processing and a decrease during later stages of the calculation process. However, our study does not provide conclusive evidence for (or against) this hypothesis.

Limitations

Regarding limitations of this study, the type of matching between the operations might have been problematic: Tasks of addition and subtraction processing were matched via the problem size which was defined as O1 + O2 + Result. However, this formula can be reduced to the following depending on operation

$$\begin{aligned}\text{Addition}: \text{Problemsize} &= O1 + O2 + \text{Result} \\ &= O1 + O2 + O1 + O2 = 2 * \text{Result} \\ \text{Subtraction}: \text{Problemsize} &= O1 + O2 + \text{Result} \\ &= O1 + O2 + O1 - O2 = 2 * O1\end{aligned}$$

This means that on average, the size of the first operand in subtraction was equal in size to the result in addition. On average, the first operand in subtraction was larger than the first operand in addition. Considering the effects from the numerical domain showing that single numerosities can induce shifts of attention, we cannot fully exclude that the effects of this study might be caused by the numerosities of the operands. However, under this assumption, the attentional effect should have been stronger in subtraction compared to addition. We observed the contrary, arguing against the idea that the attentional shift was driven by the size of the operands. Moreover, the results of Liu et al. (2017) add to this argument. They found spatial biases both when operands (Experiment 2) and when the second operand and the proposal (Experiment 3) were matched.

Second, with the design of this study, we cannot be fully certain what type of processing caused the spatial shift of attention during addition processing. Despite aiming for measuring processes during the calculation phase of the arithmetic problem-solving process, we cannot be sure that participants were calculating exactly during the time window of the TOJ task. Time-resolved recording of concurrent brain signatures (e.g., EEG) might be helpful in future experiments to determine interindividual time windows of arithmetic processes.

Finally, although it could be shown by manipulating spatial attention and thereby affecting mental arithmetic, that spatial biases are part of the solving process (and no mere epiphenomena) (Masson et al., 2017; Masson & Pesenti, 2016; Mathieu et al., 2016; Wiemers et al., 2014), there is still the possibility that some nonspatial or nonarithmetical attentional mechanisms are responsible for the observed attentional shift to the right for addition. Hence, further research will need to clarify whether general non-spatial attentional mechanisms, like the cognitive load induced by the dual-task situation of the arithmetic TOJ task of this study, have an impact on spatial biases.

Conclusion

In conclusion, this study showed that the active production of solutions to addition problems induces attentional shifts to the right, suggesting that visuospatial attention mechanisms are recruited during mental (symbolic) calculation. Thereby, this study extends earlier findings by showing that simple and complex arithmetic, as well as carry and noncarry arithmetic might share a mechanism (e.g., an approximate processing step) that is responsible for these biases. Interestingly, we did not observe attentional shifts to the left for subtraction processing. Further research is needed to clarify whether this is due to the involvement of addition strategies or due to an increased working memory load that attenuates spatial processing in subtraction.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: M.G. received no financial support for the research, authorship, and/or publication of this article. Participants' remuneration was supported by a DFG grant (grant no. KN 959-2) to André Knops.

ORCID iDs

Maria Glaser https://orcid.org/0000-0002-3934-9299
André Knops https://orcid.org/0000-0001-6357-9951

Supplementary material

The supplementary material is available at: qjep.sagepub.com

Notes

1. For consistency, the term "simple" will be used for arithmetic problems that involve single digits, that is, one-digit ± one-digit. The term "moderate" will be used to refer to problems of the form: two-digit ± one-digit. To refer to problems that involve only two-digit operands (two-digit ± two-digit), we use the term "complex."
2. Note that for the baseline TOJ task, there was actually no need for a delay variation because no stimulus except for the fixation dot was presented before. However, we decided to include the variation here because inferential statistics necessitate a fully crossed design, that is, a realisation of all factors (e.g., delay) in all levels of the factors (e.g., operation: baseline, addition, and subtraction).
3. For simplification, the term "carry" will be used both for the "carrying" process in addition and "borrowing" process in subtraction.
4. Due to experimental error, some of the subtraction tasks contained results that were multiples of 11. As analysis revealed no significant differences in accuracies compared to the remaining subtraction stimuli, $t(30) = .42$, $p = .68$, these tasks and the according trials were kept for later analysis.
5. The need for a random-effects model was tested by comparing a model that contained only the intercept to a model that additionally contained a random effect of subjects (i.e., an adjustment to the intercept grouped by subjects). This test revealed a significant variance in intercepts across subjects, $\chi^2(1) = 113.22$, $p < .0001$.
6. Again, the need for a random effects model was tested by comparing a model that contained only the intercept to a model that additionally contained random intercepts by subjects. This test revealed a significant variance in intercepts across subjects, $\chi^2(1) = 251.09$, $p < .0001$.
7. We assume that the failed manipulation of arithmetic accuracies is mainly due to the termination criterion described in the Methods section of Experiment 2: Via that criterion, we artificially decreased variance on the bottom of the arithmetic performance spectrum. As this was done in the same way in the carry and noncarry conditions, it might have resulted in a differential reduction of variances between the carry conditions contaminating the carry-manipulation of the arithmetic accuracies.

References

Ashcraft, M. H., & Guillaume, M. M. (2009). Mathematical cognition and the problem size effect. *Psychology of Learning and Motivation*, *51*, 121–151.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Barrouillet, P., & Fayol, M. (1998). From algorithmic computing to direct retrieval: Evidence from number and alphabetic arithmetic in children and adults. *Memory & Cognition*, *26*(2), 355–368. https://doi.org/10.3758/bf03201146

Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account? *Cognition*, *128*(1), 35–44.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bowers, D., & Heilman, K. M. (1980). Pseudoneglect: Effects of hemispace on a tactile line bisection task. *Neuropsychologia*, *18*(4–5), 491–498. https://doi.org/10.1016/0028-3932(80)90151-7

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Campbell, J. I. (2008). Subtraction by addition. *Memory & Cognition*, *36*(6), 1094–1102. https://doi.org/10.3758/MC.36.6.1094

Campbell, J. I., & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition*, *86*(1), 71–96.

Casarotti, M., Michielin, M., Zorzi, M., & Umiltà, C. (2007). Temporal order judgment reveals how number magnitude affects visuospatial attention. *Cognition*, *102*(1), 101–117.

Cavdaroglu, S., & Knops, A. (2016). Mental subtraction and multiplication recruit both phonological and visuospatial resources: Evidence from a symmetric dual-task design. *Psychological Research*, *80*(4), 608–624.

Chen, Q., & Verguts, T. (2012). Spatial intuition in elementary arithmetic: A neurocomputational account. *PLOS ONE*, *7*(2), Article e31180. https://doi.org/10.1371/journal.pone.0031180

Choi, M.-H., Kim, H.-S., Yoon, H.-J., Lee, J.-C., Baek, J.-H., Choi, J.-S., . . .Chung, S.-C. (2017). Increase in brain activation due to sub-tasks during driving: fMRI study using new MR-compatible driving simulator. *Journal of Physiological Anthropology*, *36*(1), 11.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1–2), 1–42.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*(1), 83–120.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*(5416), 970–974. https://doi.org/10.1126/science.284.5416.970

Deschuyteneer, M., De Rammelaere, S., & Fias, W. (2005). The addition of two-digit numbers: Exploring carry versus no-carry problems. *Psychology Science*, *47*(1), 74–83.

DeStefano, D., & LeFevre, J. A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, *16*(3), 353–386.

Eimer, M., & Grubert, A. (2015). A dissociation between selective attention and conscious awareness in the representation of temporal order information. *Consciousness and Cognition*, *35*, 274–281.

Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition*, *123*(3), 392–403.

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*(6), 555–556.

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition*, *28*(5), 774–782.

Galfano, G., Rusconi, E., & Umiltà, C. (2006). Number magnitude orients attention, but not against one's will. *Psychonomic Bulletin & Review*, *13*(5), 869–874. https://doi.org/10.3758/bf03194011

Geary, D. C., Frensch, P. A., & Wiley, J. G. (1993). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Psychology and Aging*, *8*(2), 242–256.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, *6*(6), 435–448.

Just, M. A., Keller, T. A., & Cynkar, J. (2008). A decrease in brain activation associated with driving when listening to someone speak. *Brain Research*, *1205*, 70–80.

Kingdom, F. A. A., & Prins, N. (2016). *Psychophysics: A practical introduction*. Academic Press.

Klein, E., Huber, S., Nuerk, H.-C., & Moeller, K. (2014). Operational momentum affects eye fixation behaviour. *The Quarterly Journal of Experimental Psychology*, *67*(8), 1614–1625. https://doi.org/10.1080/17470218.2014.902976

Knops, A. (2018). Neurocognitive evidence for spatial contributions to numerical cognition. In A. Henik & W. Fias (Eds.), *Heterogeneity of function in numerical cognition* (pp. 211–232). Elsevier; Academic Press.

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science*, *324*(5934), 1583–1585. https://doi.org/10.1126/science.1171599

Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect. *Attention, Perception, & Psychophysics*, *71*(4), 803–821. https://doi.org/10.3758/APP.71.4.803

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468.

LeFevre, J.-A., DeStefano, D., Penner-Wilger, M., & Daley, K. E. (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *60*(3), 209–220.

Li, M., Liu, D., Li, M., Dong, W., Huang, Y., & Chen, Q. (2018). Addition and subtraction but not multiplication and division cause shifts of spatial attention. *Frontiers in Human Neuroscience*, *12*, Article 183.

Li, S.-X., & Cai, Y.-C. (2014). The effect of numerical magnitude on the perceptual processing speed of a digit. *Journal of Vision*, *14*(12), 18. https://doi.org/10.1167/14.12.18

Liddle, E. B., Jackson, G. M., Rorden, C., & Jackson, S. R. (2009). Lateralized temporal order judgement in dyslexia. *Neuropsychologia*, *47*(14), 3244–3254. https://doi.org/10.1016/j.neuropsychologia.2009.08.007

Lindemann, O., & Tira, M. D. (2011). Operational momentum in numerosity production judgments of multi-digit number problems. *Zeitschrift für Psychologie*, *219*(1), 50–57. https://doi.org/10.1027/2151-2604/a000046

Liu, D., Cai, D., Verguts, T., & Chen, Q. (2017). The time course of spatial attention shifts in elementary arithmetic. *Scientific Reports*, *7*(1), 921. https://doi.org/10.1038/s41598-017-01037-3

Masson, N., Letesson, C., & Pesenti, M. (2018). Time course of overt attentional shifts in mental arithmetic: Evidence from gaze metrics. *Quarterly Journal of Experimental Psychology*, *71*(4), 1009–1019. https://doi.org/10.1080/17470218.2017.1318931

Masson, N., & Pesenti, M. (2014). Attentional bias induced by solving simple and complex addition and subtraction problems. *The Quarterly Journal of Experimental Psychology*, *67*(8), 1514–1526. https://doi.org/10.1080/17470218.2014.903985

Masson, N., & Pesenti, M. (2016). Interference of lateralized distractors on arithmetic problem solving: A functional role for attention shifts in mental calculation. *Psychological Research*, *80*(4), 640–651. https://doi.org/10.1007/s00426-015-0668-7

Masson, N., Pesenti, M., & Dormal, V. (2017). Impact of optokinetic stimulation on mental arithmetic. *Psychological Research*, *81*(4), 840–849. https://doi.org/10.1007/s00426-016-0784-z

Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., & Prado, J. (2016). Running the number line: Rapid shifts of attention in single-digit arithmetic. *Cognition*, *146*, 229–239. https://doi.org/10.1016/j.cognition.2015.10.002

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics*, *69*(8), 1324–1333. https://doi.org/10.3758/bf03192949

McCrink, K., & Hubbard, T. (2017). Dividing attention increases operational momentum. *Journal of Numerical Cognition*, *3*(2), 230–245.

McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology*, *103*(4), 400–408. https://doi.org/10.1016/j.jecp.2009.01.013

Montefinese, M., Turco, C., Piccione, F., & Semenza, C. (2017). Causal role of the posterior parietal cortex for two-digit mental subtraction and addition: A repetitive TMS study. *NeuroImage*, *155*, 72–81. https://doi.org/10.1016/j.neuroimage.2017.04.058

Myachykov, A., Ellis, R., Cangelosi, A., & Fischer, M. H. (2016). Ocular drift along the mental number line. *Psychological Research*, *80*(3), 379–388. https://doi.org/10.1007/s00426-015-0731-4

Naert, L., Bonato, M., & Fias, W. (2018). Asymmetric spatial processing under cognitive load. *Frontiers in Psychology*, *9*, Article 583. https://doi.org/10.3389/fpsyg.2018.00583

Nicholls, M. E. R., Loftus, A. M., & Gevers, W. (2008). Look, no hands: A perceptual task shows that number magnitude induces shifts of attention. *Psychonomic Bulletin & Review*, *15*(2), 413–418. https://doi.org/10.3758/pbr.15.2.413

Poncelet, P. E., & Giersch, A. (2015). Tracking visual events in time in the absence of time perception: Implicit processing at the ms level. *PLOS ONE*, *10*(6), Article e0127106.

Pratt, N., Willoughby, A., & Swick, D. (2011). Effects of working memory load on visual selective attention: Behavioral

and electrophysiological evidence. *Frontiers in Human Neuroscience*, *5*, Article 57. https://doi.org/10.3389/fnhum.2011.00057

Prins, N. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. *Frontiers in Psychology*, *9*, Article 1250.

Ranzini, M., Lisi, M., Blini, E., Pitteri, M., Treccani, B., Priftis, K., & Zorzi, M. (2015). Larger, smaller, odd or even? Task-specific effects of optokinetic stimulation on the mental number space. *Journal of Cognitive Psychology*, *27*(4), 459–470.

Ranzini, M., Lisi, M., & Zorzi, M. (2016). Voluntary eye movements direct attention on the mental number space. *Psychological Research*, *80*(3), 389–398. https://doi.org/10.1007/s00426-015-0741-2

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*(2), 274–278.

Rousselle, L., & Noël, M.-P. (2008). Mental arithmetic in children with mathematics learning disabilities: The adaptive use of approximate calculation in an addition verification task. *Journal of Learning Disabilities*, *41*(6), 498–513.

Seyler, D. J., Kirk, E. P., & Ashcraft, M. H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1339–1352. https://doi.org/10.1037/0278-7393.29.6.1339

Shaki, S., Pinhas, M., & Fischer, M. H. (2018). Heuristics and biases in mental arithmetic: Revisiting and reversing operational momentum. *Thinking & Reasoning*, *24*(2), 138–156.

Stoianov, I., Kramer, P., Umiltà, C., & Zorzi, M. (2008). Visuospatial priming of the mental number line. *Cognition*, *106*(2), 770–779. https://doi.org/10.1016/j.cognition.2007.04.013

Team, R. C. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Uittenhove, K., Thevenot, C., & Barrouillet, P. (2016). Fast automated counting procedures in addition problem solving: When are they used and why are they mistaken for retrieval? *Cognition*, *146*, 289–303.

Vossen, A. Y., Ross, V., Jongen, E. M., Ruiter, R. A., & Smulders, F. T. (2016). Effect of working memory load on electrophysiological markers of visuospatial orienting in a spatial cueing task simulating a traffic situation. *Psychophysiology*, *53*(2), 237–251. https://doi.org/10.1111/psyp.12572

Wiemers, M., Bekkering, H., & Lindemann, O. (2014). Spatial interferences in mental arithmetic: Evidence from the motion–arithmetic compatibility effect. *The Quarterly Journal of Experimental Psychology*, *67*(8), 1557–1570.

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem-size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (p. 331–345). Psychology Press.

Zhu, R., Luo, Y., You, X., & Wang, Z. (2018). Spatial bias induced by simple addition and subtraction: From eye movement evidence. *Perception*, *47*(2), 143–157. https://doi.org/10.1177/0301006617738718

frontiers in Psychology

GENERAL COMMENTARY
published: 06 February 2019
doi: 10.3389/fpsyg.2019.00160

Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect

Daniele Didino[1]*, *Pedro Pinheiro-Chagas*[2], *Guilherme Wood*[3,4] and *André Knops*[1,5,6]

[1] Department of Psychology, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany, [2] Laboratory of Behavioral and Cognitive Neuroscience, Stanford Human Intracranial Cognitive Electrophysiology Program, Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, United States, [3] Department of Psychology, University of Graz, Graz, Austria, [4] BioTechMed-Graz, University of Graz, Graz, Austria, [5] CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, Paris, France, [6] University Paris Descartes, Sorbonne Paris Cité, Paris, France

Keywords: operational momentum, mental arithmetic, numerical cognition, spatial biases, approximate mental calculation, number-to-line mapping tasks

A Commentary on

Commentary: The Developmental Trajectory of the Operational Momentum Effect
by Fischer, M. H., Miklashevsky, A. A., and Shaki, S. (2018). Front. Psychol. 9:2259. doi: 10.3389/fpsyg.2018.02259

OPEN ACCESS

Edited by:
Elena Nava,
Università degli Studi di Milano Bicocca, Italy

Reviewed by:
John Opfer,
The Ohio State University, United States

***Correspondence:**
Daniele Didino
daniele.didino@gmail.com

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *07 December 2018*
Accepted: *16 January 2019*
Published: *06 February 2019*

Citation:
Didino D, Pinheiro-Chagas P, Wood G and Knops A (2019) Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect. Front. Psychol. 10:160. doi: 10.3389/fpsyg.2019.00160

Fischer et al. (2018) (henceforth: FM&S) raised theoretical and methodological criticisms against our study (Pinheiro-Chagas et al., 2018) on the development of the operational momentum effect (OM). Here, we will refute their criticisms and argue for a more precise definition of the OM as the operation-induced misestimation of arithmetic problem outcomes.

First, FM&S advocate the idea that zero-problems (e.g., 6+0) would be ideally suited to reveal OM. FM&S ask "*how does* [the attentional shift] *account explain larger OM with zero problems?*" In Pinhas and Fischer (2008) task , zero problems only required to map a number (the first operand) onto a labeled line, since these problems are solved by means of rules (i.e., N+0 = N, N−0 = N) rather than mental calculation (Butterworth et al., 2001; Campbell and Metcalfe, 2007). Therefore, FM&S's question is not valid because its premise (i.e., zero-problems produce OM) is not valid. Since zero and non-zero problems do not invoke the same strategies, merging their respective biases will not be helpful in elucidating the underlying mechanisms. The attentional shift account aims to describe the operation-specific outcome misestimations caused by mentally combining (at least) two numerosities. FM&S further argue that a stronger bias for zero problems compared to non-zero problems (Pinhas and Fischer, 2008; Shaki et al., 2018) invalidates the compression account of the OM "*because the logarithm of zero is not defined.*" This argumentation is flawed because FM&S mix up logarithm as a mathematical function (not defined for zero, indeed) with logarithm as a model (coding scheme) to describe the compressed internal scale of the representation of magnitudes (Nieder and Miller, 2003; Harvey et al., 2013). In the latter case, the logarithmic function is used as mathematical approximation of the relation between external physical magnitude and its internal representation. However, it makes no sense to assume that cortical circuits actually compute the faithful "mathematical log transformation" of a given sensory information. The intensity of external physical stimuli is internally represented via non-linear spatio-temporal neural codes (e.g., rate code, population code). Basing their criticism on the restriction of the mathematical definition of the logarithm to positive real numbers, FM&S conflate the mathematical definition with the neural and cognitive representation of magnitudes. Moreover, even assuming that the cognitive

system would actually be bound to this particular *mathematical* formulation of the relation between physical stimulus magnitude and sensation, another framework has been put forward that does define a mathematical solution of zero magnitudes. Stevens's power function (with positive real exponents smaller than 1) can provide identical predictions and is defined for zero. In sum, the fact that "*the logarithm of zero is not defined*" does not invalidate the compression account nor seems the use of zero problems ideal for investigating OM.

Second, we argued that the attentional shift account and the heuristic account provide equivalent predictions. Fischer and colleagues criticize this by stating that it is in conflict with results from McCrink and Hubbard and cite: "*… the use of heuristics is generally increased when attention is decreased*" (McCrink and Hubbard, 2017, p. 240). Our interpretation of McCrink and Hubbard's manuscript was based on the idea that these two accounts "*are actually so deeply intertwined that they are indistinguishable*" (p. 240) and on the fact that McCrink and Hubbard's findings "*can be best described with a heuristics-via-spatial-shifts account*" (p. 241).

Third, FM&S criticize that the downward (upward) movement of addends (subtrahends) would be inconsistent with "*the vertical MNL*" and ask "*why* […] *operations along a horizontal MNL* [were] *primed with vertical movements?*" We argue that these movements actually mimic our daily experience: adding objects from the top into a box (downward movement) and subtracting them from inside a box to the top (upward movements). Any effect of this supposed inconsistency between physical vertical movements of the operands and attentional movement on the MNL should have weakened, eliminated or even reversed the OM. Yet, we did not observe such interference. They also reasoned that the center-to-top movement of the subtrahends "*removed attention from the place of mentally simulating the outcome, thus impeding subtraction.*" First, this conclusion is inconsistent with findings from previous studies (McCrink et al., 2007; McCrink and Hubbard, 2017), where OM was observed despite subtrahends moving to the right (i.e., inconsistently with the *horizontal MNL*). Second, FM&S conflate mental simulation of addition and subtraction with attentional focus in external space. After all, the outcomes are estimated in the participants' minds—not in external space where no numerical information is present at that point in time.

Finally, the idea that in our previous studies "*the normal ingredients of OM are dis-ordered or diluted*" originates from the divergent definition of the OM. In line with the original definition by McCrink et al. (2007), we propose that OM emerges during mental calculation, rather than rule application or arithmetic fact retrieval, and refers to the numerical deviation in estimated outcomes of arithmetic operations (e.g., addition vs. subtraction), rather than biases resulting from mapping outcomes to a non-numerical dimension. In number-to-line mapping tasks, participants locate addition and subtraction outcomes on a labeled line (Pinhas and Fischer, 2008) or modify the length of a line proportionally to addition and subtraction outcomes (Shaki et al., 2015, 2018). These paradigms do not measure outcome deviations, but rather they require an additional transformation process where the outcome is converted into another physical dimension (number to position or length). Both tasks can be subject to strategical (e.g., use of reference points; Barth and Paladino, 2011; Slusser et al., 2013; Sasanguie et al., 2016; but see Opfer et al., 2016) or procedural biases (e.g., perceptual hysteresis). Therefore, any observed biases may arise from the additional transformation process rather than the calculation process itself. Results from procedures that analyse only the final location on a labeled line (Pinhas and Fischer, 2008) or the length of a segment (Shaki et al., 2015, 2018) must be interpreted cautiously because they are not measuring OM but biases that may well take place after the calculation process and have their origin in the transformation algorithm.

AUTHOR CONTRIBUTIONS

DD and AK wrote the manuscript. PP-C and GW provided critical revision. All authors approved the final version of the manuscript for submission.

ACKNOWLEDGMENTS

This work was supported by a grant (DI 2361/1-1) from Deutsche Forschungsgemeinschaft (DFG, German Research Council) to DD.

REFERENCES

Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x

Butterworth, B., Zorzi, M., Girelli, L., and Jonckheere, A. R. (2001). Storage and retrieval of addition facts: the role of number comparison. *Q. J. Exp. Psychol. Sect. A* 54, 1005–1029. doi: 10.1080/713756007

Campbell, J. I. D., and Metcalfe, A. W. S. (2007). Arithmetic rules and numeral format. *Eur. J. Cogn. Psychol.* 19, 335–355. doi: 10.1080/09541440600717610

Fischer, M. H., Miklashevsky, A., and Shaki, S. (2018). Commentary: the developmental trajectory of the operational momentum effect. *Front. Psychol.* 9:2259. doi: 10.3389/fpsyg.2018.02259

Harvey, B. M., Klein, B. P., Petridou, N., and Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science* 341, 1123–1126. doi: 10.1126/science.1239052

McCrink, K., Dehaene, S., and Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Percept. Psychophys.* 69, 1324–1333. doi: 10.3758/BF03192949

McCrink, K., and Hubbard, T. (2017). Dividing attention increases operational momentum. *J. Numer. Cogn.* 3, 230–245. doi: 10.5964/jnc.v3i2.34

Nieder, A., and Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157. doi: 10.1016/S0896-6273(02)01144-3

Opfer, J. E., Thompson, C. A., and Kim, D. (2016). Free versus anchored numerical estimation: a unified approach. *Cognition* 149, 11–17. doi: 10.1016/j.cognition.2015.11.015

Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003

Pinheiro-Chagas, P., Didino, D., Haase, V. G., Wood, G., and Knops, A. (2018). The developmental trajectory of the operational momentum effect. *Front. Psychol.* 9:1062. doi: 10.3389/fpsyg.2018.01062

Sasanguie, D., Verschaffel, L., Reynvoet, B., and Luwel, K. (2016). The development of symbolic and non-symbolic number line estimations: three developmental accounts contrasted within cross-sectional and longitudinal data. *Psychol. Belg.* 56, 382–405. doi: 10.5334/pb.276

Shaki, S., Pinhas, M., and Fischer, M. H. (2018). Heuristics and biases in mental arithmetic: revisiting and reversing operational momentum. *Think. Reason.* 24, 138–156. doi: 10.1080/13546783.2017.1348987

Shaki, S., Sery, N., and Fischer, M. H. (2015). 1 + 2 is more than 2 + 1: violations of commutativity and identity axioms in mental arithmetic. *J. Cogn. Psychol.* 27, 471–477. doi: 10.1080/20445911.2014.973414

Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Cerebral Cortex, 2018; 1–13

doi: 10.1093/cercor/bhy163
Original Article

ORIGINAL ARTICLE

Evidence for a Posterior Parietal Cortex Contribution to Spatial but not Temporal Numerosity Perception

Seda Cavdaroglu[1] and André Knops[2,3]

[1]Department of Psychology, Humboldt-Universität zu Berlin, 10099 Berlin, Germany, [2]CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, Paris, France and [3]University Paris Descartes, Sorbonne Paris Cité, 75005 Paris, France

Address correspondence to André Knops, CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, University Paris Descartes, 46, Rue Saint Jaques, 75005 Paris, France. Email: knops.andre@gmail.com

Abstract

Posterior parietal cortex (PPC) is thought to encode and represent the number of objects in a visual scene (i.e., numerosity). Whether this representation is shared for simultaneous and sequential stimuli (i.e., mode independency) is debated. We tested the existence of a common neural substrate for the encoding of these modes using fMRI. While both modes elicited overlapping BOLD response in occipital areas, only simultaneous numerosities significantly activated PPC. Unique activation for sequential numerosities was found in bilateral temporal areas. Multivoxel pattern analysis revealed numerosity selectivity in PPC only for simultaneous numerosities and revealed differential encoding of presentation modes. Voxel-wise numerosity tuning functions for simultaneous numerosities in occipital and parietal ROIs revealed increasing numerosity selectivity along an occipito-to-parietal gradient. Our results suggest that the parietal cortex is involved in the extraction of spatial but not temporal numerosity and question the idea of commonly used cortical circuits for a mode-independent numerosity representation.

Key words: approximate number sense, fMRI, IPS, number tuning

Humans as well as other animals are endowed with a system that allows them to approximately estimate the number of items in a visual scene (i.e., the numerosity). In humans, the horizontal portion of the intraparietal sulcus (hIPS) in posterior parietal cortex (PPC) is claimed to be the neural substrate of this approximate number system (Dehaene et al. 2004). Yet, whether numerical information from different formats (e.g., Arabic numerals vs. nonsymbolic dot arrays), modes (simultaneous: items spread in space or sequential: items spread in time), or modalities (e.g., visual or auditory) converge on a unitary, abstract representation in hIPS is debated (Cohen Kadosh and Walsh 2009).

Evidence from behavioral and neuroimaging studies in humans comes together with monkey neurophysiology findings in support of an abstract number representation. Human imaging studies repeatedly revealed BOLD increase in bilateral hIPS in numerical tasks employing different presentation formats (Piazza et al. 2007; Eger et al. 2009), modes (Castelli et al. 2006; Piazza et al. 2006; Dormal et al. 2010) and modalities (Eger et al. 2003). Electrophysiology studies reported number-selective neurons in the ventral intraparietal sulcus (VIP)—the putative IPS homolog—of nonhuman primates (NHP) that code for numerosities from 1 to 5 independent of presentation mode (Nieder et al. 2006) and modality (Nieder 2012). Behaviorally, numerosity adaptation across modalities in humans supports the idea of a generalized sense of numbers (Arrighi et al. 2014).

On the other hand, several findings question the idea that identical PPC circuits integrate numerosity information both across space and time, as suggested by Dehaene and Changeux (1993). First, being part of the dorsal stream, PPC plays a pivotal role in the processing of spatial information (Kravitz et al. 2011). In line with this, 2 studies (Shafritz et al. 2002; Xu and

Chun 2006) reported that the BOLD signal in IPS increases when an increasing number of objects are presented over space (simultaneously). However, no change in parietal activity was reported when a variable number of items were presented in the same location (Shafritz et al. 2002; Xu and Chun 2006), suggesting that PPC integrates numerical information over space but not across time. Second, evidence for shared numerosity tuning for different presentation formats and modes predominantly comes from electrophysiological studies in NHPs and small numerosities from 1 to 5 (Nieder and Miller 2004; Nieder et al. 2006; Tudusciuc and Nieder 2007). Only few studies tested the notion of a mode independent numerosity representation in humans (Castelli et al. 2006; Dormal et al. 2006; Piazza et al. 2006). Yet, none of these studies provided conclusive evidence due to confounds between numerosity and frequency (Dormal et al. 2010) or duration, or increased risk for false positive results from noncorrected data (Castelli et al. 2006). In light of these findings, it is still to be investigated if identical neural circuits in PPC contribute to the encoding of both sequential and simultaneous numerosities in a way that goes beyond common task activation.

Extant computational models of numerosity perception diverge on the notion of a labeled-line coding of numerosity (quantity is coded by the location of the activation in a population of linearly ordered neurons) but agree on a summation-coding instance (more quantity is coded by larger summed activity) and the idea that numerosity is abstracted from low-level visual features during encoding. Two prominent computational models of simultaneous numerosity extraction propose a hierarchy of number-sensitive and number-selective processing steps (Dehaene and Changeux 1993; Verguts and Fias 2004). First, spatial location of objects is coded in an object–location map. The activity in these units changes monotonically with increasing number of objects, reflecting summation coding. The summed activation is then fed into the next instance that contains number-selective units. Activity of these units decreases monotonically as numerical distance between preferred and actual numerosity increases. This dovetails with number-selective neurons in monkey area VIP (Nieder and Miller 2004). Importantly, this model suggests that as one moves up in the processing hierarchy, the importance of visual features like object-size and location should decrease and the importance of numerosity should increase. In humans, Roggeman et al. (2011) found a hierarchical organization along the occipital-to-parietal pathway for numerosities in the subitizing range (i.e., 1–5), in line with this computational model. BOLD signal in superior occipital cortex and the adjacent transition region between occipital and parietal cortex monotonically increased with numerosity. Areas in posterior superior parietal lobule (PSPL) and IPS, on the other hand, exhibited numerosity tuning such that BOLD signal decreased as numerical distance between preferred and presented numerosities increased. A more recent model of spatial numerosity perception used deep networks with 2 hidden layers that were trained to reproduce visual input numerosities (Stoianov and Zorzi 2012). As in Dehaene and Changeux (1993) and Verguts and Fias (2004), the response of units in hidden layer 2 was unaffected by non-numerical features of the stimuli such as size or density of the input images, thereby providing a computational instantiation of a visual sense of numbers (Burr and Ross 2008; Anobile et al. 2016), that emerged during unsupervised learning.

While computational studies foresee that simultaneous numerosity perception should be independent of visual features of the stimuli, Gebuis and Reynvoet (2012a, b) reported that the performance in numerical comparison tasks depends on the congruity of numerosity with visual features like object size, convex hull, and total area occupied. In line with this, monkey electrophysiology and human imaging suggest that object size and numerosity representations are intermingled in PPC (Tudusciuc and Nieder 2007; Harvey et al. 2013). Although most studies try to control for non-numerical features using multiple stimulus sets with different visual features (see Dehaene et al. 2005 for a discussion), it is very difficult, if not impossible, to decorrelate numerosity from all other sensory features (see Leibovich et al. 2016 for a discussion). Hence, more empirical data is needed to investigate whether previously observed numerosity tuning in simultaneous mode reflects the abstract numerosity information from a visual scene or results from the weighted integration of several non-numerical dimensions like area, density, and dot-size by Gebuis et al. (2016).

For sequential numerosities, on the other hand, researchers assume the involvement of an accumulator that reflects increasing numerosity with increasing activity, potentially assisted by mechanisms that keep track of the serial position of an item in a sequence (Nieder et al. 2006; Dormal et al. 2010). A vast number of NHP perceptual decision-making studies found activity in lateral intraparietal area (LIP) to be closely correlated with evidence accumulation over time (Shadlen and Newsome 2001; Hanks et al. 2015), even if their functional significance remains unclear since inactivation of these circuits does not affect decision-making performance (Katz et al. 2016). Whether in humans the accumulator instance for sequential mode coincides with the computational mechanisms for the encoding of simultaneous numerosities remains an open question.

The use of study designs that fail to disentangle domain-general processes (e.g., response selection) from numerosity processing further undermines the soundness of existing evidence for a mode-independent numerosity representation in PPC. It has long been known that the parietal cortex is involved in various aspects of task related processing ranging from working memory and attention to response selection (Koenigs et al. 2009; Malhotra et al. 2009; Shomstein 2012; Dean et al. 2012). Hence, common BOLD increase in numerical tasks does not necessarily imply that the underlying representation for different formats and modes is the same. Yet, human imaging studies using multivoxel pattern analysis (MVPA) endorse simultaneous numerosity encoding independent of response/task related processing (Eger et al. 2009, 2015; Dormal et al. 2010; Bulthé et al. 2014; Castaldi et al. 2016). Based on BOLD signal patterns from PPC these authors were able to decode the numerosities seen by the participants using MVPA. These results contrast with human imaging studies showing that parietal BOLD increase disappeared when response and task related factors are well controlled in numerical tasks (Göbel et al. 2004; Shuman and Kanwisher 2004; Cavdaroglu et al. 2015). Taken together, it still remains elusive whether identical PPC circuits integrate numerosity information in the absence of domain-general task requirements.

In this study, we investigated the neural basis of simultaneous and sequential numerosity perception to answer the questions outlined above. Specifically, we tested how simultaneous and sequential numerosities are encoded in the absence of response/task related processing using a task that probed comparison of numerosities only at random points throughout the experiment (Fig. 1). We used MVPA to inquire if there is a

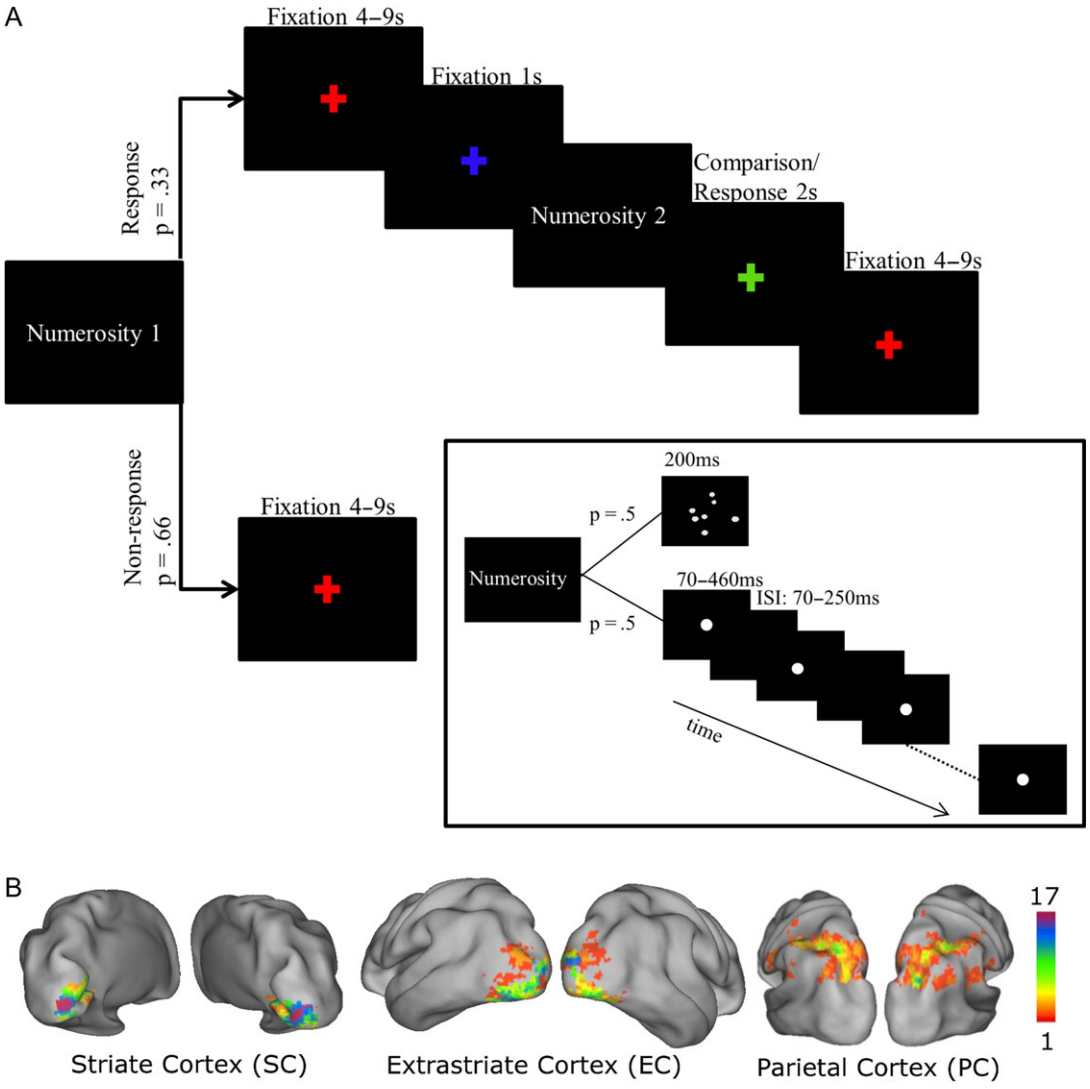

Figure 1. (*A*) Schema of the experimental procedure. Top panel depicts a response trial. After the presentation of the first numerosity (Numerosity 1), the color of the fixation-cross changed from red to blue indicating that participants were supposed to compare the numerosity before the blue fixation-cross with the numerosity coming after the fixation-cross (Numerosity 2). Participants responded by pressing the right or left button while the fixation-cross was green. The green fixation-cross was displayed for 2 s. After that, a new trial started. Bottom panel depicts a nonresponse trial. The color of the fixation-cross remained red until the next numerosity appeared. Once the trial was over, the red fixation-cross was replaced by a new numerosity and a new trial started. The inset depicts example numerosity in simultaneous (top) and sequential (bottom) mode. Both modes appeared with equal probability ($P = 0.05$). (*B*) The ROIs used for MVPA. The color coding indicates in how many participants a given voxel was activated. Although individualized ROIs were used for each site (i.e., SC, EC, and PC), there was a reasonable consistency over participants.

common coding scheme for simultaneous and sequential numerosities, which was not employed by previous studies investigating mode-independence. In addition, we used numerosities larger than 4 to see whether the previously reported gradient can also be observed for numerosities outside the subitizing range. This is important, since accumulating evidence suggests that subitizing and estimation of numerosities outside the subitizing range are 2 distinct processes, potentially hinging on different neural architectures (Piazza et al. 2002, 2011; Revkin et al. 2008; Burr et al. 2010; Anobile et al. 2012). Last but not least, in simultaneous mode, we tested how encoding of non-numerical visual features in the occipitoparietal pathway changes along with numerosity using MVPA to study the specificity of previously reported PPC tuning for numerosities.

Materials and Methods

Participants

Overall, 20 healthy right-handed participants underwent fMRI scanning after giving written informed consent. Three of them were excluded from further analysis due to excessive motion (more than the size of one voxel between subsequent volumes) or abortion of the experiment. The data from the remaining 17 participants were analyzed subsequently (8 males, mean age: 27.35, standard deviation (SD) = 4.64). All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric diseases. The study was approved by Bernstein Center for Advanced Neuroimaging (BCAN, No. 165) and the Ethical committee of Humboldt Universität zu Berlin. Participants were reimbursed 24 €.

Stimuli

Participants were engaged in a nonsymbolic numerosity processing task. The numerosities were presented on a black background using white dots. Spatial enumeration was probed by presenting simultaneous numerosities, presented as spatially scattered sets of dots (dot clouds). Temporal enumeration was probed by presenting sequential numerosities. These were scattered over time by repeatedly flashing (on–off) a single white dot in the center of the screen. Four numerosities (5, 7, 11, and 16) outside the subitizing range were used. These numerosities were chosen as they had approximately equal distance from each other on logarithmic scale. Simultaneous numerosities were created using a set of Matlab scripts as described in Dehaene et al. (2005). The scripts were adapted such that the sensory properties of dot arrays (i.e., convex hull, density, diameter, and total area) were written out during stimulus creation. Sequential numerosities were created using the method described in Cavdaroglu et al. (2015).

More specifically, non-numerical sensory features of simultaneous numerosities (i.e., dot-arrays) were controlled by 2 sets. In one set, the dot-size was kept constant whereas in the other set total area was kept constant. This way, the intensive (e.g., dot size and interitem spacing) and extensive (e.g., total luminance and total area) features of stimuli were balanced over the whole stimulus set (see Dehaene et al. 2005 for a discussion).

The non-numerical features of sequential numerosities ("flickers") were controlled in 4 sets. Single dot duration and total duration increased with numerosity in set 1 and decreased with numerosity in set 4. The interval between dots (ISI) increased with numerosity in set 2 and decreased with numerosity in set 3. Frequency (numerosity divided by total duration) increased with numerosity in sets 3 and 4 and decreased with numerosity in sets 1 and 2. Hence, participants could not rely on a single sensory cue (i.e., duration, frequency, or ISI) to extract numerosity. The individual dots were presented for a maximum duration of 270 ms to prevent counting. Only in set 4 we used dot durations longer than 270 ms as well. It was not possible to create a set of trials where total duration decreases with numerosity otherwise. This threshold was chosen based on previous studies which showed that participants cannot rely on verbal strategies (e.g., counting) within this time frame (Piazza et al. 2006; Tokita and Ishiguchi 2011). Random jitters were introduced in sequential numerosities to prevent periodicity that may lead to the perception of rhythms. The length of the jitter depended on the single dot duration. It was calculated such that after the subtraction of that jitter, the duration of the single dot was 40 ms (i.e., jitter = [dot duration – 40 ms]). This procedure guaranteed that 1) each individual dot remained distinguishable from the previous or subsequent dot and 2) when the duration of a single dot was longer than 270 ms, participants could not reliably count because the remaining stimuli in the sequence would still appear at a sufficiently high presentation rate to prevent counting. The size of dots was constant for a given numerosity sequence and was chosen randomly such that it matched the total area occupied by each dot-array in simultaneous stimuli. This ensured balanced illumination between simultaneous and sequential numerosities.

Stimuli were generated and presented using Matlab (MathWorks) and Psychtoolbox (Brainard 1997; Pelli 1997; Kleiner et al. 2007) and were projected with an LCD projector (60 Hz frame rate) onto a translucent screen in the bore of the scanner and viewed through a mirror mounted on the head coil. The duration of each dot and ISI in sequential stimuli was calculated as multiples of the refresh rate of the monitor (60 Hz) and the presentation of all the dots was synchronized with vertical refresh of the projector.

Experimental Task and Design

To separate decision and response related activations from numerosity perception, participants responded only in one-third of trials (henceforth "response trials"). In response trials, 2 numerosities from the same mode (simultaneous or sequential) were presented one after the other (Fig. 1). Participants indicated via left (first) or right (second) button press which of the 2 sequentially presented numerosities was numerically larger. The numerosities in a given response trial differed by 25% to balance difficulty across numerosities. That is, while the first numerosity in a given response trial was drawn from the set comprising 5, 7, 11, or 16, the second numerosity in that trial could be either 25% smaller or larger than the first numerosity. For example, when the first numerosity in a given trial was a temporally scattered sequence of 7 dots flashed in the center of the screen (i.e., sequential numerosity 7), the second numerosity in that trial would have comprised either 5 or 9 dots flashed in the center of the screen. In the remaining two-thirds of trials, no response was required (henceforth "nonresponse trials").

Upon presentation of a given numerosity, the participants did not know whether they would have to make a comparison with that numerosity later. This information was conveyed via the color of the fixation cross only after they were exposed to the numerosity. If the color of the fixation cross that followed the numerosity changed from red to blue (i.e., response trial), participants had to compare it with the upcoming numerosity. If the fixation-cross remained red until the next numerosity appeared, they were instructed to forget the previous numerosity and concentrate on the new one (i.e., nonresponse trials, a new trial begins if the color of the fixation-cross remains red). This way, we encouraged participants to pay attention to the numerical dimension of stimuli throughout the experiment without having any comparison or response related confounds in numerosity perception in nonresponse trials (Fig. 1).

The experiment had a fast event-related design. The timing of stimuli was optimized using simulation with fMRI design software (efMRI V9) and a stochastic design (http://archive.is/rhI2t). This type of design allows for shorter scanning periods with greater statistical power than deterministic designs (i.e., fixed ISI) or purely random ISIs (Dale 1999; Friston et al. 1999). The order of conditions and the length of the interstimulus interval (ISI) were determined using an exponential function (Dale 1999). Specifically, the ISI was randomized from an exponential distribution, taking into account the minimum ISI of 4000 ms, maximum ISI of 9000 ms and an average ISI of 6000 ms (Friston et al. 1999). The time of the jittered fixation-cross was adjusted accordingly. Double-Gamma HRF emulation was used to emulate the SPM hemodynamic response function (HRF). Five conditions were passed into the software to get the optimized presentation time for simultaneous nonresponse, simultaneous response, sequential nonresponse, sequential response, and null event (i.e., fixation) trials. Trials were randomly distributed between 4 numerosities used in the experiment (i.e., 5, 7, 11, and 16) within each condition.

The duration of null events was fixed at 1.4 s, which was the average duration of all trials. Simultaneous numerosities were

presented for 200 ms to avoid eye movements and counting. Sequential numerosities had a total duration between 630 and 4870 ms. The duration of the fixation-cross (i.e., ISI) after each response trial varied between 4071 and 8872 ms, and was identical for sequential and simultaneous numerosities. The duration of the fixation-cross between the first and second numerosity in response trials was chosen randomly from ISI durations used in between each stimulus trial. The experiment consisted of 8 blocks in total. In each block, there were 64 nonresponse trials (half simultaneous), 32 response trials (half simultaneous), and 8 null events which lasted in total ~9 min. Hence, the total duration of the main fMRI task was ~72 min.

In nonresponse trials, an equal number of trials were drawn from each stimulus set. That is, in 1 block, there were 16 numerosities from each simultaneous set and 8 numerosities from each sequential set (i.e., 16 × 2 sets = 32 simultaneous numerosities and 8 × 4 sets = 32 sequential numerosities). In response-trials, an equal number of trials were drawn from each stimulus set. Importantly, the first and second numerosities were always drawn from different sets to make sure that participants could not rely on non-numerical sensory features while they were comparing the 2 numerosities. Furthermore, both response and nonresponse trials had an equal amount of trials per numerosity and an equal number of stimuli were drawn from each set.

Localizer Task

To independently determine functional ROIs for multivariate analysis, a 12 min functional localizer was created using Psychtoolbox (Brainard 1997; Pelli 1997; Kleiner et al. 2007) and presented after the numerical task. The task is an adapted version of the localizer described in Cavdaroglu et al. (2015) with an additional visual motion localizer appended. It consisted of reading, date recall, mental subtraction, object grasping, house roof color naming, saccade formation, motion, and rest conditions.

Reading, subtraction and date recall conditions were presented using an optimized rapid event-related design (see Cavdaroglu et al. 2015 for all the details about timing). Ten simple sentences ("Bears are fond of salmon and honey"), subtraction problem sentences ("Calculate eleven minus five") (translated from Pinel et al. 2007) and novel date recall sentences ("The date of New Year's Eve is ____"), were intermixed with 10 rest periods, for a total of 40 trials. In all 3 conditions, participants were instructed to silently read the sentences and mentally generate an answer when necessary (subtraction and date recall) without giving an explicit response. In the rest condition, a blank screen with a central fixation dot was presented.

Object grasping, saccades and roof color naming blocks were presented using an optimized epoch design. Black and white illustrations of graspable objects (e.g., scissor, cup; courtesy of Philippe Pinel), multidirectional (360°) saccade targets and photographs of houses with different roof colors, were presented. In object grasping trials, participants were instructed to mentally imagine grasping the objects with their dominant (right) hand. In saccade trials, 3 saccades were made through following a saccade target (+). In house roof color naming, participants were instructed to silently name the roof color. All trials were alternated with jittered fixation trials, with a minimum ISI of 4000 ms.

The visual motion localizer was added as we found increases in the BOLD signal in MT during processing of nonsymbolic numerosities in a previous experiment (Cavdaroglu et al. 2015). The motion localizer was based on the MT localizer described in Takemura et al. (2012). In total, 200 white dots (0.25°) were presented on a black background with a circular aperture of 20° diameter centered at the fixation point. In a 12 s motion block, the dots moved inwards and outwards at a speed of 8°/s. The motion block was followed by a 12 s stationary block. Each dot lasted for 10 frames and it was replaced at a random position once the life time ended. Moreover, the dots that crossed the borders of the circle during outward motion were replaced at random locations within the circle as well as the dots that reached the central fixation during inward motion. Nine pairs of motion and nonmotion blocks were run in total.

fMRI Data Acquisition

Functional images were acquired at the Berlin Center for Advanced Neuroimaging (BCAN) with a 3 T Siemens TIM Trio scanner (Siemens, Erlangen), using a 12-channel head coil. Before the experiment, a T1-weighted image (MPRAGE) was collected as high-resolution anatomical reference (TR = 1900 ms, TE = 2.52 ms, flip angle = 9°, FOV = 256 mm × 256 mm × 192 mm, resolution = 1 mm). T2*-weighted gradient-echo echo-planar images were collected during the experiment (TR = 2500 ms, TE = 25 ms, flip angle = 82°, FOV = 190 mm × 190 mm, resolution = 2.5 mm, slices = 42 slices with a 20% distance factor; interleaved acquisition order). Finally, T2*-weighted gradient-echo echo-planar images were collected during the localizer task (TR = 2000 ms, TE = 30 ms, flip angle = 78°, FoV = 192 mm × 192 mm, resolution = 3 mm, slices = 33 with a 25% distance; descending acquisition order). The first 2 images in each series served to guarantee stable magnetization and were not recorded. After the acquisition of the anatomical image as well as before the localizer, a magnetic field mapping sequence was run to correct for inhomogeneities in the magnetic field (TR = 400 ms, TE = 5.19 ms/7.65 ms, flip angle = 60°, FOV = 192 mm × 192 mm, resolution = 3 mm, slice gap = 25%, slices = 33).

fMRI Data Analysis

Images were analyzed using Statistical Parametric Mapping software (SPM8; Wellcome Trust Center for Neuroimaging; http://www.fil.ion.ucl.ac.uk/spm/). Functional images were first reoriented to the anatomical reference and then corrected for inhomogeneities in the magnetic field. Subsequent preprocessing included slice-timing correction (where middle image in the time series was taken as the reference), spatial realignment and unwarping, coregistration to the unwarped mean image, segmentation, normalization to standard Montreal Neurological Institute (MNI) space, and smoothing (FWMH = 6 × 6 × 6 mm^3).

After preprocessing, a general linear mode (GLM) based on numerosity was defined using a canonical HRF. The numerosity model included a regressor for each numerosity (5, 7, 11, and 16); separately for simultaneous/sequential modes and response/nonresponse trials as well as a regressor for null events. In response trials, the first and second numerosities were also modeled separately. Thus, the numerosity-GLM had 33 regressors in total along with 6 movement parameters from preprocessing to capture signal variations due to head motion. The event-related numerosity regressors were locked to the onset of the numerosity presentation. The null events were used as baseline for the contrasts in univariate analysis.

All the contrasts reported in this paper were corrected with FDR at $P = 0.05$ on cluster level, $P = 0.001$ on voxel level with minimum cluster size 15 using xjView (http://www.alivelearn.net/xjview).

Unsmoothed images from the preprocessed data were used for multivariate analysis to preserve the maximal amount of spatial information. Pattern classification was performed using

linear support vector machines (SVM) on The Decoding Toolbox (Hebart et al. 2015) with regularization parameter [C] set to 1 (LIBSVM 3.12, Chang and Lin 2011). A one-block-out cross validation scheme was employed. That is, one experimental block was left as the test data and the remaining 7 blocks were used to train the classifier. The left-out block was iterated over all 8 blocks and an average decoding accuracy estimate was obtained at the end.

To investigate how numerosity-specific the pattern recognition results were in simultaneous trials, the stimuli were reorganized for each sensory feature (i.e., convex hull, density, diameter, and total area) such that there were 4 categories for the respective sensory feature. That is, instead of labeling the dot arrays based on the number of dots (i.e., 5, 7, 11, and 16), we labeled them with the corresponding category (e.g., based on how big the total area is) in 4 different models that were based on the convex hull, total area, density, or diameter of the dots in the stimulus. Since perfect balancing of the number of trials in each category was not possible in all cases, we corrected for the remaining numerical imbalances between different categories by using the balanced accuracies (that are provided by The Decoding Toolbox) during the statistical testing of multivariate analysis results for sensory features.

Analysis of the Localizer Data and ROI Extraction

Preprocessing of the localizer data was identical to the functional data besides the reference slice used for slice-timing correction (first image) and the order of slice-timing correction and spatial alignment (here, spatial alignment and unwarping was performed before slice-timing correction due to differences in slice acquisition order). After preprocessing, the localizer task was modeled by a canonical HRF and a GLM was defined that included a regressor for each condition (houses, objects, dates, reading, subtraction, saccades, motion, and fixation) and 6 motion parameters from preprocessing to capture signal variations due to head motion.

For MVPA, bilateral parietal ROIs were extracted from the combination of *F*-contrast (main task) on a subject-by-subject basis and subtraction minus reading contrast (localizer task) on group level within a mask of parietal cortex (WFU PickAtlas, Maldjian et al. 2004, 2003). Within these masks, the 500 most active voxels were chosen as subjective ROIs (Fig. 1*B*). The group level "subtraction minus reading" contrast was used, as there were not enough voxels for most subjects on individual level.

Finally, to investigate how the representation of numerosity and other sensory features evolve along the visual hierarchy, we created 2 ROIs separating striate from extrastriate areas of the visual cortex. The first visual ROI was a combination of "houses minus rest" contrast (localizer task) on subject-by-subject basis within a mask of "occipital cortex minus striate cortex" (Anatomy Toolbox, Eickhoff et al. 2005, 2006, 2007; Fig. 1*B*). The second visual ROI was a combination of "houses minus rest contrast" (localizer task) on a subject-by-subject basis within a mask of striate cortex (V1, WFU atlas, Maldjian et al. 2003, 2004; Fig. 1*B*). Hence, the first visual ROI included the extrastriate areas whereas the second visual ROI included only the striate cortex. For both ROIs, the 500 most active voxels within these masks were chosen as subjective ROIs.

Tuning Curves

Similar to the analysis of single neuron numerosity tuning (Nieder 2012; Viswanathan and Nieder 2013), we determined for each participant which numerosity a given voxel responded to maximally by searching for the maximal beta weight from the above described model containing all numerosities in simultaneous mode. Since one of the aims of this study was to analyze whether the extraction of numerosities outside subitizing range is organized along an occipital-to-parietal gradient, we defined 6 nonoverlapping ROIs in each hemisphere that covered the entire dorsal pathway from striate to parietal areas (see left inset in Fig. 2*D*). All ROIs were boxes including 768 voxels (640 mm^3), collapsed across the 2 hemispheres. The lower 3 ROIs had an extension of 16 × 12 × 4 voxels centered on the following coordinates (*x*, *y*, *z*; from occipital to parietal): (0, −90, 6); (0, −90, 18); (0, −80, 30). The upper 3 ROIs had an extension of 8 × 12 × 4 voxels in each hemisphere, centered on the following coordinates (*x*, *y*, *z*; from occipital to parietal): (±15, −70, 42); (±20, −60, 54); (±25, −50, 66).

Along this gradient, we computed the voxel-wise tuning functions and averaged across numerosities to determine the average numerosity tuning at each level of the gradient (see Serences et al. 2009 for an example). That is, we centered the numerosity-specific tuning curves on the preferred numerosity and pooled across preferred numerosities. For example, the BOLD response for numerosity 7 in voxels that respond maximally to numerosity 5 (approximate log distance: 0.15) is pooled with responses to numerosity 16 in voxels that respond maximally to numerosity 11 (approximate log distance: 0.16). The resulting numerosity tuning functions indicate the degree to which voxels in each ROI change their response as a function of numerical distance between preferred and presented numerosity. For each ROI, we computed 2 linear regressions on the numerosity tuning functions, one for negative numerical deviations from the preferred numerosity and one for positive. In a last step, we averaged the 2 regression coefficients to compute an intuitive measure of numerosity filter precision along the ROI gradient. All tuning curve analyses are based on 14 participants only, since we excluded participants who did not show at least one activated voxel in every ROI.

Results

Behavioral Results

The mean accuracy was 79.46% (SD 8.52%) for simultaneous response trials and 72.67% (SD 6.75%) for sequential response trials. In both modes, participants performed significantly above chance ($t_{(16)}$ = 14.238, $P < 0.001$ for simultaneous and $t_{(16)}$ = 13.852, $P < 0.001$ for sequential). We submitted behavioral accuracies to a repeated measures ANOVA with factors mode (simultaneous, sequential) and numerosity (5, 7, 11, and 16) and found a main effect of mode (Fig. 2, $F_{(1,16)} = 13.761$, $P = 0.002$), numerosity ($F_{(1,16)} = 16.271$, $P < 0.001$) as well as an interaction between mode and numerosity ($F_{(1,16)} = 5.034$, $P = 0.004$). Post hoc tests revealed that in simultaneous trials, the comparison accuracy for numerosity 7 was significantly higher than the accuracy for numerosity 16 ($t_{(1,16)} = 5.22$, $P = 0.001$; Bonferroni corrected). In sequential trials, comparison accuracies for numerosity 7 and 11 were significantly higher than for numerosity 5 ($t_{(16)} = 5.912$, $P = 0.009$; $t_{(16)} = 5.748$, $P < 0.001$, respectively; Bonferroni corrected) and the accuracy for 11 was significantly higher than 16 ($t_{(16)} = 4.188$, $P = 0.005$; Bonferroni corrected). Participants were significantly more accurate in simultaneous compared with sequential response trials ($t_{(16)} = 4.485$, $P < 0.001$).

fMRI Results

Univariate Analysis

The BOLD signal during nonresponse trials was captured by contrasting all numerosities against baseline, irrespective of

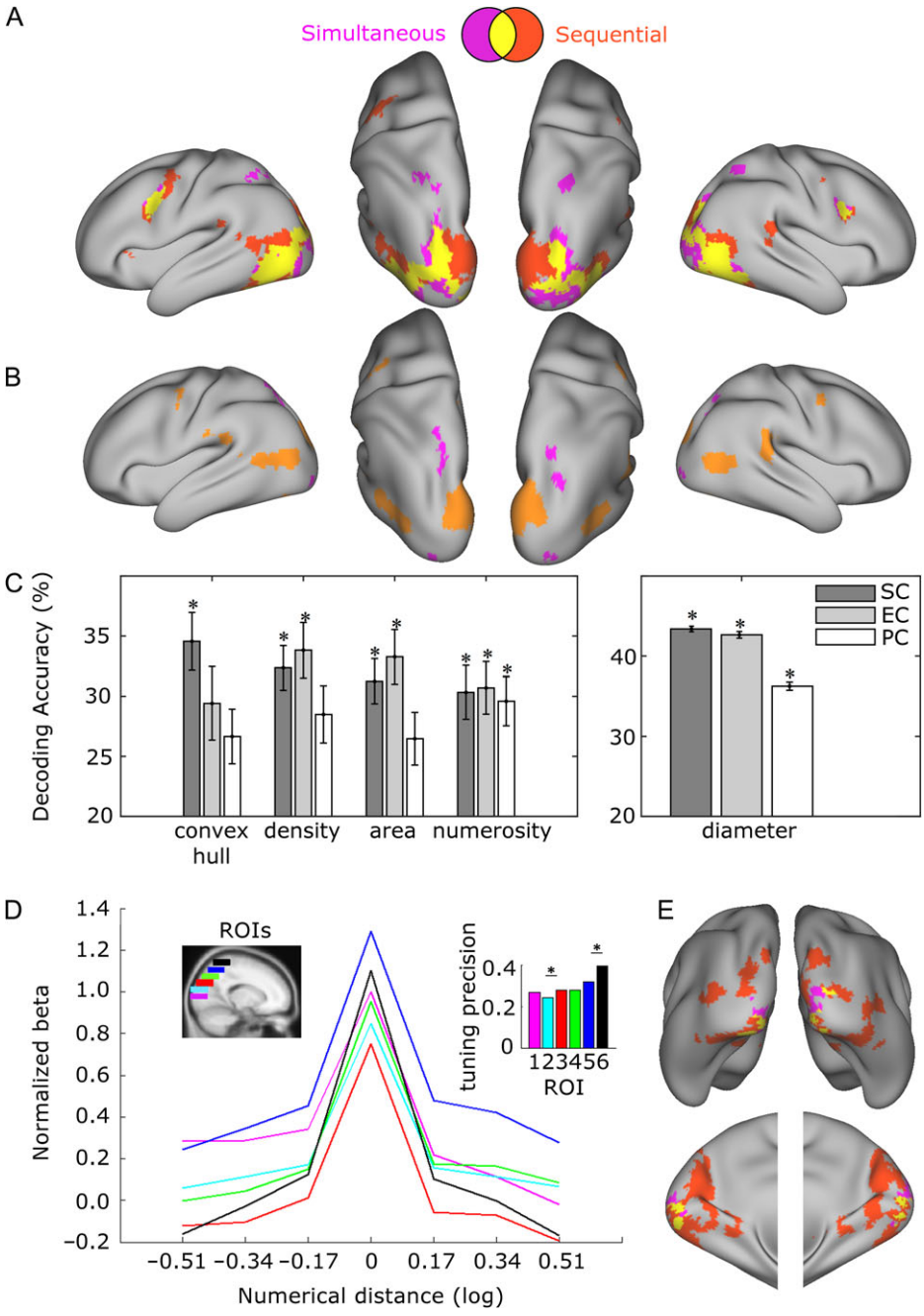

Figure 2. (*A*) Depiction of areas where the BOLD signal increased significantly for simultaneous (purple) and sequential (orange) numerosities in nonresponse trials. Overlapping activations are shown in yellow. (*B*) Depiction of areas where there was a greater BOLD increase for simultaneous compared with sequential (purple) or sequential compared with simultaneous (orange) numerosities. (*C*) Bar graph depicts the average decoding accuracy obtained from the MVPA. The graph on the left depicts the results for convex hull, density, area, and numerosity, each of which had 4 categories. The chance level was determined by permutation analysis. While sensory measures as well as numerosity were decoded significantly from visual ROIs (i.e., SC and EC), only the decoding accuracy for numerosity was significant in PC. The graph on the right depicts the decoding accuracy for diameter, which had 3 categories. An equal number of trials was chosen per diameter category to have a balanced sample for MVPA. The chance level was again determined by permutation analysis. The decoding accuracy for diameter was significant both in visual ROIs and PC. (*D*) Normalized beta-weights for all 6 ROIs (depicted on the brain within the inset) as a function of log distance between numerosities in simultaneous format. The beta values follow a tuning-profile in all the ROIs but the precision of tuning (i.e., slope of the tuning curves) increases as one moves from visual to parietal areas. (*E*) Depiction of areas where the BOLD signal increased more for large numerosities (11 and 16) compared with small numerosities (5 and 7) for simultaneous and sequential numerosities. Only visual cortex exhibited summation coding like activity for both simultaneous and sequential numerosities. All activations FDR corrected at $P = 0.05$ on cluster level, $P = 0.001$ on voxel level, cluster size 15. Left, top, and right views (respectively) of the inflated Human Connectome Project atlas (group average S1200) using Connectome Workbench software (Marcus et al. 2011).

numerosity and separately for simultaneous and sequential trials. For simultaneous numerosities in nonresponse trials, BOLD signal increased significantly in bilateral visual areas, bilateral intraparietal lobule, left-hemispheric superior parietal lobule, and bilateral frontal gyrus (purple color in Fig. 2*A* and Supplementary Table S1). For sequential numerosities in

nonresponse trials, BOLD signal increased significantly in bilateral primary visual areas, right-hemispheric superior temporal sulcus, left insula and precentral gyrus, and right-hemispheric BA 44 (orange color in Fig. 2A and Supplementary Table S1). Only in the frontal cortex (BA 44) and visual cortex (V5 and Area 18) did the BOLD signal increased when we inclusively masked simultaneous and sequential nonresponse trials (yellow areas in Fig. 2A, Supplementary Table S2).

Since previous studies found parietal involvement during sequential numerosity processing, the absence of parietal activation may simply be due to a lack of statistical power. To increase statistical power for sequential numerosities, we included 13 participants (healthy adults; 4 males; mean age = 26.3 ± 6.29 years) from a previous experiment who performed the same task on the same visual sequential numerosity stimuli (for more details see Cavdaroglu et al. 2015). While activity in occipital areas remains stable across the 2 studies, no parietal activation can be observed for sequential numerosities in the absence of active response preparation despite considerably increased statistical power (cf. Supplementary Fig. S1). No such analysis was done for simultaneous numerosities since participants in the previous study (Cavdaroglu et al. 2015) were presented with auditory numerosities instead.

To further probe brain areas that were more activated for sequential or simultaneous numerosities, respectively, we contrasted both modes against each other. Simultaneous numerosities evoked more activity in bilateral parietal cortex, bilateral area V3v, and right middle occipital gyrus (purple in Fig. 2B). Areas that were more active during encoding of sequential numerosities include bilateral occipital cortex (middle occipital gyrus, left area 18 and area 4p), middle cingulate cortex, left insula, and bilateral precentral sulcus (orange in Fig. 2B). Peak coordinates and cluster sizes are reported in Supplementary Table S3.

The BOLD signal for response trials was captured by contrasting the second numerosity in response trials against nonresponse trials separately for simultaneous and sequential numerosities. As the comparison (and response) came right after the presentation of the second numerosity, this contrast included comparison/response related activity. We observed the classic frontoparietal task-positive network (Fox et al. 2005); see Supplementary Fig. S2A). The BOLD signal increased prominently in the parietal cortex (as well as other areas) for response trials both in simultaneous and sequential mode

In order to investigate the areas that show summation coding like activity, we subtracted the BOLD signal for smaller numerosities (5 and 7) from the BOLD signal for larger numerosities (11 and 16). Only visual cortex exhibited summation coding like activity for both simultaneous and sequential numerosities (Fig. 2E).

Multivariate Analysis

For multivariate analysis, we chose ROIs from PPC, extrastriate cortex and striate cortex based on a combination of localizer and task activity (see Materials and Methods for details; see Fig. 1B for ROIs). To test whether the decoding accuracies in parietal cortex were significantly different from chance classification, we ran a permutation analysis with 1000 cycles where the labels of training data were shuffled and randomized. We tested the average MVPA accuracies per participant against the accuracies from the permutation analyses for both modes (simultaneous and sequential). While the decoding accuracy for simultaneous numerosities was significantly higher than chance in the parietal ROI ($t_{(16)} = 2.25$, $P = 0.039$), the decoding accuracy for sequential numerosities did not reach significance ($t_{(16)} = 0.44$, $P = 0.66$; see Fig. 2C for graphical depiction of decoding accuracies and see Supplementary Fig. S2B for confusion matrices).

To test whether the decoding accuracies in visual and parietal cortices were significantly different from chance classification, we run a permutation analysis with 1000 cycles where the labels of training data were shuffled and randomized. We tested the MVPA accuracies per participant against the accuracies from the permutation analyses for all measures (convex hull, density, diameter, total area, and numerosity) and ROIs (striate, extrastriate, and parietal). As one of the stimulus sets for simultaneous numerosities had constant diameter, there was an imbalance in the number of trials for that diameter category in MVPA analysis. While half of the trials had the same diameter, the other half had 4 different diameter values where the diameter decreased with increasing numerosity (i.e., the set where total area was constant, hence diameter was decreasing with increasing numerosity). To overcome this, we grouped the diameters in the set where total area was constant in 2 categories (number 5 and 7 one category as their diameter was closer to each other, 11 and 16 another) and picked equal number of trials from the other set where diameter size was constant. Hence, only for diameter, we had 3 categories instead of four. Decoding accuracies for numerosity and diameter were significant in all ROIs (numerosity: $t_{(16)} = 2.39$, $P = 0.03$; $t_{(16)} = 2.6$, $P = 0.02$; $t_{(16)} = 2.24$, $P = 0.03$; diameter: $t_{(16)} = 5.5$, $P < 0.001$; $t_{(16)} = 4.46$, $P < 0.001$; $t_{(16)} = 3.92$, $P = 0.001$ for striate, extrastriate, and parietal, respectively; see Fig. 2C for graphical depiction of decoding accuracies and Supplementary Fig. S2B for confusion matrices). Decoding accuracies for total area and density were significant only in striate and extrastriate areas (total area: $t_{(16)} = 3.36$, $P = 0.004$; $t_{(16)} = 3.6$, $P = 0.002$; $t_{(16)} = 0.667$, $P = 0.51$; density: $t_{(16)} = 3.88$, $P = 0.001$; $t_{(16)} = 3.79$, $P = 0.002$; $t_{(16)} = 1.45$, $P = 0.167$ for striate, extrastriate, and parietal, respectively). Finally, decoding accuracies for convex hull were significant only in the striate visual cortex ($t_{(16)} = 4.02$, $P < 0.001$; $t_{(16)} = 1.45$, $P = 0.17$; $t_{(16)} = 0.68$, $P = 0.51$ for striate, extrastriate, and parietal, respectively).

To investigate if any of the ROIs had a mode-independent representation, we tested whether the classifier could discriminate presentation modes (simultaneous and sequential). Again, we statistically validated the resulting accuracies against the accuracies obtained from the permutation analysis with 128 cycles—which was the highest possible amount of permutations—where labels of training data were shuffled and randomized. Interestingly, decoding accuracies for presentation mode were significantly above chance in all the ROIs tested (striate: $t_{(16)} = 24.87$, $P < 0.001$; extrastriate: $t_{(16)} = 15.93$, $P < 0.001$; parietal: $t_{(16)} = 5.15$, $P < 0.001$) indicating separate representations for simultaneous and sequential numerosities.

We observed overlapping activity for simultaneous and sequential numerosities in left and right precentral gyrus. The assumption that these regions provide the abstract convergence of sequential and simultaneous numerosities stipulates 1) significant activation and 2) numerosity specificity as tested with MVPA that 3) generalizes across modes (i.e., sequential to simultaneous, simultaneous to sequential). However, for none of these regions, we were able to significantly decode numerosity (precentral gyrus: $P = 0.4338$ for simultaneous, $P = 0.5156$ for sequential). This contradicts conditions 2) and 3).

Tuning Profiles

Similar to the analysis of single neuron numerosity tuning (Nieder 2012; Viswanathan and Nieder 2013), we determined for

each participant which numerosity a given voxel responded to maximally and computed voxel-wise tuning curves along an occipital-to-parietal gradient. First, we observed number-selective voxels at each step within the gradient from occipital to parietal cortex (Fig. 2D). We found a significantly positive regression slope in all ROIs (all $t[13] > 14$ with $Ps < 0.0001$). However, the slopes varied as a function of ROI ($F[5, 65] = 9.093$, $P = 0.0001$). Pairwise comparisons between neighboring ROIs revealed that the slope was significantly larger in ROI 3 compared with ROI 2 ($t[13] = 2.83$, $P = 0.014$). Maximal slope was observed in ROI 6, where it was larger compared with ROI 5 ($t[13] = 2.49$, $P = 0.027$). We found that numerical distance from preferred numerosity ($F[3, 39] = 426.27$, $P < 0.0001$, $\varepsilon = 0.619$) and ROI ($F[4, 65] = 4.84$, $P = 0.006$, $\varepsilon = 0.603$) had a significant impact on brain activity, that significantly interacted with each other ($F[15, 195] = 5.59$, $P = 0.0003$, $\varepsilon = 0.324$). Within each ROI we observed a significant change of activity as numerical distance between preferred and actual numerosity increased (all $Ps < 0.0001$). We found that within all ROIs except ROI 3 the comparison between numerical distances ±0.51 and ±0.34 was not significant while all other comparisons between adjacent distances were significant (Table 1). This implies that in all ROIs except ROI 3 the tuning was most marked for numerosities numerically close to the preferred numerosity and became increasingly blurred for nonpreferred numerosities deviating maximally from the preferred numerosity.

Discussion

Whether the parietal cortex hosts a mode-independent semantic representation of numerosity has long been debated. Here, we probed sequential (i.e., dots presented over time) and simultaneous (i.e., dots presented over space) numerosity perception while recording BOLD response. Going beyond previous studies, we tested numerosities outside the subitizing range and isolated perception from decision and response-related processes. We found an increase in the parietal BOLD signal during the presentation of simultaneous numerosities but not during the presentation of sequential numerosities. Using MVPA we successfully trained a classifier to decode simultaneous numerosity from the BOLD signal in the parietal cortex, providing further confirmation of numerosity selective activity in these areas. No better-than-chance classification was observed for sequential numerosities in the same ROIs. These results imply distinct underlying coding schemes for sequential and simultaneous numerosities. This idea is further supported by significant decoding of the presentation mode (i.e., simultaneous vs. sequential) in the parietal ROIs. We used ROI-based MVPA to further explore how the encoding of simultaneous numerosity and other visual features (i.e., convex hull, total area, density, and diameter) evolves from the primary visual cortex to the parietal cortex. While striate and extrastriate areas gave rise to successful classification of both non-numerical visual features and numerosity, parietal ROIs allowed for decoding of numerosity and dot diameter only, suggesting a higher-level representation in the parietal cortex beyond sensory features. We found voxel-wise numerosity tuning functions for simultaneous numerosities in occipital and parietal ROIs. Numerosity selectivity increased along an occipitoparietal gradient reaching maximal selectivity in parietal areas. We observed overlapping summation coding profiles (higher BOLD signal for large numerosities vs. small numerosities) for sequential and simultaneous numerosities in low-level visual areas only.

Table 1 Pairwise comparisons testing whether numerosity tuning is limited to the difference between preferred numerosity and directly neighboring numerical distances (last column) or whether numerosity preference extends to numerosities with larger numerical distance from preferred numerosity (first and second column).

	Numerical distances					
	±0.51 vs. ±0.34		±0.34 vs. 0.17		0.17 vs. 0	
	$t(13)$	P	$t(13)$	P	$t(13)$	P
ROI 1	−1.6130	0.1307	−3.0407	0.0095	−25.4139	<0.0001
ROI 2	−1.3049	0.2146	−2.1541	0.0506	−31.8596	<0.0001
ROI 3	−1.4337	0.1753	−2.4205	0.0309	−24.1097	<0.0001
ROI 4	−1.0519	0.3120	−2.2468	0.0427	−27.1566	<0.0001
ROI 5	−1.6776	0.1173	−3.2145	0.0068	−17.1302	0.0001
ROI 6	−1.9087	0.0786	−3.1432	0.0078	−17.0212	0.0001

Previous human neuroimaging found an occipital–parietal gradient for numerosities in the subitizing range that associated occipital areas with a location map, occipitoparietal areas with summation coding and superior parietal areas and IPS with number-selective coding (Roggeman et al. 2011). Our results complement these in several ways. First, we found selectivity in PPC for simultaneous numerosities outside subitizing range and independent from response requirements. This suggests that the number-selective coding scheme that is at the top of the model of Dehaene and Changeux (1993) may generalize to larger numerosities. Note that at the time when proposing their model, Dehaene and Changeux (1993) did not consider subitizing to be a different process from estimation (Revkin et al. 2008). Hence, our results provide the first empirical extrapolation of this model to larger numerosities in the light of recent evidence that imply a procedural distinction between small and large numerosities (Revkin et al. 2008; Anobile et al. 2014). Second, unlike striate and extrastriate areas, parietal cortex did not allow for the decoding of non-numerical stimulus features such as density or convex hull, supporting the notion of a high-level abstract number code in IPS. This notion is in line with recent neuroimaging findings that observed number-selective activity patterns in IPS in both adults (Bulthé et al. 2015; Eger et al. 2015; Castaldi et al. 2016) and adolescents (Wilkey et al. 2017). With respect to the question whether (Bulthé et al. 2015; Eger et al. 2015) or not (Castaldi et al. 2016) striate and extrastriate areas allow numerosity decoding, our results suggest that numerosity is represented in the striate and extrastriate areas as well as parietal cortex. At the same time, non-numerical stimulus features such as density and convex hull were decodable only in striate and extrastriate areas but not in parietal cortex. Together, this underlines the idea that numerosity 1) can be conceived of as a primary visual feature that influences activity patterns during early processing steps in the visual system and 2) that numerosity is abtracted away from these physical features of the visual scene as activity travels up towards IPS. The interaction between numerical and non-numerical information along the visual processing hierarchy remains an important question to unravel in the future.

Finally, unlike Roggeman et al. (2011) we did not observe areas in the occipitoparietal transition zone that exhibit a summation coding scheme for simultaneous numerosities. However, contrasting large with small simultaneous numerosities revealed summation coding in low-level visual areas, close to the occipital pole. These voxels partially overlapped with voxels that were more active for large sequential numerosities compared with small sequential numerosities. Sequential

summation coding was also observed in occipital areas along the calcarine sulcus and in superior occipital areas, most likely reflecting longer stimulation during larger numerosities. Hence, it remains an open question to what extent previously observed priority maps in superior PPC (Knops et al. 2014) contribute to a summation mechanism during numerosity extraction, as suggested by prominent computational models (Dehaene and Changeux 1993; Verguts and Fias 2004).

Monkey area VIP has been shown to contain 1) distinct neuronal circuits for the coding of simultaneous and sequential numerosities and 2) overlapping neural circuits for the maintenance of numerosities from either mode (Nieder et al. 2006). Furthermore, a recent fMRI study revealed adaptation for sequential numerosities in human IPS (Wang et al. 2015). Although our results seem to contradict these findings, it should be noted that both studies employed small numerosities (1–4 and 2–6, respectively). Moreover, fMRI adaptation and primate neurophysiology can measure neural activity on subvoxel level whereas GLM and MVPA measure the activity from tens or hundreds of voxels that contain millions of neurons (Logothetis 2008). While primate neurophysiology studies measure spiking activity from single neurons, fMRI BOLD signal correlates better with local field potentials (Goense and Logothetis 2008). These render a one-to-one mapping between MVPA analysis and adaptation studies or neurophysiology difficult. We observed overlapping activity in prefrontal cortex. This may be interpreted as the neural instantiation of a high-level integration of numerosity information across modes and modalities that has been observed on the behavioral level (Arrighi et al. 2014). In order to claim that these regions were actually providing the abstract convergence of sequential and simultaneous numerosities would require 1) significant activation and 2) numerosity specificity as tested with MVPA that 3) generalizes across modes (i.e., sequential to simultaneous and simultaneous to sequential). However, for none of these regions, we were able to significantly decode numerosity. This contradicts conditions 2) and 3). Hence we do not consider these regions to contribute to numerosity coding in our experiment. Alternatively, areas in premotor cortex and inferior frontal gyrus have recently been suggested to be involved in the maintenance of sensory information (frequency) across different modalities (tactile and visual) in working memory (Spitzer et al. 2014). These authors suggest that the role of frontal areas goes well beyond executive control functions but is more closely associated to the sensory content in working memory. Hence, these areas may provide a more abstract convergence zone for numerosity information in working memory. It remains unclear, however, why Wu et al. (2018) observed a parametric modulation of activity in precentral areas, while we did not observe a systematic modulation of activity as a function of numerosity. Our finding makes sense under the assumption that participants did not maintain the raw primary percept in WM but rather retained the abstract numerosity information. The present study was not designed to distinguish between encoding and working memory maintenance, and our data to not allow disentangling these processes. Future studies with a more stringent design are needed to clarify differential roles of prefrontal and parietal areas during encoding and working memory maintenance, and how this is associated with supramodal integration that is observed in behavior.

One may argue that encoding of sequential information is associated with higher working memory demands compared with the processing of simultaneous numerosities and that the long ISI in the current study particularly affects the maintenance of sequential numerosities. The fact that participants were overall performing better in response trials for simultaneous numerosities compared with sequential numerosities may be interpreted in this vein. On the neural level, higher working memory demands should lead to higher activity in working memory related areas. Working memory is usually associated with activity in a frontoparietal network, comprising parietal and prefrontal areas (Xu and Chun 2006; Li et al. 2014; Ma et al. 2014). Our results fit nicely with previous results (Xu and Chun 2006), that showed that even with much shorter ISIs (i.e., 1000–1200 ms), superior parietal activity in a working memory task showed significantly smaller modulation of activity as a function of set size in response to sequential presentation at a center location (comparable to the present study) compared with sequential off-center presentation (cf. Fig. 2 in Xu and Chun 2006). Inferior parietal cortex did not exhibit any modulation of activity as a function of set size with sequential presentation at center. Alternatively, the difference between sequential and simultaneous numerosities may result from higher encoding demands for sequential numerosities. However, even if encoding sequential numerosity information is more demanding compared with simultaneous encoding, this is not associated with higher parietal activity for sequential numerosities. This further undermines the idea that parietal areas play a pivotal role during the encoding of sequential numerosities.

Human neuroimaging suggested overlapping representations of sequential and simultaneous numerosities (Castelli et al. 2006; Piazza et al. 2006; Dormal et al. 2010). However, these neuroimaging studies were difficult to interpret due to confounds between numerical and non-numerical stimulus features in the sequential mode (e.g., Dormal et al. 2010 used constant duration, confounding numerosity with frequency) and the use of an active comparison task that may in and by itself activate parietal cortex that is part of a domain generalized, nonspecific network (Hugdahl et al. 2015). Our results question the assumption that in the absence of decisional and response-related requirements, numerosities from different modes converge on a common, abstract, and mode-independent representation in parietal cortex. Instead, we show that when isolating sequential numerosity perception from response requirement and carefully orthogonalizing temporal and numerical stimulus dimensions, parietal BOLD signal remains under threshold for sequential mode, even when pooling across different studies to increase statistical power.

Previous studies suggest an association between numerosity perception and formal math competencies (Feigenson et al. 2013). Despite the convergence of numerosity information from various modes and modalities that is evident from a number of behavioral studies (Arrighi et al. 2014), recent evidence found only spatially distributed numerosities to be associated with formal math skills (Anobile et al. 2018). Against this background, our results suggest that it may be parietal cortex activity that drives the association between the approximate number system and formal math skills. On a functional level, this parallels the idea that structural features of parietal cortex correlate with formal math skills (Price et al. 2016)

To conclude, while the absence of evidence may not be confounded with evidence for absence, considering the MVPA results, our study casts some doubt on the idea of a mode-independent numerosity representation in IPS. As the neural circuits for simultaneous and sequential numerosity comparison largely overlap in response trials, our results do not contradict with previous studies that used numerical tasks and reported common activation for both modes. It remains unclear, however, whether the role of parietal cortex during previous studies was to encode numerosity or to contribute to

domain-general task components such as decision making and response preparation.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Funding

Grant (KN 959/2-1) from Deutsche Forschungsgemeinschaft/German Research Council (DFG) to André Knops in support of Seda Cavdaroglu.

References

Anobile G, Cicchini GM, Burr DC. 2014. Separate mechanisms for perception of numerosity and density. Psychol Sci. 25: 265–270. https://doi.org/10.1177/0956797613501520.

Anobile G, Cicchini GM, Burr DC. 2016. Number as a primary perceptual attribute: a review. Perception. 45:1–27. https://doi.org/10.1177/0301006615602599.

Anobile G, Turi M, Cicchini G, Burr D. 2012. The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. Vision Res. 74:102–109. https://doi.org/10.1016/j.visres.2012.06.005.

Anobile G, Arrighi R, Castaldi E, Grassi E, Pedonese L, Moscoso PAM, Burr DC. 2018. Spatial nut not temporal numerosity thresholds correlate with formal math skills in children. Dev Psychol. 54:458–473. https://doi.org/10.1037/dev0000448.

Arrighi R, Togoli I, Burr DC. 2014. A generalized sense of number. Proc R Soc B. 281:20141791. https://doi.org/10.1098/rspb.2014.1791.

Brainard DH. 1997. The Psychophysics Toolbox. Spat Vis. 10: 433–436. https://doi.org/10.1163/156856897X00357.

Bulthé J, De Smedt B, Op de Beeck HP. 2014. Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. Neuroimage. 87:311–322. https://doi.org/10.1016/j.neuroimage.2013.10.049.

Bulthé J, De Smedt B, Op de Beeck HP. 2015. Visual number beats abstract numerical magnitude: format-dependent representation of Arabic digits and dot patterns in human parietal cortex. J Cogn Neurosci. 27:1376–1387. https://doi.org/10.1162/jocn.

Burr D, Ross J. 2008. A visual sense of number. Curr Biol. 18: 425–428. https://doi.org/10.1016/j.cub.2008.02.052.

Burr DC, Turi M, Anobile G. 2010. Subitizing but not estimation of numerosity requires attentional resources. J Vis. 10:1–10. https://doi.org/10.1167/10.6.20.Introduction.

Castaldi E, Aagten-Murphy D, Tosetti M, Burr D, Morrone MC. 2016. Effects of adaptation on numerosity decoding in the human brain. Neuroimage. 143:364–377. https://doi.org/10.1016/j.neuroimage.2016.09.020.

Castelli F, Glaser DE, Butterworth B. 2006. Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. Proc Natl Acad Sci U S A. 103:4693–4698. https://doi.org/10.1073/pnas.0600444103.

Cavdaroglu S, Katz C, Knops A. 2015. Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception. Neuroimage. 116:135–148. https://doi.org/10.1016/j.neuroimage.2015.04.019.

Chang C, Lin C. 2011. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol. 2:1–39.

Cohen Kadosh R, Walsh V. 2009. Numerical representation in the parietal lobes: abstract or not abstract? Behav Brain Sci. 32:313–28-73. https://doi.org/10.1017/S0140525X09990938.

Dale A. 1999. Optimal experimental design for event-related fMRI. Hum Brain Mapp. 8:109–114.

Dean HL, Hagan MA, Pesaran B. 2012. Only coherent spiking in posterial parietal cortex coordinates looking and reaching. Neuron. 73:829–841. https://doi.org/10.1016/j.neuron.2011.12.035.

Dehaene S, Changeux J-P. 1993. Development of elementary numerical abilities: a neuronal model. J Cogn Neurosci. 5: 390–407. https://doi.org/10.1162/jocn.1993.5.4.390.

Dehaene S, Izard V, Piazza M. 2005. Control over non-numerical parameters in numerosity experiments.

Dehaene S, Molko N, Cohen L, Wilson AJ. 2004. Arithmetic and the brain. Curr Opin Neurobiol. 14:218–224. https://doi.org/10.1016/j.conb.2004.03.008.

Dormal V, Andres M, Dormal G, Pesenti M. 2010. Mode-dependent and mode-independent representations of numerosity in the right intraparietal sulcus. Neuroimage. 52:1677–1686. https://doi.org/10.1016/j.neuroimage.2010.04.254.

Dormal V, Seron X, Pesenti M. 2006. Numerosity-duration interference: a Stroop experiment. Acta Psychol (Amst). 121: 109–124. https://doi.org/10.1016/j.actpsy.2005.06.003.

Eger E, Michel V, Thirion B, Amadon A, Dehaene S, Kleinschmidt A. 2009. Deciphering cortical number coding from human brain activity patterns. Curr Biol. 19:1608–1615. https://doi.org/10.1016/j.cub.2009.08.047.

Eger E, Pinel P, Dehaene S, Kleinschmidt A. 2015. Spatially invariant coding of numerical information in functionally defined subregions of human parietal cortex. Cereb Cortex. 25:1319–1329. https://doi.org/10.1093/cercor/bht323.

Eger E, Sterzer P, Russ MO, Giraud A-L, Kleinschmidt A. 2003. A supramodal number representation in human intraparietal cortex. Neuron. 37:719–726. https://doi.org/10.1016/S0896-6273(03)00036-9.

Eickhoff SB, Heim S, Zilles K, Amunts K. 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage. 32:570–582. https://doi.org/10.1016/j.neuroimage.2006.04.204.

Eickhoff SB, Paus T, Caspers S, Grosbras M-H, Evans AC, Zilles K, Amunts K. 2007. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage. 36:511–521. https://doi.org/10.1016/j.neuroimage.2007.03.060.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage. 25:1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034.

Feigenson L, Dehaene S, Spelke E. 2004. Core systems of number. Trends Cogn Sci. 8:307–314. https://doi.org/10.1016/j.tics.2004.05.002.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A. 102:9673–9678. https://doi.org/10.1073/pnas.0504136102.

Friston K, Zarahn E, Josephs O. 1999. Stochastic designs in event-related fMRI. Neuroimage. 10:607–619. https://doi.org/10.1006/nimg.1999.0498.

Gebuis T, Cohen Kadosh R, Gevers W. 2016. Sensory-integration system rather than approximate number system underlies numerosity processing: a critical review. Acta Psychol (Amst). 171:1–71. https://doi.org/10.1016/j.actpsy.2016.09.003.

Gebuis T, Reynvoet B. 2012a. The interplay between nonsymbolic number and its continuous visual properties. J Exp Psychol Gen. 141:642–648. https://doi.org/10.1037/a0026218.

Gebuis T, Reynvoet B. 2012b. The role of visual information in numerosity estimation. PLoS One. 7:e37426. https://doi.org/10.1371/journal.pone.0037426.

Goense JBM, Logothetis NK. 2008. Neurophysiology of the BOLD fMRI signal in awake monkeys. Curr Biol. 18:631–640. https://doi.org/10.1016/j.cub.2008.03.054.

Göbel SM, Johansen-Berg H, Behrens T, Rushworth MFS. 2004. Response-selection-related parietal activation during number comparison. J Cogn Neurosci. 16:1536–1551. https://doi.org/10.1162/0898929042568442.

Hanks TD, Kopec CD, Brunton BW, Duan CA, Erlich JC, Brody CD. 2015. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature. 520:220–223. https://doi.org/10.1038/nature14066.

Harvey B, Klein BP, Petridou N, Dumoulin SO. 2013. Topographic representation of numerosity in the human parietal cortex. Science. 341:1123–1126. https://doi.org/10.1126/science.1239052. (80-.).

Hebart MN, Görgen K, Haynes J. 2015. The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. Front Neuroinform. 8:88.

Hugdahl K, Raichle ME, Mitra A, Specht K. 2015. On the existence of a generalized non-specific task-dependent network. Front Hum Neurosci. 9:430. https://doi.org/10.3389/fnhum.2015.00430.

Katz LN, Yates JL, Pillow JW, Huk AC. 2016. Dissociated functional significance of decision-related activity in the primate dorsal stream. Nature. 535:285–288. https://doi.org/10.1038/nature18617.

Kleiner M, Brainard DH, Pelli DG. 2007. Thirtieth European Conference on Visual Perception, in: Perception. p. 14.

Knops A, Piazza M, Sengupta R, Eger E, Melcher D. 2014. A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. J Neurosci. 34:9857–9866. https://doi.org/10.1523/JNEUROSCI.2758-13.2014.

Koenigs M, Barbey AK, Postle BR, Grafman J. 2009. Superior parietal cortex is critical for the manipulation of information in working memory. J Neurosci. 29:14980–14986. https://doi.org/10.1523/JNEUROSCI.3706-09.2009.

Kravitz DJ, Saleem KS, Baker CI, Mishkin M. 2011. A new neural framework for visuospatial processing. Nat Rev Neurosci. 12:217–230.

Leibovich T, Katzin N, Harel M, Henik A. 2016. From "sense of number" to "sense of magnitude"—the role of continuous magnitudes in numerical cognition. Behav Brain Sci. 40:e164.

Li D, Christ SE, Cowan N. 2014. Domain-general and domain-specific functional networks in working memory. Neuroimage. 102:646–656. https://doi.org/10.1016/j.neuroimage.2014.08.028.

Logothetis NK. 2008. What we can do and what we cannot do with fMRI. Nature. 453:869–878. https://doi.org/nature06976[pii]/r10.1038/nature06976.

Ma WJ, Husain M, Bays PM. 2014. Changing concepts of working memory. Nat Neurosci. 17:347–356. https://doi.org/10.1038/nn.3655.

Maldjian J, Laurienti PJ, Burdette JH. 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. Neuroimage. 21:450–455. https://doi.org/10.1016/j.neuroimage.2003.09.032.

Maldjian J, Laurienti PJ, Kraft R. a., Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 19:1233–1239. https://doi.org/10.1016/S1053-8119(03)00169-1.

Malhotra P, Coulthard EJ, Husain M. 2009. Role of right posterior parietal cortex in maintaining attention to spatial locations over time. Brain. 132:645–660. https://doi.org/10.1093/brain/awn350.

Marcus DS, Harwell J, Olsen T, Hodge M, Glasser MF, Prior F, Jenkinson M, Laumann T, Curtiss SW, Van Essen DC. 2011. Informatics and data mining tools and strategies for the Human Connectome Project. Front Neuroinform. 5:1–12. https://doi.org/10.3389/fninf.2011.00004.

Nieder A. 2012. Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. Proc Natl Acad Sci USA. 109:11860–11865. https://doi.org/10.1073/pnas.1204580109.

Nieder A, Diester I, Tudusciuc O. 2006. Temporal and spatial enumeration processes in the primate parietal cortex. Science. 313:1431–1435. https://doi.org/10.1126/science.1130308. (80-.).

Nieder A, Miller EK. 2004. A parieto-frontal network for visual numerical information in the monkey. Proc Natl Acad Sci U S A. 101:7457–7462. https://doi.org/10.1073/pnas.0402239101.

Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis. 10:437–442.

Piazza M, Fumarola A, Chinello A, Melcher D. 2011. Subitizing reflects visuo-spatial object individuation capacity. Cognition. 121:147–153. https://doi.org/10.1016/j.cognition.2011.05.007.

Piazza M, Mechelli A, Butterworth B, Price CJ. 2002. Are subitizing and counting implemented as separate or functionally overlapping processes? Neuroimage. 15:435–446. https://doi.org/10.1006/nimg.2001.0980.

Piazza M, Mechelli A, Price CJ, Butterworth B. 2006. Exact and approximate judgements of visual and auditory numerosity: an fMRI study. Brain Res. 1106:177–188. https://doi.org/10.1016/j.brainres.2006.05.104.

Piazza M, Pinel P, Le Bihan D, Dehaene S. 2007. A magnitude code common to numerosities and number symbols in human intraparietal cortex. Neuron. 53:293–305. https://doi.org/10.1016/j.neuron.2006.11.022.

Pinel P, Thirion B, Meriaux S, Jobert A, Serres J, Le Bihan D, Poline J-B, Dehaene S. 2007. Fast reproducible identification and large-scale databasing of individual functional cognitive networks. BMC Neurosci. 8:91. https://doi.org/10.1186/1471-2202-8-91.

Price GR, Wilkey ED, Yeo DJ, Cutting LE. 2016. The relation between 1st grade grey matter volume and 2nd grade math competence. Neuroimage. 70:773–779. https://doi.org/10.1097/OGX.0000000000000256.Prenatal.

Revkin SK, Piazza M, Cohen L, Vi P. 2008. Does subitizing reflect numerical estimation? Psychol Sci. 19:607–614.

Roggeman C, Santens S, Fias W, Verguts T. 2011. Stages of nonsymbolic number processing in occipitoparietal cortex disentangled by fMRI adaptation. J Neurosci. 31:7168–7173. https://doi.org/10.1523/JNEUROSCI.4503-10.2011.

Serences JT, Saproo S, Scolari M, Ho T, Muftuler LT. 2009. Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. Neuroimage. 44:223–231. https://doi.org/10.1016/j.neuroimage.2008.07.043.

Shadlen MN, Newsome WT. 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol. 86:1916–1936. https://doi.org/10.3410/f.1001494.23207.

Shafritz KM, Gore JC, Marois R. 2002. The role of the parietal cortex in visual feature binding. Proc Natl Acad Sci U S A. 99: 10917–10922. https://doi.org/10.1073/pnas.152694799.

Shomstein S. 2012. Cognitive functions of the posterior parietal cortex: top-down and bottom-up attentional control. Front Integr Neurosci. 6:38. https://doi.org/10.3389/fnint.2012.00038.

Shuman M, Kanwisher N. 2004. Numerical magnitude in the human parietal lobe: tests of representational generality and domain specificity. Neuron. 44:557–569. https://doi.org/10.1016/j.neuron.2004.10.008.

Spitzer B, Fleck S, Blankenburg F. 2014. Parametric alpha- and beta-band signatures of supramodal numerosity information in human working memory. J Neurosci. 34:4293–4302. https://doi.org/10.1523/JNEUROSCI.4580-13.2014.

Stoianov I, Zorzi M. 2012. Emergence of a "visual number sense" in hierarchical generative models. Nat Neurosci. 15:194–196. https://doi.org/10.1038/nn.2996.

Takemura H, Ashida H, Amano K, Kitaoka A, Murakami I. 2012. Neural correlates of induced motion perception in the human brain. J Neurosci. 32:14344–14354. https://doi.org/10.1523/JNEUROSCI.0570-12.2012.

Tokita M, Ishiguchi A. 2011. Temporal information affects the performance of numerosity discrimination: behavioral evidence for a shared system for numerosity and temporal processing. Psychon Bull Rev. 18:550–556. https://doi.org/10.3758/s13423-011-0072-2.

Tudusciuc O, Nieder A. 2007. Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. Proc Natl Acad Sci U S A. 104:14513–14518. https://doi.org/10.1073/pnas.0705495104.

Verguts T, Fias W. 2004. Representation of number in animals and humans: a neural model. J Cogn Neurosci. 16:1493–1504. https://doi.org/10.1162/0898929042568497.

Viswanathan P, Nieder A. 2013. Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. Proc Natl Acad Sci USA. 110:11187–11192. https://doi.org/10.1073/pnas.1308141110.

Wang L, Uhrig L, Jarraya B, Dehaene S. 2015. Representation of numerical and sequential patterns in macaque and human brains. Curr Biol. 25:1966–1974. https://doi.org/10.1016/j.cub.2015.06.035.

Wilkey ED, Barone JC, Mazzocco MMM, Vogel SE, Price GR. 2017. The effect of visual parameters on neural activation during nonsymbolic number comparison and its relation to math competency. Neuroimage. 159:430–442. https://doi.org/10.1016/j.neuroimage.2017.08.023.

Wu YH, Uluç I, Schmidt TT, Tertel K, Kirilina E, Blankenburg F. 2018. Overlapping frontoparietal networks for tactile and visual parametric working memory representations. Neuroimage. 166:325–334. https://doi.org/10.1016/j.neuroimage.2017.10.059.

Xu Y, Chun MM. 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. Nature. 440:91–95. https://doi.org/10.1038/nature04262.

frontiers in Psychology

ORIGINAL RESEARCH
published: 17 July 2018
doi: 10.3389/fpsyg.2018.01062

The Developmental Trajectory of the Operational Momentum Effect

Pedro Pinheiro-Chagas[1,2†], Daniele Didino[3*†], Vitor G. Haase[4,5,6,7], Guilherme Wood[8,9] and André Knops[3,10,11]*

[1] Cognitive Neuroimaging Unit, CEA DRF/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin Center, Orsay, France, [2] Laboratory of Behavioral and Cognitive Neuroscience, Stanford Human Intracranial Cognitive Electrophysiology Program, Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, United States, [3] Department of Psychology, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany, [4] Developmental Neuropsychology Laboratory (LND), Department of Psychology, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [5] Programa de Pós-Graduação em Neurociências, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [6] Department of Psychology, Graduate Program in Psychology, Cognition and Behavior – Graduate Program in Neuroscience, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [7] Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino, Universidade Federal de São Carlos, São Carlos, Brazil, [8] Department of Psychology, University of Graz, Graz, Austria, [9] BioTechMed-Graz, University of Graz, Graz, Austria, [10] CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, Paris, France, [11] University Paris Descartes, Sorbonne Paris Cité, Paris, France

OPEN ACCESS

Edited by:
Maciej Haman,
University of Warsaw, Poland

Reviewed by:
Christine Schiltz,
University of Luxembourg, Luxembourg
Mojtaba Soltanlou,
Eberhard Karls Universität Tübingen, Germany

****Correspondence:***
Daniele Didino
didiwoda@hu-berlin.de
Pedro Pinheiro-Chagas
ppinheirochagas@gmail.com

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *22 December 2017*
Accepted: *06 June 2018*
Published: *17 July 2018*

Citation:
Pinheiro-Chagas P, Didino D, Haase VG, Wood G and Knops A (2018) The Developmental Trajectory of the Operational Momentum Effect. Front. Psychol. 9:1062. doi: 10.3389/fpsyg.2018.01062

Mental calculation is thought to be tightly related to visuospatial abilities. One of the strongest evidence for this link is the widely replicated operational momentum (OM) effect: the tendency to overestimate the result of additions and to underestimate the result of subtractions. Although the OM effect has been found in both infants and adults, no study has directly investigated its developmental trajectory until now. However, to fully understand the cognitive mechanisms lying at the core of the OM effect it is important to investigate its developmental dynamics. In the present study, we investigated the development of the OM effect in a group of 162 children from 8 to 12 years old. Participants had to select among five response alternatives the correct result of approximate addition and subtraction problems. Response alternatives were simultaneously presented on the screen at different locations. While no effect was observed for the youngest age group, children aged 9 and older showed a clear OM effect. Interestingly, the OM effect monotonically increased with age. The increase of the OM effect was accompanied by an increase in overall accuracy. That is, while younger children made more and non-systematic errors, older children made less but systematic errors. This monotonous increase of the OM effect with age is not predicted by the compression account (i.e., linear calculation performed on a compressed code). The attentional shift account, however, provides a possible explanation of these results based on the functional relationship between visuospatial attention and mental calculation and on the influence of formal schooling. We propose that the acquisition of arithmetical skills could reinforce the systematic reliance on the spatial mental number line and attentional mechanisms that control the displacement along this metric. Our results provide a step in the understanding of the mechanisms underlying approximate calculation and an important empirical constraint for current accounts on the origin of the OM effect.

Keywords: operational momentum, approximate addition, approximate subtraction, children, development, attentional shift account, compression account, heuristic account

INTRODUCTION

Adults and children (Barth et al., 2006), and even infants (Wynn, 1992), are able to perform approximate mental calculation, which consists in the capacity to add or subtract numbers expressed in non-symbolic notations (e.g., dots). This skill requires to estimate the numerosity (i.e., cardinality) of two sets of elements and to encode it on an internal representation on which cognitive processes operate to generate the approximate outcome of the calculation. Growing evidence (McCrink et al., 2007; Pinhas and Fischer, 2008; Knops et al., 2009b; McCrink and Wynn, 2009; Lindemann and Tira, 2011; Chen and Verguts, 2012; Knops et al., 2013, 2014; Klein et al., 2014; Marghetis et al., 2014; Pinheiro-Chagas et al., 2017) shows that approximate addition and subtraction are subjected to an Operational Momentum (hereafter, OM) effect: results of addition are overestimated and results of subtraction are underestimated. Although an OM effect has been found in infants (McCrink and Wynn, 2009) and an inverse OM effect emerged in 6/7 years old children (Knops et al., 2013), no studies investigated the developmental trajectory of this effect. Therefore, it is still unclear how the OM effect evolves during the acquisition of formal mathematical knowledge. The relevance of the OM effect lies in the knowledge it provides regarding the cognitive mechanisms involved in the representation and the manipulation of non-symbolic numerical magnitudes. In this study, we aimed to measure how the OM effect evolves in children between 8 and 12 years of age. Moreover, the developmental trajectory of the OM effect can also provide evidence in favor of or against the current accounts proposed to explain this effect.

A prerequisite to perform approximate mental calculation is the capacity to estimate and manipulate numerical quantities, which is a phylogenetically ancient cognitive tool that humans share with other animals (Flombaum et al., 2005; Cantlon and Brannon, 2007; Piazza, 2010) and that arises early in life (Xu and Spelke, 2000; Izard et al., 2009). A widely accepted view (Dehaene, 1997) assumes that the mental representation of numerical magnitudes takes the form of an analog mental number line (hereafter, MNL). In the last decades, evidence has been collected to support the idea that on the MNL numerosities are spatially oriented in ascending order from left to right (Dehaene et al., 1993; Fias and Fischer, 2005; Hubbard et al., 2005; Rugani and Sartori, 2016; de Hevia et al., 2017). The SNARC effect (spatial numerical association of response codes; Dehaene et al., 1993) is often interpreted as evidence for the functional association between numbers and space: in a parity judgment tasks, where participant have to decide whether a displayed number is odd or even, left-hand responses are faster for relatively small number and right-hand responses for relatively large numbers (Dehaene et al., 1993; Fias and Fischer, 2005; Hubbard et al., 2005). Since the magnitude of the number is not relevant for the task, this spatial bias is assumed to reflect the automatic activation of the spatial mapping of magnitudes on the MNL (but for an alternative account see Santens and Gevers, 2008). The functional association between visuospatial processing and numerical magnitudes is additionally suggested by the mounting evidence showing that a shift of spatial attention can be induced by number processing (Sallilas et al., 2008; Ranzini et al., 2015, 2016; for a review see Fischer and Knops, 2014). It is worth noting that a functional association also emerges between shifts of spatial attention and mental arithmetic (Masson and Pesenti, 2014, 2016; Mathieu et al., 2016, 2017; Masson et al., 2017a,b). Moreover, converging evidence from behavioral (Izard and Dehaene, 2008), computational (Dehaene and Changeux, 1993), and neurophysiological studies (Nieder and Miller, 2003) suggests that the MNL is logarithmically compressed, which means that the representational overlap between adjacent quantities increases proportionally to their size, in accordance with the Weber–Fechner law (see Piazza et al., 2010).

Approximate calculation also follows the Weber–Fechner law (Barth et al., 2006; Dehaene, 2007), but it also shows an additional response bias, that is the OM effect. Three mutually not exclusive mechanisms have been proposed to explain the OM effect: attentional shift account, heuristic account, and compression account. However, none of them aimed to describe how this effect changes over development. Evidence shows that the neural network that supports mental calculation undergoes substantial functional changes during development and reaches an adult-like configuration only during adolescence (Rosenberg-Lee et al., 2011; Soltanlou et al., 2017, 2018; Arsalidou et al., 2018; Peters and De Smedt, 2018). Therefore, in order to fully understand the cognitive mechanisms lying at the core of the OM effect it is important to measure its developmental dynamics and to evaluate whether the current accounts are able to explain these age-related changes. In what follows, we introduce these accounts of the OM effect and discuss the developmental trajectories predicted by each of them.

It has been proposed that mental calculation is grounded in neural circuits that originally evolved for processing visuospatial information (Anderson, 2007; Dehaene and Cohen, 2007; Knops et al., 2009a). Moreover, various evidence supports the existence of a functional relationship between visuospatial attention (i.e., shift of spatial attention) and mental calculation (Masson and Pesenti, 2014, 2016; Mathieu et al., 2016, 2017; Masson et al., 2017a,b). In line with these studies, the *attentional shift account* proposes that the OM effect is the result of this functional relationship (McCrink et al., 2007; Knops et al., 2009b; Pinheiro-Chagas et al., 2017). The central assumption of the attentional shift account hypothesizes that non-symbolic addition and subtraction are implemented by shifting spatial attention on a spatially oriented MNL. During approximate calculation, the first operand is mapped on the MNL, then the attentional focus shifts from the current position (i.e., the point corresponding to the magnitude of the first operand) to a new position (i.e., the point corresponding to the magnitude of the result) by a distance corresponding to the magnitude of the second operand. The OM effect is produced by a bias in the attentional shift, that is the attentional focus moves too far along the MNL in the direction of the operation, generating an overestimation and an underestimation of the result of addition and subtraction, respectively. Strong evidence for the hypothesis that visuospatial attention is co-opted during mental calculation is provided by the overlap in the posterior superior parietal lobule (PSPL)

of the neural activity associated with left/right saccades (i.e., visuospatial orientation) and mental calculation (Knops et al., 2009a).

McCrink and Wynn (2009) proposed the *heuristic account* to explain the finding that the OM effect also affects performance in 9 months old infants. This account assumes that infants adopted a simple heuristic to solve the problems: "if adding, accept larger outcomes," "if subtracting, accept smaller outcomes." For addition, this heuristic approach might encourage infants to perceive larger outcomes as more plausible compared smaller ones, and vice versa for subtraction. Recently, McCrink and Hubbard (2017) interpreted the finding that the OM effect increased in adults when available attentional resources were limited by dividing attention between two concurrent tasks as further evidence for the heuristics account. However, the heuristic account and the attentional shift account are deeply intertwined and can be considered as a single mechanism (i.e., heuristics-via-spatial-shifts account), that is the heuristic decision results from the visuospatial system (McCrink and Hubbard, 2017). Therefore, we will only focus on the attentional shift account, assuming that the two accounts provide equivalent predictions.

The attentional shift account has been developed to explain the OM effect in adults. Therefore, no predictions or hypotheses were proposed regarding how the attentional shifts on the MNL that accompany addition and subtraction emerge and whether they undergo substantial changes during development. Here, we propose that formal schooling (i.e., acquiring arithmetical skills) could reinforce (or even contribute to develop) the idea that addition is related with shifts toward larger numbers and subtraction toward smaller numbers. Namely, although mental calculation might be implemented as an attentional shift on the MNL before formal schooling, repeated exposition to spatial-numerical associations (e.g., the number line) might consolidate a systematic movement direction during the acquisition of arithmetical skills. Moreover, the systematic association between operations and results (i.e., when adding, the result is always larger than both operands; when subtracting, the result is always smaller than the first operand), that children are exposed to, could boost the attentional shift on the MNL. The influence of the attentional shift in the estimation of the result might increase with age and in turn a larger and more systematic bias would emerge. Therefore, one may predict an increasing OM effect during childhood. Moreover, it is worth noting that the co-opting of visuospatial attention during mental calculation seems to increase with age. In fact, significant functional changes associated with the neural activity elicited by symbolic arithmetic problem-solving have been found between 2nd and 3rd graders, that is 7–9 years old children (Rosenberg-Lee et al., 2011). During the processing of symbolic arithmetic problems, 3rd grade children showed greater activity in brain regions related to visuospatial attentional processes (posterior parietal cortex: intraparietal sulcus, superior parietal lobule, and angular gyrus) and high-order visual processing (ventral visual areas: lingual gyrus, right lateral occipital cortex, and right parahippocampal gyrus), compared to 2nd grade children.

The *compression account* has been proposed by McCrink et al. (2007) and deploys the logarithmic compression of the MNL to explain the OM effect. This compressed metric would generate a systematic operational bias in the direction of the operation due to the implementation of a linear arithmetic operation (i.e., addition or subtraction) on a logarithmically scaled mental representation. This mechanism acts in three steps. First, the operands are encoded as logarithmically compressed magnitudes on the MNL. Second, the logarithmic transformation is undone, which means that the operands are uncompressed to a linear scale. Third, the two uncompressed operands are added or subtracted. The OM effect results from the inaccuracy of the uncompression process. If the uncompression is ineffective the arithmetic operation is performed on logarithmic values and thus the generated outcome corresponds to an extreme overestimation or underestimation for addition and subtraction, respectively. If the uncompression is highly accurate the operation is performed on the linear scale, in which case the generated outcome corresponds (approximately) to the arithmetically correct result. A more plausible scenario is to assume that the actual degree of uncompression lies between these two extreme possibilities. An example can help describe this idea. If uncompression fails, adding two operands (e.g., 26 and 14) corresponds to adding their logarithmically compressed internal representation, that is $\log(26) \approx 3.26$ and $\log(14) \approx 2.64$, respectively. Since adding the logarithm of two numbers is equivalent to multiplying their linear values, the system generates an extreme overestimation of the correct result: $\log(26) + \log(14) \approx 5.9$, which in linear scale corresponds to $e^{5.9} \approx 26 \times 14 \approx 364$. However, the actual approximate addition performed by the system is much more accurate (see for example McCrink et al., 2007), and thus the uncompression is to some extent carried out and the generated outcome is much closer to the correct result. The same reasoning is valid to explain the mechanisms underpinning the underestimation of subtraction outcomes.

What developmental trajectory of the OM effect is expected according to the compression account? This account focuses on the logarithmic compression of the MNL. A large body of evidence suggests that the representational metric of the MNL shifts from a logarithmic to a linear scale during childhood (Siegler and Opfer, 2003; Siegler and Booth, 2004; Booth and Siegler, 2006, 2008; Laski and Siegler, 2007; Opfer and Siegler, 2007 but for a different interpretation see Barth and Paladino, 2011). The logarithmic-to-linear shift of the MNL implies that the compression of this magnitude representation decreases with age and probably with accumulation of experience in formal mathematics teaching. Therefore, the uncompression of the operands, performed before the approximate mental calculation, starts from a highly logarithmic scale in young children and from a more linear scale in adults. The degree of uncompression required to generate an accurate outcome is thus greater in young children and this in turn could lead to a stronger OM effect. The compression account therefore predicts that the size of the OM effect is higher in young children and decreases with age to reach an adult-like pattern in older children. It is worth noting that, as discussed below, the inverse OM effect (i.e., overestimation

of subtraction problems) found in 6/7 years old children (Knops et al., 2013) already provides evidence against this account.

MATERIALS AND METHODS

The sample and the tasks analyzed in the present paper were administered to children as part of a larger study conducted in Brazil (for a more precise description of this larger study see Pinheiro-Chagas et al., 2014).

Participants

One hundred seventy-two children from first to sixth grade were recruited from private and public schools in Brazil. Ten children were not able to perform non-symbolic numerical tasks, as shown by the fact that they failed to perform a non-symbolic number comparison task (this task is not reported here, for a more detailed description of this task see Pinheiro-Chagas et al., 2014). In that non-symbolic number comparison task, children had an accuracy less than 55% and a poor fit ($R^2 < 0.2$) in the estimation of the Weber fraction, and thus were excluded from the study. These ten children were also not included in the present analyses. The final sample consisted of 162 children (66 boys, 96 girls) between 8 and 12 years of age (mean = 9.7 years, $SD = 1.1$; 8 years old: 24 children, 9 years old: 54, 10 years old: 50, 11 years old: 20, 12 years old: 14). Informed written consent was obtained from the parents and oral consent from the children. This study was approved by the ethics review board of the Federal University of Minas Gerais, Brazil (COEP–UFMG).

All children performed above the 25th percentile in the spelling (mean = 110.08, $SD = 8.13$, range = [85, 126]) and arithmetic (mean = 108.92, $SD = 11.41$, range = [86, 134]) subtests of the TDE (Teste de Desempenho Escolar; Stein, 1994) and had a normal intelligence (mean = 110.61, $SD = 10.55$, range = [86, 134]), as measured by Raven's Colored Progressive Matrices (Angelini et al., 1999).

Tasks

Non-symbolic Estimation Task

In this task children were asked to estimate and report verbally the numerosity of a set of dots visually presented on a computer screen. Dots were displayed in black within a white circle, which was presented against a black background. The following numerosities were presented: 10, 16, 24, 32, 48, 56, or 64 dots. Each numerosity was presented five times (in a different configuration), resulting in a total of 35 trials. The same numerosity never appeared in consecutive trials. Each trial started with a fixation point (i.e., a white cross at the center of the screen) presented for 500 ms, followed by the onset of the set of dots which remained on the screen until spacebar was pressed or for up to 1000 ms. During the presentation of the dots, as soon as the child responded, the examiner, who was seated next to the child, pressed the spacebar on the keyboard and typed the child's answer. The next trial started after an intertrial interval of 700 ms, which consisted of a black screen. Dots were displayed on the screen for up to 1000 ms only to prevent counting. To prevent the use of non-numerical features, total dot area was held constant across the trials and thus it could not be used as a clue to estimate the different numerosities. The average dot-size of the dots was selected so that the total area remained constant, but the dot-size of each dot could vary with a normal distribution with the mean selected to provide constant area across the trials. Therefore, while the average dot-size covaried negatively with numerosity, the dot-size of the single dots could not be used as a cue to evaluate the numerosity of the set. To avoid memorization effects due to the repetition of a specific numerosity, on each trial, the stimuli were randomly chosen from a set of 10 precomputed images with the given numerosity. To exclude extreme responses, the normalized mean estimated value was calculated for each child and each of the seven presented numerosities, then responses ±3 SD from the mean estimated value were considered outliers and excluded from the analysis (3.5% of the trials). Children's number acuity was measured in term of individual mean coefficient of variation (i.e., separately for each numerosity, the ratio of standard deviation and mean chosen value).

Non-symbolic Approximate Calculation Task

This task has been adapted from Knops et al. (2013) study. Children were asked to solve approximate addition and subtraction problems with operands and proposed results presented in a non-symbolic notation (i.e., sets of dots). Problems are reported in **Table 1**. Eight addition and eight subtraction problems were generated. Both arithmetic operations had the same range of possible outcomes: 10, 16, 26, 40. To prevent the subjects from memorizing the problems, the operands were randomly "jittered" by adding a random value r, with $r \in J$ and $J = [-1, 0, 1]$. For each correct outcome, seven response alternatives were generated as round ($c \times 2.5^{i/3}$), where c is the correct result and $i = [-3, -2, -1, 0, 1, 2, 3]$. To avoid a strategy of always selecting the response alternative falling in the middle of the proposed range, only five of the seven generated alternatives were presented in a trial (see **Table 1**). In one half of the trials, the presented responses were the upper five (henceforth, high range), and thus the correct outcome was the second smallest numerosity. In the other half, the presented responses were the lower five (henceforth, low range), and thus the correct outcome was the fourth smallest numerosity. Each trial was repeated twice and thus the total number of trials was 64: 2 operations (addition and subtraction) × 8 problems × 2 ranges (high and low) × 2 repetitions. To prevent the use of non-numerical features, total dot area and dot-size were manipulated as in the non-symbolic estimation task. To avoid memorization effects due to the repetition of a specific numerosity, on each trial, the stimuli were randomly chosen from a set of 10 precomputed images with the given numerosity. Trials without response and trials where the selected response was ±3 SD from the normalized mean chosen values (calculated combining addition and subtraction) were considered outliers and excluded from the analysis (3.1% of the trials). To analyze the OM effect, for each child and for each operation (addition vs. subtraction), mean chosen value, standard deviation, and coefficient of variation (i.e., the ratio of standard deviation and mean chosen value) were calculated for each of the four correct outcomes.

TABLE 1 | Operands, correct outcome (C) and deviant (D) outcomes presented in the non-symbolic arithmetic problems.

Operands		Correct results and deviant proposed outcomes						
		1/2.5	1/1.8	1/1.4	1	1.4	1.8	2.5
Addition								
5	5	4	5	7	10	14	18	25
6	4	4	5	7	10	14	18	25
8	8	6	9	12	16	22	29	40
10	6	6	9	12	16	22	29	40
13	13	10	14	19	26	35	48	65
18	8	10	14	19	26	35	48	65
20	20	16	22	29	40	54	74	100
26	14	16	22	29	40	54	74	100
Subtraction								
16	6	4	5	7	10	14	18	25
20	10	4	5	7	10	14	18	25
24	8	6	9	12	16	22	29	40
32	16	6	9	12	16	22	29	40
40	14	10	14	19	26	35	48	65
52	26	10	14	19	26	35	48	65
62	22	16	22	29	40	54	74	100
80	40	16	22	29	40	54	74	100
Range								
	Low	D	D	D	C	D		
	High			D	C	D	D	D

The last two rows report the set of outcomes presented in the two ranges.

To provide a child-friendly paradigm, problems were embedded in a story of a monkey having a box of balls (**Figure 1**). Each trial started with the drawing of the monkey's face presented for 500 ms. After the offset of the monkey's face, an empty brown box (against a black background) appeared at the bottom of the screen and a first set of red dots moved into the box. The first set of dots appeared at the top of the screen and moved toward the box until the dots disappeared inside it. For addition problems, a second set of red dots appeared at the top of the screen and disappeared inside the box in the same way. For subtraction problems, a set of red dots moved out of the box and disappeared at the top of the screen. Both for the first and the second sets, the duration of the dots movement (from the appearance to the disappearance) was 1000 ms. After the second set of dots disappeared, the box was replaced by the top-view of five boxes that contained five different sets of dots (i.e., five responses alternatives). Two boxes appeared on the left of the screen, two on the right, and one on the top. Children were asked to click with the left-key of the mouse on the box containing the set of dots which numerosity was the closest to the correct outcome of the operation. The beginning of the response active period was indicated by the appearance of the mouse pointer on top of a green star in the center of the screen. A training period consisting of two trials preceded the testing phase. In the training period, there was no time limit for the response and feedback was provided by a frame around the chosen box. The appearance of a green frame indicated a correct response, whereas a red frame indicated an incorrect response. If the response was incorrect, the child was asked to choose another box, and this procedure was repeated until the correct box was chosen. Before testing phase, the children were asked if they had understood the task, and if not, the training was repeated until they confirmed that they understood the task. In the testing phase, children had a maximum of 10,000 ms to select the box and the chosen box was indicated by a neutral blue frame (i.e., no feedback provided). Addition and subtraction problems were presented in different blocks counterbalanced across participants.

Data Analysis

All analyses were performed using R-project software (R Core Team, 2015) and RStudio software (RStudio Team, 2015). In the following analyses, ANOVAs were Greenhouse-Geisser corrected (Greenhouse and Geisser, 1959) when the assumption of sphericity was violated; uncorrected degrees of freedom and epsilon values (εGG) are reported. In the *post hoc* analyses all *p*-values have been corrected with Holm's method (Holm, 1979). For the OM effect, effect sizes are reported following the recommendation of Lakens (2013). Additional analyses of children's performance (absolute error) and of the operational bias (ratio) are reported in the Appendix A.

RESULTS

The results of all the ANOVAs performed on the tasks are reported in the Appendix B (Supplementary Table S2).

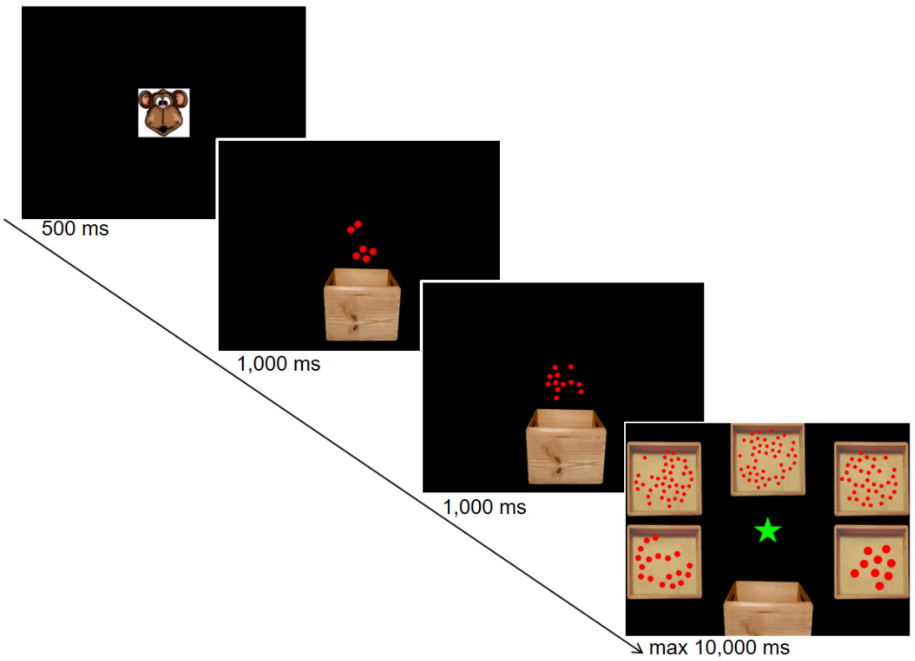

FIGURE 1 | Trial sequence of the non-symbolic approximate calculation task. The example shows the screenshots from a non-symbolic addition trial. During the response period, the five response alternatives were presented in a circle-like shape around the center of the screen (i.e., green star) with two boxes on the left of the screen, two on the right, and one on the top.

Non-symbolic Estimation Task

The first analysis aims to evaluate the performance of children in the non-symbolic number estimation task. Mean chosen numerosity and CV were analyzed with a repeated measure ANOVA with displayed numerosity (i.e., 10, 16, 24, 32, 48, 56, and 64 dots) as within-subject factor and age (i.e., 8 to 12 years old) as between-subject factor. Mean chosen numerosities significantly increased with displayed numerosity [$F(6,942) = 313.45$, $p < 0.001$, $\varepsilon GG = 0.27$, generalized $\eta^2 = 0.47$]. However, as shown in **Figure 2**, and in line with adults' behavior (Knops et al., 2014), children underestimated the larger displayed numerosities. To verify whether this pattern was statistically significant a repeated measure correlation (Bakdash and Marusich, 2017) was performed between numerical difference (chosen numerosity minus displayed numerosity) and displayed numerosity. There was a strong negative correlation between numerical difference and displayed numerosity [$r_{rm}(971) = -0.57$, 95% CI = [−0.61, −0.53], $p < 0.001$], that is the discrepancy between displayed and chosen values increased with numerosity (**Figure 2**). In the ANOVA, neither the main effect of age nor the interaction was significant.

On the basis of the assumption that mental numerosity representation is subjected to the Weber–Fechner law, the CV should not covary with displayed numerosity (i.e., the CV should be constant across numerosities). As shown in **Figure 2**, the CV is lowest for the displayed numerosity 10 and increases with displayed numerosity [$F(6,942) = 11.04$, $p < 0.001$, $\varepsilon GG = 0.92$, $\eta^2_G = 0.05$]. To further explore the relationship between CV and displayed numerosity, we performed a repeated measure correlation (Bakdash and Marusich, 2017) between these two variables. A weak positive correlation emerged [$r_{rm}(971) = 0.16$, 95% CI = [0.10, 0.22], $p < 0.001$], showing that the CV slightly increases with displayed numerosity. The ANOVA also revealed that the CV decreased with age [$F(4,157) = 5.26$, $p < 0.001$, $\eta^2_G = 0.04$; see **Figure 2**] but no interaction was observed [$F(24, 942) < 1$]. This indicates that the overall accuracy increased with age.

To account for putative effects of inflated variance due to small number of trials in each displayed numerosity, we repeated these analyses using the *z*-transformed scores. For both mean chosen numerosity and CV, we calculated the standardized z-scores over all displayed numerosity for each child. The mean *z*-scores were entered into a repeated measure ANOVA with age as between-subject factor. Similar results emerged. In fact, age significantly influenced CV [$F(4,157) = 5.37$, $p < 0.001$] but not mean chosen numerosity [$F(4,157) < 1$].

Distribution of Responses in Approximate Addition and Subtraction

In each trial, the set of five proposed alternatives was sampled from either the lower range of responses (alternatives from 1 to 5, see **Table 1**) or the higher range (alternatives from 3 to 7, see **Table 1**). Therefore, the correct outcome was either the second (high range) and the fourth (low range) smaller proposed alternative. If children were able to solve the calculation, the response pattern should show a non-flat distribution centered on the correct outcome (i.e., second or fourth smaller alternative for high and low range, respectively).

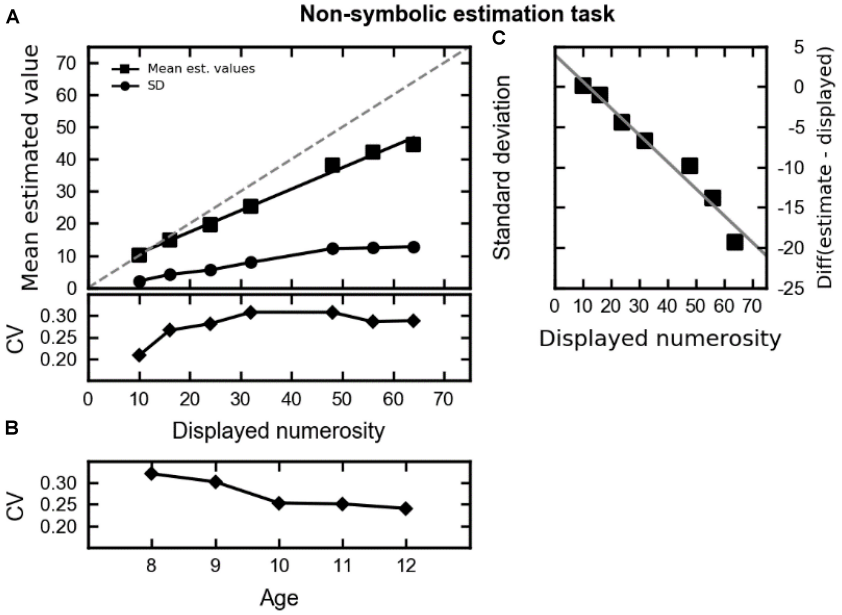

FIGURE 2 | (A) The top part shows the mean chosen numerosities (squares; the black line represents the regression model) and standard deviation (circles) plotted against the displayed numerosity. The gray dashed line represents perfect performance. The lower part reports the mean CV (coefficients of variation) plotted against the displayed numerosity. **(B)** The mean CV plotted against the age groups. **(C)** The difference between chosen numerosity and displayed numerosity plotted against the displayed numerosity. The gray line represents a regression model between the variables.

Mean (arcsine-transformed) percentage of choice was analyzed with a repeated-measure ANOVA with response category (i.e., 1 to 5), range (i.e., low vs. high), and operation (i.e., addition vs. subtraction) as within-subject factors and age (i.e., 8 to 12 years old) as between-subject factor. Results are reported in Supplementary Table S2 (see Appendix B). In particular, both the operation × range × response category interaction [$F(4,628) = 141.89$, $p < 0.001$, εGG = 0.95, generalized $\eta^2 = 0.16$] and the age × range × response category interaction [$F(16,628) = 1.71$, $p = 0.048$, εGG = 0.89, generalized $\eta^2 = 0.01$] were significant. Moreover, the four-way interaction showed a tendency toward significance [$F(16,628) = 1.54$, $p = 0.085$, εGG = 0.95, generalized $\eta^2 < 0.01$]. The tendency of the four-way interaction and **Figure 3** suggest that the performance was different in the two operations. Therefore, to further explore this pattern, two additional ANOVAs were performed on mean percentage of choice with response category and range as within-subject factors and age as between-subject factor, separately for addition and subtraction.

For addition, the main effect of response category was significant [$F(4,628) = 22.06$, $p < 0.001$, εGG = 0.89, generalized $\eta^2 = 0.06$]. Moreover, the age × response category [$F(16,628) = 2.19$, $p = 0.007$, εGG = 0.89, generalized $\eta^2 = 0.03$], the range × response category interaction [$F(4,628) = 223.06$, $p < 0.001$, εGG = 0.87, generalized $\eta^2 = 0.43$] and the three-way interaction [$F(16,628) = 2.07$, $p = 0.012$, εGG = 0.87, generalized $\eta^2 = 0.03$] were significant (**Figure 3**).

For subtraction, only the main effect of response category [$F(4,628) = 19.18$, $p < 0.001$, εGG = 0.89, generalized $\eta^2 = 0.07$] and the age × response category interaction [$F(16,628) = 2.02$, $p = 0.014$, εGG = 0.89, generalized $\eta^2 = 0.03$] were significant, whereas neither the range × response category interaction [$F(4,628) = 2.07$, $p = 0.087$] nor the three-way interaction [$F(16,628) < 1$] reached significance (**Figure 3**). The response distribution for subtraction was flatter, showing that children found more difficult to perform approximate subtraction.

Children's Performance in Approximate Calculation

In order to evaluate children's performance in approximate addition and subtraction, mean chosen response and standard deviation were analyzed with a repeated-measure ANOVA with correct outcome (i.e., 10, 16, 26, and 40) and operation (i.e., addition vs. subtraction) as within-subject factors and age (i.e., 8–12 years old) as between-subject factor. For mean chosen response, the main effect of correct outcome was significant [$F(3,471) = 1685.80$, $p < 0.001$, εGG = 0.60, $\eta^2_G = 0.76$]. Mean chosen responses increased with correct outcome (mean responses: 12.0, 17.3, 24.1, and 32.9 for the outcomes 10, 16, 26, and 40, respectively). Mean chosen responses were greater for addition (mean = 23.2) than for subtraction (mean = 19.9) [$F(1,157) = 93.49$, $p < 0.001$, $\eta^2_G = 0.12$]. Moreover, all the two-way interactions were significant: correct outcome × operation [$F(3,471) = 131.81$, $p < 0.001$, εGG = 0.72, $\eta^2_G = 0.12$], correct outcome × age [$F(12,471) = 2.03$, $p = 0.049$, εGG = 0.60, $\eta^2_G = 0.01$], operation × age [$F(4,157) = 6.24$, $p < 0.001$, $\eta^2_G = 0.04$]. Interestingly, the three-way interaction was also significant [$F(12,471) = 2.78$, $p = 0.004$, εGG = 0.72, $\eta^2_G = 0.01$]. As shown in **Figure 4**, mean chosen values were overestimated for addition compared to subtraction, and this difference was greater

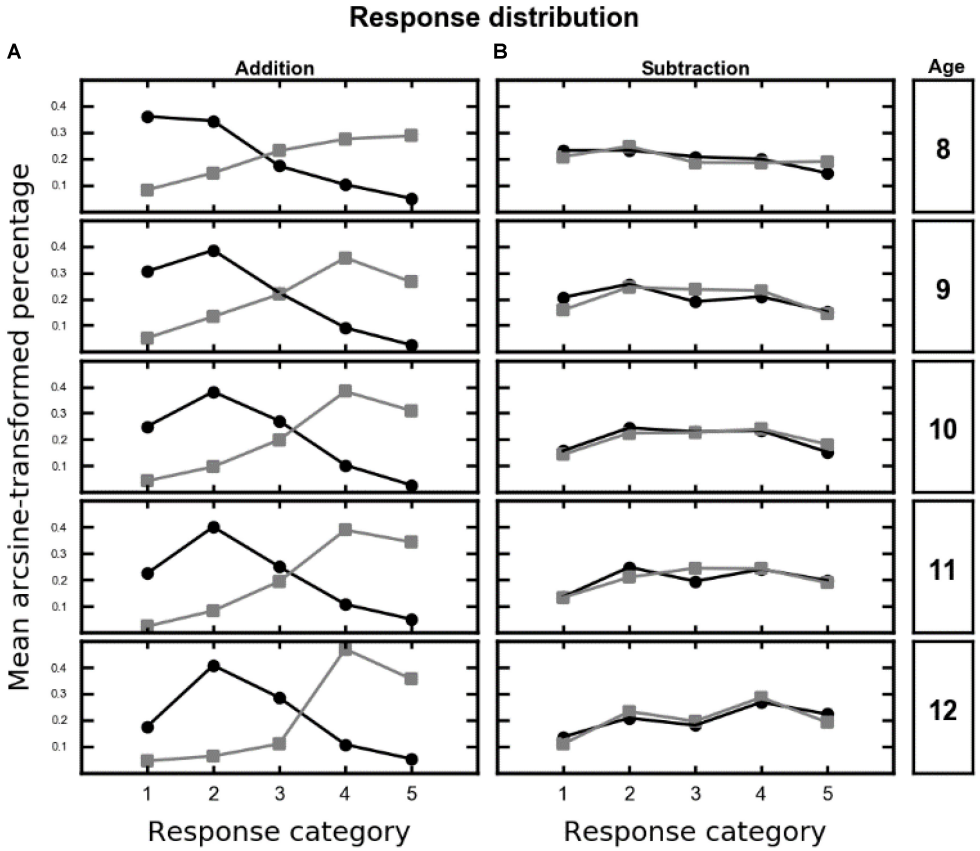

FIGURE 3 | Mean (arcsine-transformed) percentage of choice across the response category (*x*-axis) as a function of range (high: black circles, low: gray squares) and age (from 8 to 12, rows), for addition **(A)** and subtraction **(B)**. For high range the correct outcome is the response category 2, for low range the correct outcome is the response category 4.

for larger numerosities and increased with age. This pattern reflects the OM effect and will be further investigated in the following section.

Standard deviation significantly increased with correct outcome [$F(3,471) = 275.66$, $p < 0.001$, $\varepsilon GG = 0.82$, $\eta^2_G = 0.35$]. However, this increase followed a different pattern in the two operations, as shown by the correct outcome by operation interaction [$F(3,471) = 18.17$, $p < 0.001$, $\varepsilon GG = 0.88$, $\eta^2_G = 0.02$], see **Figure 4**. No other main effects or interactions were significant.

To investigate whether children's mental numerosity representation follows Weber–Fechner law, a third ANOVA was performed on CV with correct outcome and operation as within-subject factors and age as between-subject factor. The main effect of correct outcome was significant [$F(3,471) = 5.88$, $p < 0.001$, $\varepsilon GG = 0.90$, $\eta^2_G = 0.01$] [outcomes 10: mean CV (*SD*) = 0.32 (0.09); outcome 16: 0.31 (0.09); outcome 26: 0.33 (0.09); outcome 40: 0.30 (0.07)]. Moreover, the CV was also significantly smaller for addition (mean = 0.30, *SD* = 0.08) than for subtraction (mean = 0.33, *SD* = 0.08) [$F(1,157) = 30.28$, $p < 0.001$, $\eta^2_G = 0.03$]. Finally, the interaction between correct outcome and operation was significant [$F(3,471) = 7.46$, $p < 0.001$, $\varepsilon GG = 0.96$, $\eta^2_G = 0.01$], see **Figure 4**. To further investigate this interaction, we performed a repeated measure correlation between correct outcome and CV, separately for each operation. For addition, no correlation emerged between CV and correct outcome [$r_{rm}(485) = 0.005$, 95% CI = [−0.08, 0.09], $p = 0.91$]. For subtraction, a weak negative correlation emerged [$r_{rm}(485) = -0.17$, 95% CI = [−0.25, −0.08], $p < 0.001$], showing that mean CV slightly decreased with correct outcome, and thus the variability of the chosen response did not increase proportionally with the mean of the chosen response. These results are not perfectly consistent with the assumption that the underlying mental numerosity representation follows the Weber–Fechner law. However, since the CV did not covary with correct outcome in addition and only weakly correlated with it in subtraction (explained variance: 2.89%), the overall performance did not substantially deviate from this assumption.

Operational Momentum Effect

To investigate the developmental trajectory of the OM effect, the mean response bias was analyzed with a repeated-measure ANOVA with operation as within-subject factor and age as between-subject factor. Response bias was calculated as the mean difference between the logarithm of the chosen response and the logarithm of the correct outcome. Response bias was significantly

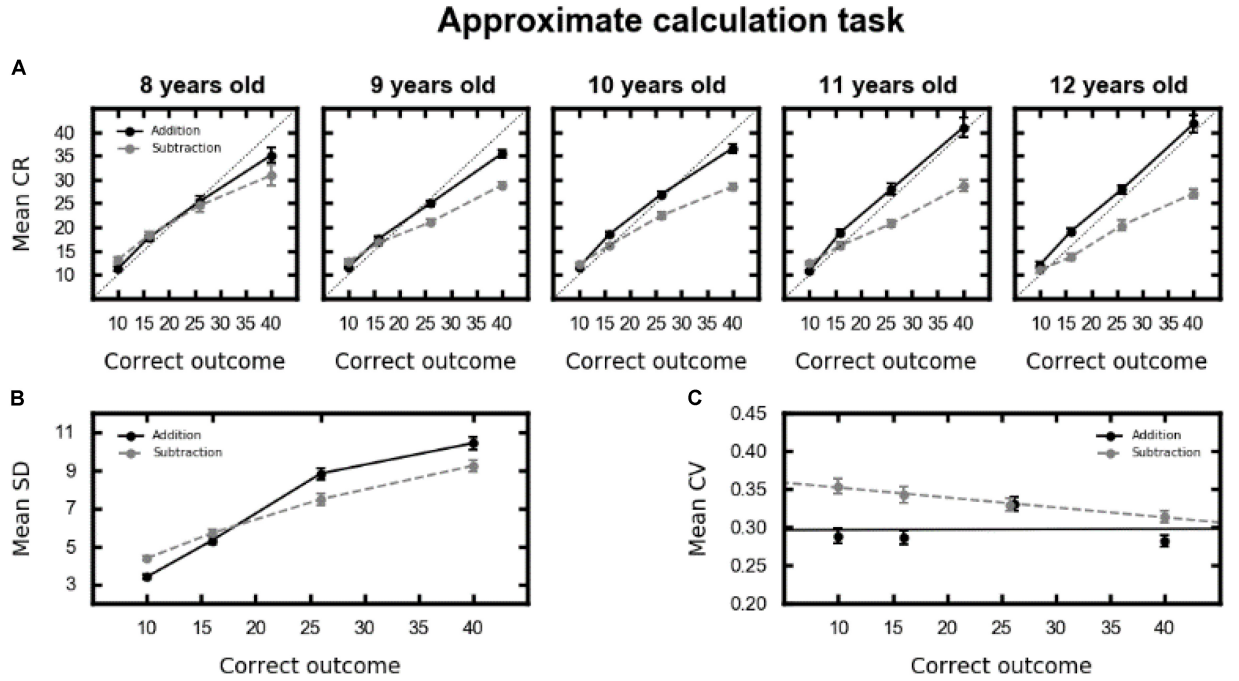

FIGURE 4 | (A) Mean chosen response (CR) as a function of correct outcome (*x*-axis), operation (addition in black, subtraction in gray), and age (columns). The black dotted lines represent perfect performance. **(B)** Mean standard deviation (SD) as a function of correct outcome (*x*-axis) and operation (addition in black, subtraction in gray), collapsed across all ages. **(C)** Mean coefficients of variation (CV) as a function of correct outcome (*x*-axis) and operation (addition in black, subtraction in gray, the lines represent the regression models), collapsed across all ages. In all plots, error bars represent the standard error of the mean.

different between addition (−0.0004, $SD = 0.05$) and subtraction (−0.06, $SD = 0.08$) [$F(1,157) = 60.2$, $p < 0.001$, $\eta^2_G = 0.17$]. The age by operation interaction was also significant [$F(4,157) = 4.45$, $p = 0.002$, $\eta^2_G = 0.06$]. As shown in **Figure 5**, the OM effect monotonically increased with age[1], from no effect for younger children to a strong effect for older children (see **Table 2** for *post hoc* comparison and effect sizes). To further explore the addition and subtraction response biases separately, a second set of one-sample *t*-tests have been performed to evaluate whether they significantly differed from zero (biases significantly different from zero are shown in bold in **Table 2**). As shown in the table, only subtraction biases for the age groups from 9 to 12 were significantly different from zero [all $ts < -4.97$, all $ps < 0.01$].

In Appendix A, we report an additional set of analyses that by and large confirms these findings.

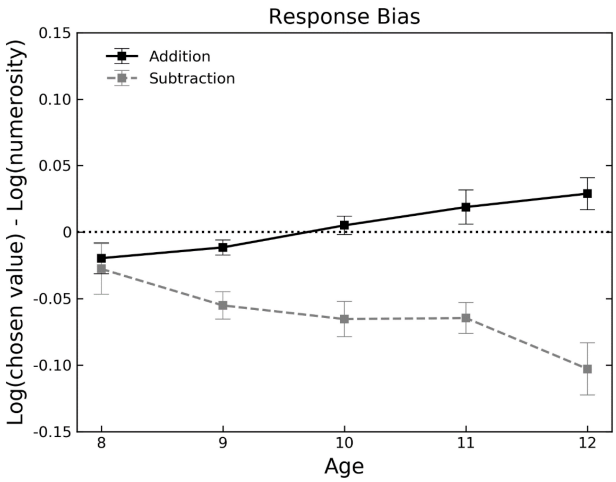

FIGURE 5 | Mean response bias (i.e., difference between the logarithm of the chosen response and the logarithm of the correct outcome) as a function of age and operation (addition in black, subtraction in gray dashed). Error bars represent the standard error of the mean. The horizontal dotted line represents no bias.

DISCUSSION

This study aimed to investigate the developmental trajectory of the OM effect in children aged from 8 to 12 years old and to assess whether the current accounts are able to predict these age-related changes. Concerning the non-symbolic estimation task, consistent with previous research (Izard and Dehaene, 2008; Knops et al., 2014; but for overestimation see Mejias and Schiltz, 2013), children underestimated the cardinality of displayed numerosities and this underestimation increased

[1]Since the sample size is unequal in the different age groups, we also performed two Spearman's correlation analyses between mean response bias and age (in months), separately for addition and subtraction. For addition, there was significant positive correlation [$r = 0.31$, $p < 0.001$]. For subtraction, there was significant negative correlation [$r = -0.24$, $p = 0.002$].

TABLE 2 | *T*-tests comparing the response bias between addition and subtraction in the different age groups.

Age group	*N*	Addition		Subtraction		*t*	df	*p*-value	Cohen's d_z	Hedges' g_{av}
		Mean	*SD*	Mean	*SD*					
8	24	−0.020	0.057	−0.028	0.094	0.4	23	>0.1	0.08	0.10
9	54	−0.012	0.041	**−0.055**	0.075	3.61	53	0.005	0.49	0.71
10	50	0.005	0.048	**−0.065**	0.093	4.55	49	<0.001	0.64	0.94
11	20	0.019	0.058	**−0.065**	0.052	4.52	19	0.002	1.01	1.46
12	14	0.029	0.045	**−0.103**	0.073	5.04	13	0.002	1.35	2.04

All p-values have been corrected with Holm's method. For the calculation of the effect sizes (Cohen's d_z and Hedges' g_{av}) refers to Lakens (2013). Mean response biases significantly different from zero (i.e., one-sample t-tests, separately computed for each operation and age group) are in bold, all ps < 0.01.

with numerosity. Although the CV significantly increased with numerosity, the correlation between the two variable was weak ($r_{rm} = 0.16$). Moreover, both mean estimated values and standard deviation increased with displayed numerosity. This suggests that children's performance was by and large well captured by Weber–Fechner law, even if the CV was not perfectly linear across the entire numerical range. In line with previous findings that suggest that the Weber fraction decreases with age (Piazza et al., 2010; Halberda et al., 2012), the coefficient of variation also significantly decreased with age. Deviations may be due to non-numerical features of the stimulus set, for example. Further studies are needed to fully explain these inconsistencies.

In the approximate addition task, the distribution of responses clearly peaked around the correct outcome showing that children were able to solve these problems. The response distribution for subtraction problems, however, showed a different pattern. The distribution was flat for younger children (8 years old, see **Figure 3**) and in general the two ranges (low vs. high, see **Table 1**) were almost overlapped. Therefore, children found subtraction problems more difficult to solve compared to addition problems, in line with adults (Knops et al., 2009b). However, for subtraction problems, the significant main effect of response category and **Figure 3** suggest that children (at least in the age groups from 9 to 12) did not respond at random but rather selected more often values in the center of the response category range (i.e., 2, 3, 4) compared to the extremes (i.e., 1 and 5). This suggests that children might have used a different strategy to perform subtraction compared to addition. Despite the lower performance on subtractions problems, a clear OM effect emerged in our sample. Importantly, for addition the increase of the OM effect was accompanied by an increase in overall accuracy (see **Figure 3**). That is, while younger children made more and non-systematic errors, older children made less but systematic errors. Interestingly, the OM effect monotonically increases with age. While no effect was present in younger children (8 years-olds), the OM effect (i.e., the relative difference between the estimated responses in addition and subtraction) increased with age. In what follows, we first summarize the findings related to the evolution of the OM effect during childhood, and then we will discuss the implications of these findings for the current accounts of the OM effect (i.e., compression account and attentional shift account).

McCrink and Wynn (2009) found that 9 months old infants exhibit an OM effect similar to that found in adults. Although the similarity between the OM effect found in infants (McCrink and Wynn, 2009) and adults (McCrink et al., 2007; Knops et al., 2009b) would suggest that the OM effect results from inherited mechanisms (since infants are not yet affected by cultural practices) and remains constant during development, a more complex pattern emerges if we consider a previous study (Knops et al., 2013) and the findings reported in the current paper. In fact, contrary to the expected continuity of the OM effect during development, Knops et al. (2013) found an inverse OM effect in 6/7 years old children: subtraction was significantly overestimated compared to addition. Finally, our results showed a monotonic increase of the OM effect with age. This complex developmental pattern indicates that the evolution of the OM effect is not linear. In fact, a standard OM effect emerges in infants (McCrink and Wynn, 2009), an inverse OM effect was found in 6/7 years old children (Knops et al., 2013), and our results show no OM in 8 years old children and a monotonically increasing OM effect from 9 to 12 years old.

How well do the current accounts predict the developmental-related changes of the OM effect? The *compression account* (McCrink et al., 2007) predicts that, due the logarithmic-to-linear shift of the MNL during childhood (Siegler and Opfer, 2003; Siegler and Booth, 2004; Booth and Siegler, 2006, 2008; Laski and Siegler, 2007; Opfer and Siegler, 2007; but for a different perspective see Barth and Paladino, 2011), the OM effect decreases with age. Our result clearly points in the opposite direction showing an increase of the OM effect.

In line with the recycling theory (Dehaene and Cohen, 2007; see also the redeployment theory, Anderson, 2007), which proposes that arithmetic calculation is grounded on the recycling of neural circuits that originally evolved for processing visuospatial information, the *attentional shift account* assumes that the OM effect is driven by the functional relationship between visuospatial attention and mental arithmetic. Strong evidence for the idea that visuospatial attention is co-opted during mental calculation is provided by the fact that the neural activity associated with left/right saccades (i.e., visuospatial orientation) and mental calculation overlap in the posterior superior parietal lobule (Knops et al., 2009a). Using fMRI data, these authors showed that a multivariate classifier algorithm trained to classify the neural activity elicited by leftward and rightward saccades was able to generalize to approximate arithmetic. Without further training, this algorithm was able to distinguish between addition and subtraction

by classifying approximate additions as rightward saccades. The activation of the same neural areas during rightward saccades and approximate addition speaks in favor of the recruitment of attentional shift mechanisms during mental calculation. This hypothesis stipulates a functional coupling between eye movements and arithmetic. A recent study provided confirmatory evidence for this notion (Klein et al., 2014). Participants' eye movements after the first saccade were observed to move to the right during addition problems and to the left in subtraction problems when asked to indicate the location of the result on a labeled line (Klein et al., 2014). Moreover, the redeployment of visuospatial attention during mental calculation seems to be enhanced during formal schooling (Rosenberg-Lee et al., 2011). Finally, on the behavioral level, too, even if spatial-numerical association already emerges in preschoolers, the evidence is mixed. For example, White et al. (2012) found that the SNARC effect emerged during the 2nd year of schooling in British students, that is at around 7 years of age, while 6-year-olds did not show a significant SNARC effect (see also Gibson and Maurer, 2016). Moreover, Yang et al. (2014) found a SNARC effect in kindergarteners (age range: 4.8–6.4 years), 2nd, 3rd, 5th, and 6th graders, while 1st and 4th graders did not show a significant effect (see also Patro and Haman, 2012). Hoffmann et al. (2013) also found mixed evidence for the emergence of the SNARC effect. While all children in the second-term (mean age: 5.8 years old) showed a SNARC effect, in the first-term group (5.5 years old) the effect emerged when a magnitude comparison task preceded a digit color judgment task but not when the task order was inverted. Moreover, in the magnitude comparison task the size of the SNARC effect was related to proficiency with Arabic numbers. This developmental pattern suggests that the spatial-numerical association is still immature in young children. We propose that formal schooling could bolster spatial-numerical associations and hence reinforce movement direction during addition (toward larger numbers) and subtraction (toward smaller numbers). Attentional shifts may implement the core cognitive function to carry out the shifts along the spatial mental number representation and may be affected in at least two ways by the emerging spatial-numerical associations. Either the amount of displacement in the direction of the operation on the MNL increases (i.e., generate a larger and/or more systematic bias) or the variance of displacement is reduced while the overall amplitude remains constant. Therefore, the attentional shift account predicts an increasing OM effect during childhood. Consistent with this prediction, we found a monotonous increase of the OM effect with age.

Although the attentional shift account is consistent with our results, a more complex picture emerges if the results from previous studies are taken into account. In fact, the inverse OM effect found in 6/7 years old children (Knops et al., 2013) is neither explained nor predicted by this account. However, Knops et al. (2013) showed that the direction of the OM effect was related to reorienting attention in a Posner paradigm. The reorientation effect was calculated as the difference in reaction times between valid (i.e., the target stimulus appeared on the left or right of a bidirectional arrow previously presented in the center of the screen) and invalid trials (i.e., the target stimulus appeared opposite the pointing direction of a single-headed arrow). In their study, children who exhibited a smaller reorientation effect (i.e., more proficient to reorient attention after an invalid cue) also had a more regular OM effect (i.e., addition overestimated compare to subtraction). As those authors suggested, it can be hypothesized that the OM effect relies on a fully developed attentional system and on a robust functional association between visuospatial attention and mental calculation. Alternatively, it may suggest that inhibitory control of saccadic eye movements plays a crucial role for the association between attention and arithmetic. We can only speculate as to why an inverse OM effect emerges in 6/7 years old children and the youngest age group of our sample does not show any effect. The more immature attentional system (Rueda et al., 2004; Konrad et al., 2005) and the weaker functional connection between visuospatial processing and mental calculation (Rosenberg-Lee et al., 2011) might be at the origin of the inverse OM effect and its absence in younger children. Namely, the implementation of approximate addition and subtraction would not be yet supported by operation-specific, systematic attentional shifts on the MNL that produce misestimation in the direction of the operation.

The presence of a standard OM effect in infants (McCrink and Wynn, 2009) challenges the idea that the OM effect monotonically increases during childhood due to the consolidation of the engagement of visuospatial processing during mental calculation. However, this contradiction strongly relies on the idea that the development of cognitive performance always reflects linear developmental trajectories. However, as put forward by Siegler (1996), behavior may reflect the prevalence of heuristics and biases that wax and wane over time. That is, while infants may respond according to a given heuristic, the very same heuristic may be less influential during later periods in life. In children, performance in approximate calculation tasks may be performed with the support of the visuospatial system (i.e., the shift of the attentional focus on the MNL), while in infants the heuristic decision may result from simpler processes rather than from more sophisticated attentional mechanisms. Namely, in children (or adults) and infants the heuristic decision might result from different mechanisms. However, more evidence on the development of the OM effect is needed to unravel the cognitive mechanisms that drive the OM at different ages.

This study has some limitations. First, children's performance in subtraction was low compared to addition. The higher difficulty to estimate the result of approximate subtraction could be due to the use of different strategies to perform the two operations. To better understand how children perform approximate calculation, future research should further investigate this difference in performance. Second, despite the fairly large sample, 6/7 years old children were not included, that is the age group that showed the inverse OM effect. Future studies should include a larger age range in order to confirm the inverse OM effect and to further investigate the development of this effect. Third, we did not include any task to measure visuospatial attention. Future studies should investigate whether there is a correlation between the developmental trajectories of visuospatial attention and of the OM effect. Finally, the effect of education

is also accompanied by the maturation of neural network that supports mental calculation. In the analysis we focused on age, future research, however, should also disentangle the influence of age (neural maturation) and grade (education) on the OM effect. These two independent factors could make distinct contribution at various stages of development.

To sum up, we provided a novel finding on the developmental trajectory of the OM effect in children from 8 to 12 years old. The OM effect monotonically increases with age. This developmental pattern is inconsistent with the compression account. On the other hand, the attentional shift account provides a possible explanation of these results based on the functional relationship between visuospatial attention and mental calculation and on the effect of the acquisition of arithmetical skills during formal schooling. The attentional shift account leads to new predictions about a correlation between visuospatial processing and mental calculation which can be addressed in future studies. Our results provide an important empirical constraint to further explore the origin of the OM effect.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of ethics review board of the Federal University of Minas Gerais, Brazil (COEP–UFMG) with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. Informed written consent was obtained from the parents and oral consent from the children. The protocol was approved by the ethics review board of the Federal University of Minas Gerais, Brazil (COEP–UFMG).

AUTHOR CONTRIBUTIONS

PP-C, VH, GW, and AK designed the research. PP-C performed the research. DD and PP-C analyzed the data. DD drafted the manuscript. DD, PP-C, AK, VH, and GW contributed to write and revise the paper.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (DFG) and the Open Access Publication Fund of Humboldt-Universität zu Berlin.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01062/full#supplementary-material

REFERENCES

Anderson, M. L. (2007). Evolution of cognitive function via redeployment of brain areas. *Neuroscientist* 13, 13–21. doi: 10.1177/1073858406294706

Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., and Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven – Escala Especial*. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia, 435.

Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., and Pascual-Leone, J. (2018). Brain areas associated with numbers and calculations in children: meta-analyses of fMRI studies. *Dev. Cogn. Neurosci.* 30, 239–250. doi: 10.1016/j.dcn.2017.08.002

Bakdash, J. Z., and Marusich, L. R. (2017). Repeated measures correlation. *Front. Psychol.* 8:456. doi: 10.3389/fpsyg.2017.00456

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., and Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition* 98, 199–222. doi: 10.1016/j.cognition.2004.09.011

Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x

Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 41, 189–201. doi: 10.1037/0012-1649.41.6.189

Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x

Brown, S. W. (1985). Time perception and attention: the effects of prospective versus retrospective paradigms and task demands on perceived duration. *Percept. Psychophys.* 38, 115–124. doi: 10.3758/BF03198848

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Percept. Psychophys.* 59, 1118–1140. doi: 10.3758/BF03205526

Cantlon, J. F., and Brannon, E. M. (2007). Basic math in monkeys and college students. *PLoS Biol.* 5:e328. doi: 10.1371/journal.pbio.0050328

Chen, Q., and Verguts, T. (2012). Spatial intuition in elementary arithmetic: a neurocomputational account. *PLoS One* 7:e31180. doi: 10.1371/journal.pone.0031180

de Hevia, M. D., Veggiotti, L., Streri, A., and Bonn, C. D. (2017). At birth, humans associate "Few" with Left and "Many" with right. *Curr. Biol.* 27, 3879–3884. doi: 10.1016/j.cub.2017.11.024

Dehaene, S. (1997). *The Number Sense*. New York, NY: Oxford University Press.

Dehaene, S. (2007). "Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation," in *Sensorimotor Foundations of Higher Cognition: Attention and Performance*, eds P. Haggard, Y. Rossetti, and W. Kawato (Oxford: Oxford University Press), 527–574.

Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371

Dehaene, S., and Changeux, J. P. (1993). Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.* 5, 390–407. doi: 10.1162/jocn.1993.5.4.390

Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398. doi: 10.1016/j.neuron.2007.10.004

Fias, W., and Fischer, M. H. (2005). "Spatial representation of number," in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (Hove: Psychology Press), 43–54.

Fischer, M. H., and Knops, A. (2014). Attentional cueing in numerical cognition. *Front. Psychol.* 5:1381. doi: 10.3389/fpsyg.2014.01381

Flombaum, J. I., Junge, J. A., and Hauser, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition* 97, 315–325. doi: 10.1016/j.cognition.2004.09.004

Gibson, L. C., and Maurer, D. (2016). Development of SNARC and distance effects and their relation to mathematical and visuospatial abilities. *J. Exp. Child Psychol.* 150, 301–313. doi: 10.1016/j.jecp.2016.05.009

Glicksohn, J., and Hadad, Y. (2012). Sex differences in time production revisited. *J. Individ. Differ.* 33, 35–42. doi: 10.1027/1614-0001/a000059

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., and Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11116–11120. doi: 10.1073/pnas.1200196109

Hoffmann, D., Hornung, C., Martin, R., and Schiltz, C. (2013). Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds. *J. Exp. Child Psychol.* 116, 775–791. doi: 10.1016/j.jecp.2013.07.013

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684

Izard, V., and Dehaene, S. (2008). Calibrating the mental number line. *Cognition* 106, 1221–1247. doi: 10.1016/j.cognition.2007.06.004

Izard, V., Sann, C., Spelke, E. S., and Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10382–10385. doi: 10.1073/pnas.0812142106

Klein, E., Huber, S., Nuerk, H. C., and Moeller, K. (2014). Operational momentum affects eye fixation behaviour. *Q. J. Exp. Psychol.* 67, 1614–1625. doi: 10.1080/17470218.2014.902976

Knops, A., Dehaene, S., Berteletti, I., and Zorzi, M. (2014). Can approximate mental calculation account for operational momentum in addition and subtraction? *Q. J. Exp. Psychol.* 67, 1541–1556. doi: 10.1080/17470218.2014.890234

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., and Dehaene, S. (2009a). Recruitment of an area involved in eye movements during mental arithmetic. *Science* 324, 1583–1585. doi: 10.1126/science.1171599

Knops, A., Viarouge, A., and Dehaene, S. (2009b). Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. *Atten. Percept. Psychophys.* 71, 803–821. doi: 10.3758/APP.71.4.803

Knops, A., Zitzmann, S., and McCrink, K. (2013). Examining the presence and determinants of operational momentum in childhood. *Front. Psychol.* 4:325. doi: 10.3389/fpsyg.2013.00325

Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, J., et al. (2005). Development of attentional networks: an fMRI study with children and adults. *Neuroimage* 28, 429–439. doi: 10.1016/j.neuroimage.2005.06.065

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863

Laski, E., and Siegler, R. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Dev.* 78, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x

Lindemann, O., and Tira, M. (2011). Operational momentum in numerosity production judgments of multi-digit number problems. *J. Psychol.* 219, 50–57. doi: 10.1027/2151-2604/a000046

Marghetis, T., Núñez, R., and Bergen, B. K. (2014). Doing arithmetic by hand: hand movements during exact arithmetic reveal systematic, dynamic spatial processing. *Q. J. Exp. Psychol.* 67, 1579–1596. doi: 10.1080/17470218.2014.897359

Masson, N., and Pesenti, M. (2014). Attentional bias induced by solving simple and complex addition and subtraction problems. *Q. J. Exp. Psychol.* 67, 1514–1526. doi: 10.1080/17470218.2014.903985

Masson, N., and Pesenti, M. (2016). Interference of lateralized distractors on arithmetic problem solving: a functional role for attention shifts in mental calculation. *Psychol. Res.* 80, 640–651. doi: 10.1007/s00426-015-0668-7

Masson, N., Pesenti, M., Coyette, F., Andres, M., and Dormal, V. (2017a). Shifts of spatial attention underlie numerical comparison and mental arithmetic: evidence from a patient with right unilateral neglect. *Neuropsychology* 31, 822–833. doi: 10.1037/neu0000361

Masson, N., Pesenti, M., and Dormal, V. (2017b). Impact of optokinetic stimulation on mental arithmetic. *Psychol. Res.* 81, 840–849. doi: 10.1007/s00426-016-0784-z

Mathieu, R., Epinat-Duclos, J., Léone, J., Fayol, M., Thevenot, C., and Prado, J. (2017). Hippocampal spatial mechanisms relate to the development of arithmetic symbol processing in children. *Dev. Cogn. Neurosci.* 30, 324–332. doi: 10.1016/j.dcn.2017.06.001

Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., and Prado, J. (2016). Running the number line: rapid shifts of attention in single-digit arithmetic. *Cognition* 146, 229–239. doi: 10.1016/j.cognition.2015.10.002

McCrink, K., Dehaene, S., and Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Percept. Psychophys.* 69, 1324–1333. doi: 10.3758/BF03192949

McCrink, K., and Hubbard, T. (2017). Dividing attention increases operational momentum. *J. Numer. Cogn.* 3, 230–245. doi: 10.5964/jnc.v3i2.34

McCrink, K., and Wynn, K. (2009). Operational momentum in large number addition and subtraction by 9-month-olds. *J. Exp. Child Psychol.* 103, 400–408. doi: 10.1016/j.jecp.2009.01.013

Mejias, S., and Schiltz, C. (2013). Estimation abilities of large numerosities in Kindergartners. *Front. Psychol.* 4:518. doi: 10.3389/fpsyg.2013.00518

Mioni, G., Stablum, F., McClintock, S. M., and Grondin, S. (2014). Different methods for reproducing time, different results. *Atten. Percept. Psychophys.* 76, 675–681. doi: 10.3758/s13414-014-0625-3

Nieder, A., and Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157. doi: 10.1016/S0896-6273(02)01144-3

Opfer, J., and Siegler, S. (2007). Representational change and children's numerical estimation. *Cogn. Psychol.* 55, 169–195. doi: 10.1016/j.cogpsych.2006.09.002

Patro, K., and Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006

Peters, L., and De Smedt, B. (2018). Arithmetic in the developing brain: a review of brain imaging studies. *Dev. Cogn. Neurosci.* 30, 265–279. doi: 10.1016/j.dcn.2017.05.002

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends Cogn. Sci.* 14, 542–551. doi: 10.1016/j.tics.2010.09.008

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition* 116, 33–41. doi: 10.1016/j.cognition.2010.03.012

Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003

Pinheiro-Chagas, P., Dotan, D., Piazza, M., and Dehaene, S. (2017). Finger tracking reveals the covert stages of mental arithmetic. *Open Mind* 1, 30–41. doi: 10.1162/OPMI_a_00003

Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., et al. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One* 9:e111155. doi: 10.1371/journal.pone.0111155

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ranzini, M., Lisi, M., Blini, E., Pitteri, M., Treccani, B., Priftis, K., et al. (2015). Larger, smaller, odd or even? Task-specific effects of optokinetic stimulation on the mental number space. *J. Cogn. Psychol.* 27, 459–470. doi: 10.1080/20445911.2014.941847

Ranzini, M., Lisi, M., and Zorzi, M. (2016). Voluntary eye movements direct attention on the mental number space. *Psychol. Res.* 80, 389–398. doi: 10.1007/s00426-015-0741-2

Rosenberg-Lee, M., Barth, M., and Menon, V. (2011). What difference does a year of schooling make? Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *Neuroimage* 57, 796–808. doi: 10.1016/j.neuroimage.2011.05.013

RStudio Team (2015). *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Rugani, R., and Sartori, L. (2016). Numbers in action. *Front. Hum. Neurosci.* 10:388. doi: 10.3389/fnhum.2016.00388

Sallilas, E., El Yagoubi, R., and Semenza, C. (2008). Sensory and cognitive processes of shifts of spatial attention induced by numbers: an ERP study. *Cortex* 44, 406–413. doi: 10.1016/j.cortex.2007.08.006

Santens, S., and Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition* 108, 263–270. doi: 10.1016/j.cognition.2008.01.002

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. Oxford: Oxford University Press.

Siegler, R. S., and Booth, J. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., and Opfer, J. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438

Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A. C., et al. (2017). Increased arithmetic complexity is associated with domain-general but not domain-specific magnitude processing in children: a simultaneous fNIRS-EEG study. *Cogn. Affect. Behav. Neurosci.* 17, 724–736. doi: 10.3758/s13415-017-0508-x

Soltanlou, M., Artemenko, C., Ehlis, A. C., Huber, S., Fallgatter, A. J., Dresler, T., et al. (2018). Reduction but no shift in brain activation after arithmetic learning in children: a simultaneous fNIRS-EEG study. *Sci. Rep.* 8:1707. doi: 10.1038/s41598-018-20007-x

Stein, L. M. (1994). *TDE – Teste do Desempenho Escolar. Manual Para Aplicação e Interpretação*. São Paulo: Casa do Psicólogo, 204.

White, S., Szucs, D., and Soltész, F. (2012). Symbolic number: The integration of magnitude and spatial representations in children aged 6 to 8 years. *Front. Psychol.* 2:392. doi: 10.3389/fpsyg.2011.00392

Wynn, K. (1992). Addition and subtraction by human infants. *Nature* 358, 749–750. doi: 10.1038/358749a0

Xu, F., and Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition* 74, B1–B11. doi: 10.1016/S0010-0277(99)00066-9

Yang, T., Chen, C., Zhou, X., Xu, J., Dong, Q., and Chen, C. (2014). Development of spatial representation of numbers: a study of the SNARC effect in Chinese children. *J. Exp. Child Psychol.* 117, 1–11. doi: 10.1016/j.jecp.2013.07.013

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Journal of Numerical Cognition
jnc.psychopen.eu | 2363-8761

Editorial

Domain-General Factors Influencing Numerical and Arithmetic Processing

André Knops*[ab], Hans-Christoph Nuerk[cde], Silke M. Göbel[f]

[a] Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany. **[b]** CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, University Paris Descartes, Sorbonne Paris Cité, Paris, France. **[c]** Fachbereich Psychologie, Universität Tübingen, Tübingen, Germany. **[d]** IWM- Institut für Wissensmedien, Tübingen, Germany. **[e]** LEAD Graduate School and Research Network, Tübingen, Germany. **[f]** Department of Psychology, University of York, York, United Kingdom.

Abstract

This special issue contains 18 articles that address the question how numerical processes interact with domain-general factors. We start the editorial with a discussion of how to define domain-general versus domain-specific factors and then discuss the contributions to this special issue grouped into two core numerical domains that are subject to domain-general influences (see Figure 1). The first group of contributions addresses the question how numbers interact with spatial factors. The second group of contributions is concerned with factors that determine and predict arithmetic understanding, performance and development. This special issue shows that domain-general (Table 1a) as well as domain-specific (Table 1b) abilities influence numerical and arithmetic performance virtually at all levels and make it clear that for the field of numerical cognition a sole focus on one or several domain-specific factors like the approximate number system or spatial-numerical associations is not sufficient. Vice versa, in most studies that included domain-general and domain-specific variables, domain-specific numerical variables predicted arithmetic performance above and beyond domain-general variables. Therefore, a sole focus on domain-general aspects such as, for example, working memory, to explain, predict and foster arithmetic learning is also not sufficient. Based on the articles in this special issue we conclude that both domain-general and domain-specific factors contribute to numerical cognition. But the how, why and when of their contribution still needs to be better understood. We hope that this special issue may be helpful to readers in constraining future theory and model building about the interplay of domain-specific and domain-general factors.

Keywords: spatial-numerical associations, arithmetic, intervention studies, spatial skills, working memory, modules

Journal of Numerical Cognition, 2017, Vol. 3(2), 112–132, doi:10.5964/jnc.v3i2.159

Published (VoR): 2017-12-22.

*Corresponding author at: Humboldt-Universität zu Berlin, Department of Psychology, Unter den Linden 6, D-10099 Berlin, Germany. E-mail: knops.andre@gmail.com

Numerical cognition encompasses a broad variety of cognitive and neural processes related to the perception, understanding and manipulation of numerical content. Hence, when investigating numerical cognition, we are not looking at an encapsulated cognitive module, supported by a single neural system but rather at a widespread network of interrelated cognitive processes with complex neural underpinnings. Much like human behaviour cannot be fully understood when leaving aside the social interactions and influences, we need to understand to what extent the core numerical processes are influenced and/or mediated by domain-general factors and other domains such as spatial skills and language.

Domain-Generality Versus Domain-Specificity

Our current understanding of the terms domain-general and domain-specific factors has been shaped by discussions about whether there are domain-specific modules in the mind (Fodor, 1983) and whether infants enter the world with innately pre-specified core knowledge (for numerical cognition, see e.g. Baillargeon & Carey, 2012; Dehaene, 2001; Rugani, Vallortigara, Priftis, & Regolin, 2015, but see Núñez, 2017; Patro & Nuerk 2017, for critical valuation). Fodor, in his influential book 'The Modularity of Mind', proposed modules, described by Elman et al. (1998) as 'mental/neural systems that [..] are uniquely suited to and configured for a particular task and no other task' (page 36). Fodor listed nine key criteria that a module has to satisfy; domain-specificity is one of those key criteria, i.e. a module per definition deals exclusively with a single type of information. Ever since this proposal it has been debated whether modules exist. Language and face recognition for example have been put forward as candidates for modules. Fodor also proposed central systems that cut across modules, and called those structures domain-neutral. Today the term "domain-general processes" is more commonly used for those structures and processes.

In Fodor's definition modules need to be innately pre-specified. Along a similar vein, developmental psychologists have proposed that infants possess innately pre-specified domain-specific core knowledge (Dehaene, 2001; de Hevia, Izard, Coubart, Spelke, & Streri, 2014) which supports their early learning from experience. For example, Feigenson, Dehaene, and Spelke (2004) postulated two core systems of numerical representations, one system for representing large numerosities approximately and one system for representing small numbers of objects exactly, that are already present in preverbal infants and non-human animals.

'Modularity' and 'domain-specificity' have often been lumped together, but domain-specificity does not have to imply innateness. Clearly, specified learned systems can also be domain-specific, e.g. cycling, typing, and piano playing do not have to be innate (Elman et al., 1998). For the case of numerical cognition, it has also been proposed that domain-specific modules are a product of neural recycling, i.e., of fast and automatic learning and enculturation, which may start directly after or even before birth in humans (Verguts & Fias, 2004, for a model; Patro, Nuerk, & Cress, 2016 for an enculturation account; Schleger et al., 2014, for magnitude processing in fetuses). Furthermore, the term domain-specificity has been applied to at least five different levels: domain-specific tasks, domain-specific behaviours, domain-specific representations, domain-specific processing mechanisms and domain-specific genes (Elman et al., 1998). Articles in this special issue cover all levels except the domain-specific genetic level.

More recent discussions suggest that the distinction between domain-general and domain-specific processes might be too crude. In practice, it can be a matter of perspective. While for a researcher interested in numerical cognition, symbolic number processing might be domain-specific and spatial skills might be defined as domain-general (or at least domain-overlapping), for spatial cognition researchers spatial skills might be domain-specific (see Cornu, Hornung, Schiltz, and Martin, 2017, this issue) and symbolic thinking might be seen as a domain-general skill. At a more abstract level, the distinction between domain-general versus domain-specific might be artificial, because different processes might be better conceptualised on a continuum, with processes ordered from being relevant for fewer (but not only one) to more (but not all) domains ('domain-relevance', see Karmiloff-Smith, 2015). Thus, the terms domain-general and domain-specific might represent theoretical and rarely reached categorical endpoints on a continuum.

This is also highlighted by the range of topics in this special issue (for an overview see Table 1a/1b and Figure 1). Most competencies fall somewhere in the middle of this continuum: they are important for more than one domain, but not for all domains. For example while understanding of ordinality and magnitude processing clearly is important for numerical processing, these two competencies are not unique to the number domain. We can order letters, days of the week and process the size or magnitude of animals or space or time (see Bueti & Walsh, 2009). This has led some researchers to question whether purely domain-specific representations do actually exist. Cantlon, Platt, and Brannon (2009), for example, question the idea of a domain-specific system solely devoted to numerical processing that is independent of other types of quantity judgements.

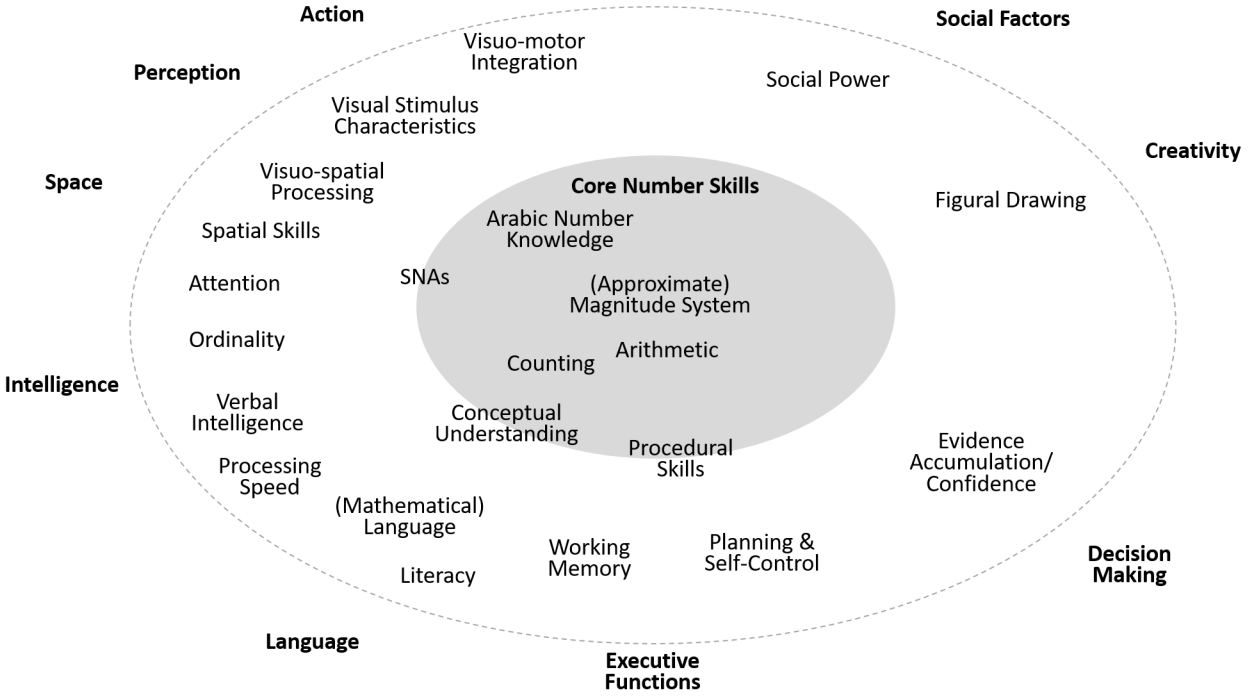

Figure 1. Schematic of the specific examples (within the dotted ellipse) of domain-general factors (in bold print outside the dotted ellipse) whose relationships to domain-specific numerical competencies (within the grey ellipse) were assessed in this special issue.

Karmiloff-Smith (2015) in a developmental framework of domain-relevance suggests how the continuum of domain-general to domain-specific processes might develop: the infant brain starts out with a number of basic-level processing tendencies. Each of these tendencies might be more relevant to the processing of certain different kinds of input over others, i.e. more relevant to a particular domain, and thus can become more domain-specific over time. Dehaene and Cohen's (2007) neuronal recycling hypothesis could be seen as a potential neural explanation as to why some factors are more domain-relevant than others for a particular domain. They propose that evolutionary-speaking recent cultural inventions such as reading, writing and arithmetic are using evolutionarily older cortical circuits that were devoted to different but similar functions such as spatial transformations, object and scene recognition. Evolutionary-speaking younger functions would then share the same structural constraints as the original functions and this could provide an explanation for some cross-domain interactions as well as for why some functions are more relevant to a particular domain.

At a more fine-grained level, some skills might be more relevant for particular numerical tasks or representations but less so for other numerical processes. For example, working memory capacity influences complex addition, particularly if it includes a carry-procedure, more strongly than it does the retrieval of rote-learned simple multiplication facts (Fürst & Hitch, 2000; Imbo, Vandierendonck, & De Rammelaere, 2007; Imbo, Vandierendonck, & Vergauwe, 2007; Seitz & Schumann-Hengsteler, 2002; for comparable contribution of verbal and spatial WM to subtraction and multiplication see Cavdaroglu & Knops, 2016).

The present special issue comprises 18 articles that address the question how numerical processes interact with domain-general factors from different angles. To provide the reader with an overview, we subsumed the contributions to this special issue under two core numerical domains with several subdomains (see Figure 1, Table 1a/1b) that are subject to domain-general influences. First, a number of contributions address the question how numbers interact with spatial factors. This includes the question how the visual system extracts numerosity from a visual scene where items are distributed in space, how directional mappings of numbers to space influence arithmetic operations, and how the spatial layout of the mental number line affects binary choice behavior and whether motor trajectories provide a direct access to this mapping. Second, a pressing question is what factors determine and predict arithmetic understanding and performance. This is important from a theoretical point of view to better understand the cognitive organization of numerical competencies. It also is of large practical importance because it can inform practitioners how to design educational curriculae and remedy measures. This second theme contains contributions that investigate predictors of mental arithmetic from a developmental perspective as well as in adults and special populations. Two studies specifically address the question what factors are best targeted in training measures in order to improve numerical competencies: domain-general or domain-specific factors.

The present results make it very clear that mental arithmetic is subject to influences from a broad variety of domain-general factors. These include but are not limited to the following: mathematical language skills, sustained attention, conceptual understanding and creativity. The importance of working memory appears to be particularly controversial since some find that working memory training does not affect arithmetic performance while others report improved numerical understanding after working memory training.

As can be seen from Figure 1, there are many different domain-general factors and domain-specific factors, which have been associated with different spatial-numerical and arithmetic effects and capabilities. In Table 1a/1b we tried to summarize the findings. The domain-general factors can be found in Table 1a. The domain-specific factors can be found in Table 1b. The numerical and arithmetic effects and capabilities can be found in different columns. So, for example, the study by Gilmore et al. (2017, this issue) is located in the working memory row and the arithmetic column, because the paper by Gilmore et al. investigated (among other factors) the influence of working memory on arithmetic.

The names of the first authors of the studies are either depicted in bold or in italic font. Bold font means that this study found an influence, while italic font means, this study did not find an influence. Often, there were different analyses reported within the same paper. In those cases, bold/italic in the table refers to the most complex analysis results (e.g., multiple regression instead of raw correlation). So, if a domain-general factor X had a raw correlation with the target variable (e.g., arithmetic performance), but no influence in the multiple regression, because this variance could be better explained by other variables in the study, then this factor X would be italic, because it does not explain unique variance.

What can be seen immediately from the overview in Table 1a/1b is that there is a mixture of bold (significant influence) and italic (no significant influence) for virtually all variables investigated in our special issue. Often there is even the same study in bold and italic in the same cell, because there was an influence of a specific domain-general factor X in one condition or for one of several age groups, but not in another condition or another age group. Sometimes outcomes depended on the target variable. For instance, Nemati et al. found an influence of planning (Tower of London), but not of self-control on the *accuracy* of arithmetic, however, an influence of self-control, but not of planning on *response times* for arithmetic.

At this point, it is very important to note that this does not mean that studies, which report opposite results (i.e., bold and italic names in the same cell), are necessarily contradicting each other. On the contrary, the studies differ in multiple aspects. They often use different operationalisations of the underlying constructs, consider different co-variates in the analyses, use different paradigms, stimuli and effects for the to-be-predicted target variables, investigate different age groups, employ different designs (e.g., experimental, correlational, interventions) and different uni- and multivariate analyses; in sum, they reflect the variety present in the field studying domain-general and domain-specific influences on numerical processing. We will discuss this in more detail later. However, what can be said after a short inspection of Table 1a/1b is that the result about the influence of a particular domain-general factor in one paradigm, with one type of stimuli or effect of interest, with a particular choice of target variables, for a particular age group in a particular design and with a particular analysis can hardly be generalized to the field as such. Rather, we need the full picture and the full variety of these manipulations to arrive at a more complete picture of the influence of domain-specific and domain-general factors on numerical cognition.

Although Table 1a/1b may seem quite complex at first sight, it reflects of course a major information reduction of the single studies to provide a rough overview. As always for such information reduction, some construct names for domain-general and domain-specific factors can be controversial. Even more importantly, for some cases, it can be discussed in which cell they should be located or not and even if the name should be in bold or italic or both (especially, if multiple analyses are run as for instance in the paper by Purpura et al., 2017, this issue). To get full insight, we, of course, recommend reading the papers. However, in the following, we will briefly summarize the major findings of the contributions in a more detailed way than in Table 1a/1b. The organization follows the above-described major domains in numerical cognition (as used in Table 1a/1b and Figure 1).

Table 1a

Domain-General Influences on Numerical Processing and Arithmetic Reported in the Studies in This Special Issue

	Numerical and arithmetic effect/capability							
Domain-general influence	**Extension: Approximate**	**Extension: Exact**	**Direction: Implicit Cardinal**	**Direction: Implicit Operations**	**Arithmetic**	**Numerical Identification**	**Counting**	**Arabic Magnitude Comparison**
Working Memory	*Honoré*	**Ramani** *Honoré*			**Gilmore** **Kroesbergen** **Nemati**[g] **Purpura** *Honoré* *Nemati* *Purpura* *Ramani*	**Ramani**	*Honoré*	*Honoré*
Executive Functions					**Nemati**[g] **Purpura** *Nemati* *Purpura*			
Intelligence		*Cornu*			**Kroesbergen** *Cornu*			
Attention				**Katz** **McCrink** *Katz*[d]	*Ashkenazi*			
Visuo-Spatial Processing	**Anobile (Crollen)**[a]	**Cornu (Crollen)**	**(Crollen) Georges**[b] *Georges*		**Cornu (Crollen)**			
Visuo-Motor Integration		*Cornu*			**Cornu**			
Language					**Ashkenazi** **Purpura** *Purpura*			
Mathematical Language[c]					**Purpura**			
Ordinality (non-numerical)			*Schröder*[f]					
Processing/Perceptual Speed					**Ashkenazi** *Purpura*			
Face Recognition			**Alonso-Diaz**					
Self Control					**Nemati**[g] *Nemati*			
Self Regulation					*Nemati*			
Social Power		**Huber**[e] *Huber*						
Creativity					**Kroesbergen**			
Socio-Economic Status					*Purpura*			

Note. Studies referenced by first author. **Bold reference:** Significant prediction/influence/intervention effect /analogous effect, when other variables in the study were considered; *italic reference:* No prediction/influence/intervention effect /analogous effect, when other variables in the study were considered. When several models were computed, we chose the best model in the manuscript (e.g., most variance explained, best fit). Other models may come to different results.

[a]Note that Crollen et al. is a review of existing studies, not presenting new empirical data. Therefore, parentheses were used.

[b]Georges et al. examined correlations, regressions, moderations. One significant raw correlation disappeared in regression and moderation analysis. In one moderation analysis an interaction between SNARC and arithmetic prevailed, but arithmetic itself did not predict SNARC.
[c]Purpura et al.: "Mathematical language has been classified in this study as a domain-general variable, but it should be noted that it also highly overlaps with domain-specific skills as it is comprised of content-specific language."
[d]Katz et al. found attention correlated with OM effects in the non-symbolic, but not in the symbolic condition.
[e]Huber et al. found an influence of social power in a number line length estimation task in the "increase" condition, but not in the "decrease" condition.
[f]Schröder et al. found non-significant correlations of non-numerical (weekdays) and numerical SNARC in three of four analyses. In the remaining analysis, they observed $p = .092$, which would be significant, when tested one-sided.
[g]In Nemati et al.'s paper Planning (Tower of London) predicted accuracy, but not RT, while Self-control predicted RT, but not accuracy. Working memory was not a significant predictor in the regression, but was a predictor in the mediation analysis.

Table 1b

Domain-Specific Influences on Numerical Processing and Arithmetic Reported in the Studies in This Special Issue

	Numerical and arithmetic effect/capability							
Domain-specific influence	**Extension: Approximate**	**Extension: Exact**	**Direction: Implicit Cardinal**	**Direction: Implicit Operations**	**Arithmetic**	**Numerical Identification**	**Counting**	**Arabic Magnitude Comparison**
Extension Approximate					**Purpura[b]** *Kroesbergen* *Purpura*			
Extension Exact					**Kroesbergen**			
Direction Implicit Cardinal					*Georges[a]*			
Arithmetic Performance			*Georges[a]*					
Counting Abilities		*Cornu*			*Cornu*			
Knowledge of Arabic Numbers		*Cornu*			**Cornu**			
Procedural Skills					**Gilmore**			
Conceptual Knowledge		**Ramani**			**Gilmore** *Ramani*	**Ramani**		
Ordering				**Macchi Cassia[c]** *Macchi Cassia*				
Symbolic Magnitude Comparison					**Kroesbergen**			

Note. Studies referenced by first author. **Bold reference:** Significant prediction/influence/intervention effect /analogous effect, when other variables in the study were considered; *italic reference:* No prediction/influence/intervention effect /analogous effect, when other variables in the study were considered. When several models were computed, we chose the best model in the manuscript (e.g., most variance explained, best fit). Other models may come to different results.
[a]Georges et al. examined correlations, regressions, moderations. One significant raw correlation disappeared in regression and moderation analysis. In one moderation analysis an interaction between SNARC and arithmetic prevailed, but arithmetic itself did not predict SNARC.
[b]Purpura et al. presents different response cart tree analyses for different age groups and for high and low performance prediction. Only mathematical language was (almost) consistently predictive in all analyses.
[c]The OM modulation was absent for size ordering, but present for ordering symbolic and non-symbolic sequences.

On the Influence of Spatial Factors and the Association Between Numbers and Space

The metaphor of the mental number line (MNL), a spatially ordered representation of numerical magnitude, is often used to describe the mental representation of cardinal values and the interaction between representations of number and space has been an active research area for decades now (Hubbard, Piazza, Pinel, & Dehaene, 2005; Van Dijck, Ginsburg, Girelli, & Gevers, 2015). Recently, a taxonomy of spatial-numerical associations (SNAs; Cipora, Schroeder, Soltanlou, & Nuerk, in press) has been proposed that is helpful in the current context to situate the different contributions to this special issue. The central distinction in this taxonomy is based on non-directional (henceforth called extensions) vs. directional associations between numerical and physical space. While an extension describes certain spatial qualities of an object (e.g., x is wide and high), directionality is topological in nature, because it only refers to an object's location within certain reference frames (e.g., x stands to the left of y). Within non-directional SNAs a distinction is made between effects based on spatial and numerical cardinal sizes on the one hand, and those based on spatial and numerical intervals (i.e., distinguished parts of the physical or numerical whole) on the other. Within directional SNAs, implicit activation of directional representation (e.g., in a parity judgment task) is distinguished from an explicit one (e.g., counting of spatially aligned objects). Within each of these subcategories of directional SNAs, the coding of cardinality, ordinality and functions form separate SNA instances.

According to this taxonomy, the approximate number system can be categorised as an extensive SNA since the activation of an approximate numerosity is conceptualised as an activation of a magnitude range on the MNL with the peak activation representing the most probable output to other cognitive systems. According to the number sense hypothesis, numerical information is internally represented by an analogue magnitude code in an approximate manner that allows for a numerical estimation of a set of items (e.g., a set of dots). The analogue magnitude code is invariant to input modality, format, and to non-numerical stimulus aspects such as density or the overall surface covered by the items. According to this approach, numerosity is a principal feature of our environment that can be directly sensed, comparable to color, contrast, or brightness. A concurrent model proposes that numerosity is derived indirectly from non-numerical stimulus dimension such as density, for example (Morgan, Raphael, Tibber, & Dakin, 2014; see also Gevers, Cohen Kadosh, & Gebuis, 2016).

Anobile, Cicchini, Pomè, and Burr (2017, this issue) tested a straight-forward prediction of the indirect model: When connecting individual items, numerosity estimates should increase due to more texture information in the high frequency range. Contrary to this prediction, the current results reveal reduced estimates after connecting individual items in a medium numerical range. When increasing the number of items, however, individuation becomes more and more difficult and texture-density mechanisms come into play. This effect is accentuated by connecting individual items. The results support the idea that approximate numerosity estimation is governed by three different regimes; a subitizing regime for very small quantities from one to about four, an estimation regime where individual items can be segregated, and a texture-density regime when the items in a set get too crowded.

Huber, Bloechle, Dackermann, Scholl, Sassenberg, and Moeller (2017, this issue) tested whether size estimations are influenced by social factors. Since participants were asked to adjust the length of a line in order

to represent previously presented Arabic digits, this falls in the category of extensive SNAs with cardinal magnitudes. Participants who had previously been associated with low social power overestimated line length compared to a control group without social power manipulation. In contrast, participants who had previously been associated with high social power were more accurate in their estimates. Together, this pattern of results may be the result of differentially experienced task demands such that high perceived task demand in the low social power group led to overestimation. On a more general note, these results provide support for the idea that perceived social power influences how we perceive the world.

When the association between numbers and space entails the relative position of one object with respect to another, Cipora and colleagues (in press) speak of directional SNAs. In a paradigm, where these associations remain implicit, Schroeder, Nuerk, and Plewnia (2017b, this issue) examine the relation between numbers and space by asking whether or not ordinal judgments of numerical and ordered sequences such as days of the week share a common metric. By analysing individual differences, Schroeder and colleagues did not observe strong evidence for a common construct. Rather, the correlations between corresponding SNARC coefficients were overall low and even vanished after standardisation. From a psychometric point of view, this very low construct validity suggests that the idea that different SNA for ordinal and cardinal metrics do not rely on one and the same underlying construct. Georges, Hoffmann, and Schiltz (2017, this issue) directly address the question whether implicit and explicit SNAs arise from a single predominant account or whether task-specific coding mechanisms underlie these SNAs. They took the SNARC effect (spatial numerical association of response codes; Dehaene, Bossini, & Giraux, 1993) in a parity judgment task (implicit) and a magnitude comparison task (explicit) as a test bed. No correlation between the SNAs from these paradigms was observed. Additionally, the implicit and explicit SNAs were predicted by different variables, namely arithmetic performance and visualization profile, respectively. The authors conclude that visuospatial coding mechanisms contribute to explicit SNAs only, hence supporting the distinction between these SNAs as proposed in the taxonomy of Cipora and colleagues (in press).

While there is no doubt that representations of number and space interact, one question that remains controversial is how the spatial layout of the hypothesized MNL can be assessed best. That is, what paradigm can be used to obtain the best possible measure of the MNL metrics? Since the SNARC effect can be explained by at least five alternative, not necessarily spatial accounts (for an overview, see Schroeder, Nuerk, & Plewnia, 2017a, Figure 5), researchers recently shifted to alternative paradigms. In one such approach the reach trajectories are recorded and the deviations between different conditions are sometimes analyzed as an index for a penetration of the underlying cognitive representation. Song and Nakayama (2008), for example, found that in a number comparison task manual movement trajectories deviated more with larger numerical distance to the reference number compared to smaller numerical distances. These results were interpreted as “direct evidence for a spatial number representation” (p. 1002). On a more general note, these results are thought to reveal “information about internal states as they unfold over time” (p. 1002). Hence, the authors assume a direct mapping between internal representations and the spatial layout of the reach.

Alternatively, one may conceive of reach trajectories as being modulated by the amount of response competition. That is, reach trajectories to either of several and simultaneously competing targets may reflect the confidence in the particular choice made, which in turn can be understood as the difference in accumulated evidence for the present options. Alonso-Diaz, Gaffin-Cahn, Mahon, and Cantlon (2017, this issue) tested these two hypotheses using (a) Arabic numerals and (b) facial expressions. The authors found that reach trajectories

were equally sensitive to stimulus similarity (i.e. numerical distance or similarity of facial expressions). These results support the domain-general response competition account and cast some doubt on the idea that the features of the internal (and mostly unconscious) mental magnitude representation are directly mapped onto external space and movements therein.

It becomes clear that we are still at the very beginning of our understanding of the contribution of domain-general and domain-specific influences to directional space-number associations. While some papers in our special issue suggest that numerical representations are part of one common and much more general mental magnitude representation, others suggest that even within the numerical domain, cardinal and ordinal associations or explicit and implicit magnitude associations are not part of one common construct. Clearly, there is still much work ahead to build a framework which incorporates all these findings.

The combination of numerical information is thought to represent a directional SNA with implicit coding of *arithmetic functions* (rather than magnitudes or ordinal sequences as in the paragraphs above). Combining numerical quantities during mental arithmetic is believed to be influenced by elementary perceptual operations such as attentional shifts (Fischer & Knops, 2014; Knops, Thirion, Hubbard, Michel, & Dehaene, 2009). This may underlie a particular bias during addition and subtraction, called the operational momentum effect (OM). OM describes the phenomenon that participants tend to overestimate the results of addition problems while underestimating subtraction problems (McCrink, Dehaene, & Dehaene-Lambertz, 2007). During the computation of the outcome, participants are thought to activate and attend different positions on the MNL, that is, they move along the MNL. The OM is thought to reflect a certain over- and undershoot of movement, caused by attentional shifts. McCrink and Hubbard (2017, this issue) investigate whether the OM effect is modulated by the overall amount of available attentional resources. The authors found that the OM effect increased when available attentional resources were limited by dividing attention between two concurrent tasks. They conclude that the increase of OM results from a heightened use of arithmetic heuristics, such as ‘addition means larger’, which have long been known in the history of word problems as the so called semantic consistency effects (e.g., Daroczy, Wolska, Meurers, & Nuerk, 2015, for a review). In contrast, Katz, Hoesterey, and Knops (2017, this issue) found that the operational momentum in non-symbolic multiplication and division was correlated with reorienting attention in a sample of healthy adults; the higher the reorientation costs, the stronger the OM effect. These results provide further evidence for a functional association between spatial attention and approximate arithmetic, as stipulated by the attentional shifts account of OM. These conflicting results can be seen as starting point for a more strategic and joint effort to investigate how domain-general processes contribute to particular empirical phenomena. Finally, Macchi Cassia, Bulf, McCrink, and de Hevia (2017, this issue) investigated the development and emergence of OM effects, i.e., whether 4-months-old infants are subject to OM-like effects. Infants were habituated to sequences of objects that changed in one or several quantitative dimension such as physical size or numerical quantity. At test, infants were presented with sequences that consisted of items that were extrapolating the previously habituated dimensional sequence (e.g., a sequence with objects of increasing size with individual sizes being larger than during habituation) or went against this expectation (e.g., a sequence with objects of increasing size with individual sizes being smaller than during habituation). Infants exhibited longer looking times (indicating surprise due to violated expectations) only when the sequences combined violations on several dimensions simultaneously. When manipulating only physical size or numerical quantity, no change in looking time was observed. These results suggest that infants’ attention is guided towards concordant information from several dimensions within a visual

scene. Alternatively, different dimensions may act together and separately add to the built-up of expectations in time.

To sum up, the results show that numerical and spatial processes interact with each other. Yet, these spatial-numerical associations are not a unitary construct. We need to differentiate between different regimes of numerosity perception (subitizing, estimation, texture perception) that are governed by the overall number of items in a scene and their spatial layout. Further, the proposed taxonomy provides a useful framework to organize the different SNAs. Implicit directional associations between number and space need to be dissociated from explicit ones, as shown for the SNARC effect in parity and magnitude judgment tasks, respectively, and different numerical representations such as cardinality or ordinality need to be dissociated. However, not only numerical attributes are associated with space, but also numerical functions. The operational momentum effect provides an exciting test bed for investigating the interaction between numerical functions (e.g. approximate arithmetic) and spatial capacities.

To sum up this section, first, our special issue shows the need for distinctions in the associations between domain-specific number capabilities (cardinality, ordinality, functions) in their relation with the more domain-general processing of space. Second, however, this special issue also seems to suggest that the number magnitude system may be part of a more general mental magnitude system (see for example, Walsh, 2003). To integrate such seemingly diverging findings is an important task for the future. We suggest that the integration and differentiation of space-magnitude associations may depend on task, sample (age), experimental context and the involvement of other domain-general factors. To distinguish the situations and processes, in which SNAs are rather part of a general magnitude system, from those in which we need further differentiation even within the numerical domain, remains a challenge for future research and theory and model development. We hope that the current special issue helps to set the necessary constraints for this endeavor. Finally, our special issue draws attention to one of the most challenging issues in experimental psychology, namely to critically question to what extent the tasks and paradigms we deploy are indeed direct and valid measures of the cognitive processes we aim to assess.

Predicting and Improving Arithmetic

(Early) prediction of arithmetic capabilities by more basic domain-specific factors (e.g., approximate number system) or domain-general factors (e.g., working memory) has been a long-time dream of cognitive and educational researchers and practitioners in arithmetic research. Identifying such building blocks and cornerstones of arithmetic development and functioning would have important consequences for education and intervention. Education – even much before formal schooling – could focus on mastering elementary building blocks of arithmetic to improve arithmetic performance and learning at large. Moreover, diagnostics could identify children who have trouble mastering the basic building blocks of arithmetic, before formal schooling, and targeted interventions may then help to improve these building blocks of later arithmetic development and the long-term outcome of those children. To identify such building blocks, prediction and intervention studies are essential – most researchers seem to agree that both domain-specific and domain-general factors predict (later) arithmetic performance. However, there is still no clear consensus on which of those factors are fundamental to arithmetic performance and arithmetical development. As we will see, about half of our special issue is devoted to the question of predicting and improving arithmetic performance and development.

To begin with, the capacity to judge the ordinal relation between objects has recently been suggested to be an important stepping stone for arithmetic performance. The paper by Vogel et al. (2017, this issue) focuses on the relationship between serial order and arithmetic in adults. They measured adults' arithmetic fluency and their ability to judge whether Arabic digits, dot patterns and letters are ordered correctly by magnitude at two time points. Adults' reaction times on the symbolic order judgment task (Arabic digits) was an independent predictor of arithmetic fluency over and above their reaction times on the non-symbolic judgment task (dot patterns) and the letter order task. In line with findings from Lyons and Beilock (2011) and Lyons, Price, Vaessen, Blomert, and Ansari (2014), this highlights that the understanding of the ordinal relationship between Arabic digits might be foundational to arithmetic performance. However, alternatively, adults' efficiency in dealing with Arabic digits might be driving the relationship between both tasks, the serial order judgment of Arabic digits and the arithmetic fluency task (Castronovo & Göbel, 2012). Indeed, in primary school children familiarity with the Arabic digit symbol system is a significant predictor of arithmetic growth (Göbel, Watson, Lervåg, & Hulme, 2014). Future studies will need to include measures of both serial order and efficiency in the processing of Arabic digits in adults and/or a measure of familiarity with the Arabic digit system in children in order to disentangle whether ordinal understanding of Arabic digits or the ease of processing Arabic digits is a stronger predictor of arithmetical development and performance.

Moving on from factors specific to the numerical domain, such as Arabic digit order, to domain-relevant skills, Cornu, Hornung, Schiltz, and Martin (2017, this issue) investigated the relative importance of spatial skills in kindergartners as longitudinal predictors of arithmetic and number line estimation. They differentiate between skills in spatial orientation, in spatial visualisation and in visuo-motor integration. They found that number line estimation in 5 year-old children was significantly predicted by their performance on the spatial orientation and visuo-motor integration tasks four months earlier. In addition, Arabic number knowledge, spatial orientation and visuo-motor integration were significant predictors of arithmetic performance four months later. The relationship between spatial orientation and arithmetic was partially mediated through children's number line estimation. This study provides evidence of the usefulness of a more fine-grained approach showing that some but not all spatial skills are important for early arithmetic development and that the importance of different spatial skills even varies between different numerical tasks. It also highlights that the theory of spatial influences on arithmetic performance and development is currently still underdeveloped. To the best of our knowledge, there is no systematic taxonomy specifying which spatial representations and processes must be distinguished, because they differentially influence arithmetic performance and/or development in general or even differentially influence different numerical and arithmetic capabilities (but see Fischer, 2012, for a more general conceptual framework).

One first step towards the development of such a taxonomy is taken by Crollen, Collignon, and Noël (2017, this issue) by reviewing findings from several atypical populations with deficits in the visual or visuo-spatial domain: early blind adults, adults with hemi-spatial neglect, children with low visuo-spatial skills, children with non-verbal learning disorder and children with William's syndrome. In their review, they ask the question whether and if so, how, the specific deficits observed in these populations affect the number domain. Results vary largely between these different atypical populations from seemingly no effects of early blindness on the development of spatial-numerical associations to severe deficits or delays on a range of numerical tasks for children with William's syndrome. Future work now needs to develop a taxonomy leading to an overarching framework with clear predictions about the relationships between specific spatial and numerical skills that can also account for the

deficits in numerical representations and arithmetic performance observed in these populations with abnormal (visuo-)spatial representations.

Several studies in this special issue took the laudable approach to investigate the relative contribution of domain-specific and domain-general factors towards arithmetic and mathematical performance within the same study. Purpura, Day, Napoli, and Hart (2017, this issue) were interested in predicting later mathematical performance in pre-schoolers, particularly for low mathematical performance. They tested children on a large battery of domain-specific and domain-general tasks including early numeracy, ANS, language and literacy tests, mathematical language, executive functions and processing speed. In the test of mathematical language children were assessed on the understanding of comparative (e.g. more, less) and spatial (e.g. near, far) language. For younger children poor performance in mathematical language, print knowledge and response inhibition was indicative of poor mathematical performance about five months later. For older children, it was poorer performance on mathematics, mathematical language and definitional vocabulary. For young children domain-general processes such as language and executive function actually allowed more accurate classification for them than their performance on number-specific tasks. A clear outcome of this study is to highlight mathematical language as a currently understudied, yet recently emerging potential candidate for early training studies.

Gilmore, Keeble, Richardson, and Cragg (2017, this issue) investigated the effect of three domain-general skills, procedural skill, conceptual understanding and working memory, on mathematical achievement within the same study. As in previous studies, they found that in 5-6 year old children all three domain-general skills are independently associated with mathematical performance. However, more importantly, they investigated the relationship between those three domain-general skills and found that they interact: the impact of better procedural skills on mathematical performance was higher for children who also had better conceptual understanding and higher working memory capacity. In addition, their study showed large inter-individual differences in children's skill profiles on these three domain-general skills: children with similar scores in a mathematical achievement test can show very different strengths and weaknesses in procedural skill, conceptual understanding and working memory. These findings highlight how important it is to examine the cognitive profile of children in more details in order to identify the type of support most likely to improve each individual child's mathematical performance.

Two further papers in this special issue investigated the role of working memory for mathematical performance in the context of other domain-general factors. Kroesbergen and Schoevers (2017, this issue) focused on the contribution of creativity and working memory to mathematical performance in 8-10 year-old children. Overall, working memory, in particular verbal working memory, was significantly associated with mathematical performance. Creativity was significantly related to performance on a standard mathematics test and a mathematical creativity test. Children's performance on the creativity test was predictive of mathematical performance even when their working memory and number sense performance was taken into account. When they split their sample of children into three groups by mathematical performance into 1) children with Mathematical Learning Difficulties, 2) typically developing children and 3) mathematically gifted children, creativity discriminated between typically developing children and mathematically gifted children with the mathematically gifted children achieving a significantly higher creativity score. Interestingly, children's visuo-spatial WM discriminated between all three groups of children: the children with Mathematical Learning Difficulties showed the lowest performance in spatial working memory, followed by the typically developing

children. Mathematically gifted children showed the highest spatial working memory scores. In sum, this study introduces a new domain-general factor predicting mathematical performance: creativity. It also confirms the importance of working memory for mathematical performance in children.

In contrast, in Nemati, Schmid, Soltanlou, Krimly, Nuerk, and Gawrilow's study (2017, this issue) on adults, working memory was no longer significantly associated with their mathematical performance once other domain-general skills were included. Planning capacities (as measured by the performance in the Tower of London task) and self-control (as measured by self-report) predicted multiplication performance in undergraduate students and working memory was no longer a significant predictor anymore, when executive functions (planning) were considered. This finding is in line with the assumption that arithmetic fact retrieval in adults primarily relies on recall from long-term memory, rather than the application of arithmetic procedures (e.g., Campbell, 1995)

Finally, Ashkenazi and Silverman (2017, this issue), in large-scale sample (N = 1322) of college students, investigated the influence of three further domain-general variables on mathematical capabilities: perception speed, attention and reading variables. Employing structural equation modeling they observed effects of perception speed and a modest effect of reading on mathematics performance. Sustained attention had some impact on selected mathematical skills (arithmetic fact retrieval and procedural knowledge), while selective attention (assessed by the attention network test) had no effect on mathematics. The authors concluded that multiple domain-general skills have an influence on mathematic performance. However, their data also point to the conclusion that not every domain-general variable affects every aspect of numerical and arithmetic capabilities in the same way.

As we laid out in the introduction, prediction studies are essential to identify targets for early education, instruction, and intervention. However, we also believe, there are some serious shortcomings currently in the literature as whole. First, in our special issue alone over twenty different predictors were tested and many more potential predictors are out there. In general, each study uses its own set of predictors. Consequently, different studies reveal different sets of predictors, which are relevant for good (later) arithmetic performance or arithmetical development. However, it is important to note that the results of a study do not only depend on the predictors included, but also on the predictors not included. For instance, Nemati et al. (2017, this issue) found correlations of working memory with arithmetic performance (albeit in adults). However, working memory failed to be a predictor, when one executive function measure (namely planning) was included. Had Nemati and colleagues not included planning, they would have published another study, which had suggested that working memory itself (not as a possible part of planning) is the most relevant predictor.

This leads us to the second point: the power of prediction studies. Large-scale longitudinal studies are hard to conduct and require a lot of effort. This is even more so the case for the age range for which finding predictors is arguably most important and of most practical relevance: from kindergarten to school. Therefore, either the sample size is often quite small for the number of predictors or the set of predictors is very limited. Both solutions can lead to misleading results and this might be one of the reasons why different prediction studies have often very different outcomes. What is needed, is a large-scale multi-center prediction study, which incorporates a large number of children and all relevant domain-specific and domain-general predictors so far found in the literature.

Our final point is that the outcome measure is often either one of many available standardized mathematics test or a curriculum-based test of mathematics. Those tests are often 'umbrella tests', i.e. measuring a large range of numerical, arithmetical and mathematical abilities without an option to distinguish between them. Consequently, those tests are frequently used without a model of its underlying representations and processes. In numerical cognition, it is virtually undisputed nowadays that different (neuro-cognitive) representations and networks are supporting different numerical processes and operations (e.g., Dehaene et al., 2003; Klein et al., 2016). Just using one big melting pot variable most certainly leads to missing important specific (longitudinal) relationships. For instance, using various, clearly specified outcome variables, Moeller, Pixner, Zuber, Kaufmann, and Nuerk (2011) found that different predictor variables predict different outcome effects. In summary, a better differentiation of both the predictor and the outcome variable/s as well as more clearly defined models about the proposed relationship between predictors and outcome variable/s are needed. It may also help to reconcile apparently different results, because such differences might not rely only on the predictors included in a study, but also on the characteristics of the outcome variable/s.

Intervention Studies

Compared to prediction studies, intervention studies, in addition to their obvious practical implications, have an important theoretical advantage. Prediction studies are by its very nature correlational (when they are longitudinal over different time points) and all variables assessed are dependent variables. In contrast, in intervention studies the variable of interest, e.g. type of training, is manipulated as independent variable and therefore findings from intervention studies allow (cautious) causal rather than only correlational conclusions. Like in every other study, this does not preclude the influence of confounded or mediating variables. However, if a particular intervention leads to a training effect, this allows the conclusion that either the variable of interest (the training) or a variable confounded with it were instrumental for the intervention outcome, e.g., for improvement in arithmetic performance.

In this special issue, two papers evaluated the effect of working memory training on numerical and arithmetic performance. Honoré and Noël (2017, this issue) trained 5-6 year old preschoolers on a visuo-spatial working memory training program (Cogmed) for five weeks. There was a small effect directly after training on Arabic number comparison and visuo-spatial working memory, but there was no effect of training on verbal working memory, counting comparison of collections and addition. The effect of training on Arabic number comparison and visuo-spatial working memory was not sustained ten weeks after the training. Ramani, Jaeggi, Daubert, and Buschkuehl (2017, this issue) compared the effectiveness of two tablet-based interventions to a no intervention control group in 6 year-old kindergartners from low income backgrounds: a domain-general (visuo-spatial working memory) intervention and a domain-specific (number board game) intervention. They found no effect of either training on children's arithmetic performance and only limited improvement of working memory performance for one of the three working memory measures for both trainings. But both interventions improved children's numerical magnitude knowledge, as assessed by the number line estimation task, with a significantly larger effect for the number-specific training. These results suggest that domain-specific training is more effective. Furthermore, they demonstrate the practicability of easy and comparably cheap, tablet-based interventions for improving kindergarten children's numerical understanding. This should encourage educators and teachers to harness the benefit of these new techniques as they become more and more accessible. However, in both intervention studies the effects of training were either non-existent, moderate or limited to

specific outcome variables. This suggests that correlations observed in prediction studies do not easily translate into learning success in intervention studies.

Summary and Conclusions

The papers in this special issue show that domain-general as well as domain-specific abilities influence numerical and arithmetic performance virtually at all levels. This special issue thus makes it very clear that for the field of numerical cognition a sole focus on one or several domain-specific factors, like the approximate number system or spatial-numerical associations, is not sufficient. Vice versa, in most studies that included domain-general and domain-specific variables, domain-specific numerical variables predicted arithmetic performance above and beyond domain-general variables. Therefore, a sole focus on domain-general aspects, such as, for example, working memory, to explain, predict and foster arithmetic learning is also not sufficient. In the two intervention studies of this special issue, effects of domain-general interventions were weak or even not existent. Therefore, by acknowledging the importance of domain-general factors for arithmetic we do most certainly not advocate a restriction on domain-general factors. Rather, we are convinced that to understand numerical and arithmetic performance, development and learning, the contribution of both domain-general and domain-specific factors must be considered. However, these contributions may not simply be linearly additive or even independent; rather their interplay and their interactions must be studied more thoroughly both concurrently and longitudinally in future research. We believe that only then the full picture of arithmetic performance, development and learning can be understood.

What is still missing in our view are thoroughly developed models that specify how and to which extent domain-general and domain-specific factors contribute to numerical and arithmetical performance, development and learning and how those factors interact. As shown in Figure 1 and Table 1a/1b, based on the contributions to this special issue we can conclude that both domain-general and domain-specific factors contribute to numerical cognition. But the how, why and when of that contribution still needs to be better understood. We hope that this special issue may be helpful to readers in constraining future theory and model building about the interplay of domain-specific and domain-general factors.

Funding

HCN was supported by the Deutsche Forschungsgemeinschaft (DFG: NU 265/3-1).

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgments

The authors have no support to report.

References

Alonso-Diaz, S., Gaffin-Cahn, E., Mahon, B. Z., & Cantlon, J. F. (2017). What is in a reach? Domain-general spatial modulation of motor responses by number representations. *Journal of Numerical Cognition, 3*(2), 112-132. doi:10.5964/jnc.v3i2.28

Anobile, G., Cicchini, G. M., Pomè, A., & Burr, D. C. (2017). Connecting visual objects reduces perceived numerosity and density for sparse but not dense patterns. *Journal of Numerical Cognition, 3*(2), 133-146. doi:10.5964/jnc.v3i2.38

Ashkenazi, S., & Silverman, S. (2017). Multiple skills underlie mathematics performance: A comprehensive perspective. *Journal of Numerical Cognition, 3*(2), 496-515. doi:10.5964/jnc.v3i2.64

Baillargeon, R., & Carey, S. (2012). Core cognition and beyond: The acquisition of physical and numerical knowledge. In S. Pauen (Ed.), *Early childhood development and later outcome* (pp. 33-65). Cambridge, United Kingdom: Cambridge University Press.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London: Series B. Biological Sciences, 364*(1525), 1831-1840. doi:10.1098/rstb.2009.0028

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition, 1*, 121-164.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13*(2), 83-91. doi:10.1016/j.tics.2008.11.007

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One, 7*(4), Article e33832. doi:10.1371/journal.pone.0033832

Cavdaroglu, S., & Knops, A. (2016). Mental subtraction and multiplication recruit both phonological and visuospatial resources: Evidence from a symmetric dual-task design. *Psychological Research, 80*, 608-624. doi:10.1007/s00426-015-0667-8

Cipora, K., Schroeder, P. A., Soltanlou, M., & Nuerk, H. C. (in press). More space, better math: Is space a powerful tool or a cornerstone for understanding arithmetic? In K. S. Mix & M. T. Battista (Eds.), *Visualizing mathematics: The role of spatial reasoning in mathematical thought.* Cham, Switzerland: Springer.

Cornu, V., Hornung, C., Schiltz, C., & Martin, R. (2017). How do different aspects of spatial skills relate to early arithmetic and number line estimation? *Journal of Numerical Cognition, 3*(2), 309-343. doi:10.5964/jnc.v3i2.36

Crollen, V., Collignon, O., & Noël, M.-P. (2017). Visuo-spatial processes as a domain-general factor impacting numerical development in atypical populations. *Journal of Numerical Cognition, 3*(2), 344-364. doi:10.5964/jnc.v3i2.44

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H.-C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology, 6*, Article 348. doi:10.3389/fpsyg.2015.00348

Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16*(1), 16-36. doi:10.1111/1468-0017.00154

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*(3), 371-396. doi:10.1037/0096-3445.122.3.371

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384-398.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*(3-6), 487-506. doi:10.1080/02643290244000239

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences of the United States of America, 111*(13), 4809-4813. doi:10.1073/pnas.1323628111

Elman, J. L., Bates, E. A., Johnson, M. J., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1998). *Rethinking innateness: A connectionist perspective on development* (Vol. 10). Cambridge, MA, USA: MIT Press.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307-314.

Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cognitive Processing, 13*(Suppl. 1), 161-164. doi:10.1007/s10339-012-0477-5

Fischer, M. H., & Knops, A. (2014). Attentional cueing in numerical cognition. *Frontiers in Psychology, 5*, Article 1381. doi:10.3389/fpsyg.2014.01381

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA, USA: MIT press.

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28*, 774-782. doi:10.3758/BF03198412

Georges, C., Hoffmann, D., & Schiltz, C. (2017). How and why do number-space associations co-vary in implicit and explicit magnitude processing tasks? *Journal of Numerical Cognition, 3*(2), 182-211. doi:10.5964/jnc.v3i2.46

Gevers, W., Cohen Kadosh, R., & Gebuis, T. (2016). Sensory integration theory: An alternative to the approximate number system. In A. Henik (Ed.), *Continuous issues in numerical cognition* (pp. 405-418). London, United Kingdom: Academic Press.

Gilmore, C., Keeble, S., Richardson, S., & Cragg, L. (2017). The interaction of procedural skill, conceptual understanding and working memory in early mathematics achievement. *Journal of Numerical Cognition, 3*(2), 400-416. doi:10.5964/jnc.v3i2.51

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science, 25*(3), 789-798. doi:10.1177/0956797613516471

Honoré, N., & Noël, M.-P. (2017). Can working memory training improve preschoolers' numerical abilities? *Journal of Numerical Cognition, 3*(2), 516-539. doi:10.5964/jnc.v3i2.54

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*(6), 435-448. doi:10.1038/nrn1684

Huber, S., Bloechle, J., Dackermann, T., Scholl, A., Sassenberg, K., & Moeller, K. (2017). Magnitude estimation is influenced by social power. *Journal of Numerical Cognition, 3*(2), 147-163. doi:10.5964/jnc.v3i2.52

Imbo, I., Vandierendonck, A., & De Rammelaere, S. (2007). The role of working memory in the carry operation of mental arithmetic: Number and value of the carry. *Quarterly Journal of Experimental Psychology, 60*, 708-731. doi:10.1080/17470210600762447

Imbo, I., Vandierendonck, A., & Vergauwe, E. (2007). The role of working memory in carrying and borrowing. *Psychological Research, 71*, 467-483. doi:10.1007/s00426-006-0044-8

Karmiloff-Smith, A. (2015). An alternative to domain-general or domain-specific frameworks for theorizing about human evolution and ontogenesis. *AIMS Neuroscience, 2*(2), 91-104.

Katz, C., Hoesterey, H., & Knops, A. (2017). A role for attentional reorienting during approximate multiplication and division. *Journal of Numerical Cognition, 3*(2), 246-269. doi:10.5964/jnc.v3i2.62

Klein, E., Suchan, J., Moeller, K., Karnath, H. O., Knops, A., Wood, G., . . . Willmes, K. (2016). Considering structural connectivity in the triple code model of numerical cognition: Differential connectivity for magnitude processing and arithmetic facts. *Brain Structure & Function, 221*(2), 979-995. doi:10.1007/s00429-014-0951-1

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science, 324*(5934), 1583-1585. doi:10.1126/science.1171599

Kroesbergen, E. H., & Schoevers, E. M. (2017). Creativity as predictor of mathematical abilities in fourth graders in addition to number sense and working memory. *Journal of Numerical Cognition, 3*(2), 417-440. doi:10.5964/jnc.v3i2.63

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*, 256-261. doi:10.1016/j.cognition.2011.07.009

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in Grades 1-6. *Developmental Science, 17*, 714-726. doi:10.1111/desc.12152

Macchi Cassia, V., Bulf, H., McCrink, K., & de Hevia, M. D. (2017). Operational momentum during ordering operations for size and number in 4-month-old infants. *Journal of Numerical Cognition, 3*(2), 270-287. doi:10.5964/jnc.v3i2.67

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics, 69*(8), 1324-1333. doi:10.3758/BF03192949

McCrink, K., & Hubbard, T. (2017). Dividing attention increases operational momentum. *Journal of Numerical Cognition, 3*(2), 230-245. doi:10.5964/jnc.v3i2.34

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H. C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities, 32*(5), 1837-1851. doi:10.1016/j.ridd.2011.03.012

Morgan, M. J., Raphael, S., Tibber, M. S., & Dakin, S. C. (2014). A texture-processing model of the 'visual sense of number'. *Proceedings of the Royal Society of London B: Biological Sciences, 281*(1790), Article 20141137. doi:10.1098/rspb.2014.1137

Nemati, P., Schmid, J., Soltanlou, M., Krimly, J.-T., Nuerk, H.-C., & Gawrilow, C. (2017). Planning and self-control, but not working memory, directly predict multiplication performance in adults. *Journal of Numerical Cognition, 3*(2), 441-467. doi:10.5964/jnc.v3i2.61

Núñez, R. E. (2017). Is there really an evolved capacity for number? *Trends in Cognitive Sciences, 21*(6), 409-424. doi:10.1016/j.tics.2017.03.005

Patro, K., & Nuerk, H.-C. (2017). Limitations of trans-species inferences: The case of spatial-numerical associations in chicks and humans. *Cognitive Science, 41*, 2267-2274. doi:10.1111/cogs.12432

Patro, K., Nuerk, H.-C., & Cress, U. (2016). The mental number line in the preliterate brain: The role of early directional experiences. *Child Development Perspectives, 10*, 172-177. doi:10.1111/cdep.12179

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., . . . Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*(1), 33-41. doi:10.1016/j.cognition.2010.03.012

Purpura, D. J., Day, E., Napoli, A. R., & Hart, S. A. (2017). Identifying domain-general and domain-specific predictors of low mathematics performance: A classification and regression tree analysis. *Journal of Numerical Cognition, 3*(2), 365-399. doi:10.5964/jnc.v3i2.53

Ramani, G. B., Jaeggi, S. M., Daubert, E. N., & Buschkuehl, M. (2017). Domain-specific and domain-general training to improve kindergarten children's mathematics. *Journal of Numerical Cognition, 3*(2), 468-495. doi:10.5964/jnc.v3i2.31

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science, 347*(6221), 534-536. doi:10.1126/science.aaa1379

Schleger, F., Landerl, K., Muenssinger, J., Draganova, R., Reinl, M., Kiefer-Schmidt, I., . . . Preissl, H. (2014). Magnetoencephalographic signatures of numerosity discrimination in fetuses and neonates. *Developmental Neuropsychology, 39*(4), 316-329. doi:10.1080/87565641.2014.914212

Schroeder, P. A., Nuerk, H.-C., & Plewnia, C. (2017a). Switching between multiple codes of SNARC-like associations: Two conceptual replication attempts with anodal tDCS in sham-controlled cross-over design. *Frontiers in Neuroscience, 11*, Article 654. doi:10.3389/fnins.2017.00654

Schroeder, P. A., Nuerk, H.-C., & Plewnia, C. (2017b). Space in numerical and ordinal information: A common construct? *Journal of Numerical Cognition, 3*(2), 164-181. doi:10.5964/jnc.v3i2.40

Seitz, K., & Schumann-Hengsteler, R. (2002). Phonological loop and central executive processes in mental addition and multiplication. *Psychologische Beiträge, 44*, 275-302.

Song, J. H., & Nakayama, K. (2008). Numeric comparison in a visually-guided manual reaching task. *Cognition, 106*(2), 994-1003. doi:10.1016/j.cognition.2007.03.014

Van Dijck, J.-P., Ginsburg, V., Girelli, L., & Gevers, W. (2015). Linking numbers to space: From the mental number line towards a hybrid account. In R. Cohen-Kadosh & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 89-105). Oxford, United Kingdom: Oxford Unversity Press.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience, 16*(9), 1493-1504. doi:10.1162/0898929042568497

Vogel, S. E., Haigh, T., Sommerauer, G., Spindler, M., Brumner, C., Lyons, I. M., & Grabner, R. (2017). Processing the order of symbolic numbers: A reliable and unique predictor of arithmetic fluency. *Journal of Numerical Cognition, 3*(2), 288-308. doi:10.5964/jnc.v3i2.55

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*(11), 483-488. doi:10.1016/j.tics.2003.09.002

Journal of Numerical Cognition
jnc.psychopen.eu | 2363-8761

Research Reports

A Role for Attentional Reorienting During Approximate Multiplication and Division

Curren Katz[a], Hannes Hoesterey[a], André Knops*[a]

[a] Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany.

Abstract

When asked to estimate the outcome of arithmetic problems, participants overestimate for addition problems and underestimate for subtraction problems, both in symbolic and non-symbolic format. This bias is referred to as operational momentum effect (OM). The attentional shifts account holds that during computation of the outcome participants are propelled too far along a spatial number representation. OM was observed in non-symbolic multiplication and division while being absent in symbolic multiplication and division. Here, we investigate whether (a) the absence of the OM in symbolic multiplication and division was due to the presentation of the correct outcome amongst the response alternatives, putatively triggering verbally mediated fact retrieval, and whether (b) OM is correlated with attentional parameters, as stipulated by the attentional account. Participants were presented with symbolic and non-symbolic multiplication and division problems. Among seven incorrect response alternatives participants selected the most plausible result. Participants were also presented with a Posner task, with valid (70%), invalid (15%) and neutral (15%) cues pointing to the position at which a subsequent target would appear. While no OM was observed in symbolic format, non-symbolic problems were subject to OM. The non-symbolic OM was positively correlated with reorienting after invalid cues. These results provide further evidence for a functional association between spatial attention and approximate arithmetic, as stipulated by the attentional shifts account of OM. They also suggest that the cognitive processes underlying multiplication and division are less prone to spatial biases compared to addition and subtraction, further underlining the involvement of differential cognitive processes.

Keywords: operational momentum, approximate calculation, spatial attention, mental number line, reorienting, mental arithmetic

Journal of Numerical Cognition, 2017, Vol. 3(2), 246–269, doi:10.5964/jnc.v3i2.62

Received: 2016-06-13. Accepted: 2017-11-17. Published (VoR): 2017-12-22.

Handling Editors: Silke Goebel, University of York, York, United Kingdom; Hans-Christoph Nuerk, Universität Tübingen, Tübingen, Germany

*Corresponding author at: Unter den Linden 6, 10099 Berlin, Germany. E-mail: andre.knops@hu-berlin.de

Numerical cognition has long been thought to represent a prime example of an abstract propositional symbolic system with no obvious reference to the outer world. For example, number names such as ‘one’, ‘two’ and so on do not bear any obvious association to the referenced cardinal meaning. Recent evidence, however, implies that numbers and mental arithmetic bear numerous associations with space and time (Fischer & Knops, 2014; Hubbard, Piazza, Pinel, & Dehaene, 2005). Beyond using space to represent numerical magnitude along a spatially oriented mental number line (MNL), it has been hypothesized that humans re-use their spatial prowess and the underlying neural resources to master numerical operations (Dehaene & Cohen, 2007; Knops, Thirion, Hubbard, Michel, & Dehaene, 2009). The cortical mechanisms in parietal cortex that serve for the transformation of spatial coordinates in different referent frames (Beck, Latham, & Pouget, 2011; Pouget, Deneve, & Duhamel, 2002) may play a pivotal role in this context. In particular, neural populations in posterior

superior parietal cortex encode object location with respect to different body parts and references. For example, during the planning of eye movements these circuits compute the head-centred position of a given object by adding the activity of neural populations that encode eye-centred coordinates and eye position (Beck et al., 2011; Pouget et al., 2002). Hence, the cortical algorithms for executing simple additions are present in parietal cortex. Basic mathematical operations might co-opt these circuits by providing numerical input instead of coordinate information (Dehaene & Cohen, 2007; Hubbard et al., 2005; Knops, Thirion, et al., 2009).

This process is not bias-free. When approximating the outcome of simple addition or subtraction problems humans are likely to provide a biased response that deviates systematically from the correct outcome. The results of addition problems tend to be overestimated while the results for subtraction problems are underestimated. This cognitive bias is referred to as operational momentum (OM; McCrink, Dehaene, & Dehaene-Lambertz, 2007) due to its resemblance to a systematic bias in the estimation of the point in space where a moving object disappeared (Hubbard, 2005, 2015). OM has initially been observed with non-symbolic operations where participants estimated the number of visual objects (dots) that would result from adding two sets of dots or subtracting one set from the other (McCrink et al., 2007). However, OM has also been found in symbolic notation (Knops, Dehaene, Berteletti, & Zorzi, 2014; Knops, Viarouge, & Dehaene, 2009) which has been interpreted as evidence for a common underlying mechanism. OM has also been observed with paradigms that require translating the cognitively generated numerical estimate into a position on a labeled line (Pinhas & Fischer, 2008), or actively producing the outcome via a dot generating manual device (Lindemann & Tira, 2011). In the labeled line task, participants indicated the position of the outcome of visually presented addition and subtraction problems on a line that was labeled with zero on the left and 10 on the right. Crucially, participants misplaced addition outcomes to the right and subtraction outcomes to the left, compared to baseline that required indicating the numbers' positions without preceding arithmetic operation. When actively producing dot patterns corresponding to the outcome of multi-digit problems, participants produced relatively larger estimates for addition problems as compared to subtraction problems with identical correct outcome (Lindemann & Tira, 2011). Together, this implies that mental arithmetic is subject to systematic biases that may have their origin in the application of spatial coordinate transformation mechanisms to numerical quantity information that can be conceived of as positions on the MNL.

More evidence for spatial contributions to mental arithmetic comes from a recent study reporting systematic interference between arm or eye movements and mental arithmetic (Wiemers, Bekkering, & Lindemann, 2014). Addition performance was impaired when participants moved their arms or eyes downward while subtraction was affected with upward movements of the arm or eye. In horizontal plane only arm but not eye movements to the left interfered with addition while arm movements to the right interfered with subtraction. During simple addition and subtraction tasks, Margethis and colleagues found systematic deviations of manual mouse pointer trajectories from the ideal path such that addition and subtraction trajectories deviated to the right and left from the ideal paths, respectively (Marghetis, Núñez, & Bergen, 2014). Interestingly, this effect was observed over and above confounding the response position with numerical magnitude (Pinhas & Fischer, 2008; Pinhas, Shaki, & Fischer, 2014) because the correct outcomes were presented both to the upper left or right relative to starting position.

Despite the growing number of studies demonstrating OM in different settings, the underlying mechanisms of the OM effect are currently debated. Three major hypotheses vie with one another. First, it has been suggested that the OM effect reflects the outcome of a simple heuristic that would associate different arithmetic operations

with expectations concerning the numerical relationship between the outcome and the operands (McCrink et al., 2007). For addition, this heuristic would predict that the result should be larger than either of both operands. For subtraction, the heuristic would predict outcomes that are smaller than the first operand. The same heuristic would hold for other arithmetic operations such as multiplication and subtraction. In particular, the observed OM in 9-month-olds supports this account because it is unlikely that at that age infants have developed a spatial representation of numerical magnitude (McCrink & Wynn, 2009). A second hypothesis assumes that OM results from the flawed logarithmic compression and decompression into a linear metric during the arithmetic process (Chen & Verguts, 2012). The mental number line is assumed to be logarithmically compressed (Dehaene, 2001; Nieder & Dehaene, 2009). The compression and decompression processes needed to transduce between linear and logarithmic scales may be flawed. In the extreme variant of the compression hypothesis, addition and subtraction operate on compressed values, that is, the decompression fails entirely. This would result in massive over- and underestimations since the sum of the logs is the log of a product (i.e. $\log(a) + \log(b) = \log(a \times b)$). An equivalent underestimation would result for subtraction since the difference of the logs is the log of a quotient (i.e. $\log(a) - \log(b) = \log(a/b)$).

Finally, spatial accounts have been proposed to account for OM. According to the spatial competition hypothesis (Pinhas & Fischer, 2008; Pinhas, Shaki, & Fischer, 2015), OM is the result of the "competing spatial biases invoked by the operands, the operation sign, and the result of an arithmetic problem" (p. 997; Pinhas et al., 2015). Operands and results activate their respective positions on the mental number line, that "compete for responses" (p. 413; Pinhas & Fischer, 2008). For subtractions, for example, the result of a given problem may be located between the operands (e.g. 7 - 2 = 5) or to the left of both operands (e.g. 7 - 4 = 3). Compared with problems involving zero as a second operand, these competing biases mitigate the observed bias towards the result which is more pronounced when the second operand does not additionally compete for responses. Consequently, Pinhas and Fischer (2008) observed the largest OM bias with zero problems. Recently, these authors observed an inverse OM when reversing the line labels such that the right end of the line was labeled with 0 and the left end with 10 (Pinhas et al., 2015), providing support for their spatial competition bias and underlining the observation that number-to-space mappings are highly flexible (Bächtold, Baumuller, & Brugger, 1998). According to a second spatial account, the OM reflects systematic biases from the deployment of the coordinate transformation system in parietal cortex that also mediates attentional shifts in space (Knops et al., 2014; Knops, Thirion, et al., 2009; Knops, Viarouge, et al., 2009; Knops, Zitzmann, & McCrink, 2013). According to this approach, approximate mental arithmetic is mediated by a dynamic interaction between positional codes on the MNL and an attentional system that shifts the spatial focus to the left or right. At the neural level this may be instantiated in the functional interactions between areas along the intraparietal sulcus and posterior, superior parietal areas. The idea is that a parietal circuit that has been proposed to combine retinal and eye position information via vector addition in order to compute positions in space may be recycled to implement mental arithmetic. The resulting positional information can be used to guide eye or hand movements and has been proposed to be the base for shifts of spatial attention. This places mental arithmetic in the realm of dynamic updating processes of spatial coordinates in parietal cortex and stipulates that the efficiency of this system is linked with arithmetic performance. Due to the approximate nature of this process the shifts may 'overshoot', leading to over- and underestimation in addition and subtraction, respectively. Interestingly, the latter approach suggests a functional coupling between eye movements and arithmetic. A recent study provided confirmatory evidence for this notion (Klein, Huber, Nuerk, & Moeller, 2014). Participants' eye movements after the first saccade were observed to move to the right during addition problems and to the

left in subtraction problems when asked to indicate the location of the result on a labeled line (but see Hartmann, Mast, & Fischer, 2015; Klein et al., 2014). A second implication from this account is that OM is not limited to addition and subtraction but generalizes to basic arithmetic transformations with natural numbers such as multiplication and division as long as they require quantity manipulations that are associated with attentional shifts along the MNL. Operations that lead to larger outcomes would be associated with rightward shifts of attention, while the opposite should hold for operations that lead to smaller outcomes. Katz and Knops (2014) investigated the OM effect in the context of multiplication and division. However, as opposed to simple addition and subtraction problems, finding the solution of simple symbolic multiplication problems is often conceptualized as verbally mediated recall from long term memory (e.g. Campbell & Xue, 2001), mixed with idiosyncratic short cuts (e.g. retrieving 9 × 7 = 63 by subtracting 7 from 70). Nevertheless, approximate estimates of the outcome might involve spatial transformations, protecting us against accepting grossly wrong solutions. For example, knowing that 32 × 8 must be a three-digit number helps excluding 40 as a possible outcome. Despite this possible role of spatial transformations in multiplication and division, neither for the standard set of multiplication tables (i.e. between 1 × 1 = 1 and 9 × 9 = 81), nor for two-digit × one-digit problems (e.g. 14 × 3 = 42) an OM effect was observed (Katz & Knops, 2014). However, for the corresponding non-symbolic problems where the quantities were presented as dot patterns a regular OM effect was observed. One crucial difference between the procedure adopted by Katz and Knops (2014) and previous studies (Knops, Viarouge, et al., 2009) was that the correct solution was presented as one of five response alternatives, potentially encouraging the engagement in exact calculation and direct retrieval rather than approximating the outcome. This may have reduced the opportunity to detect any systematic biases due to attentional shifts that accompany approximate quantity manipulations.

The aim of the current study was two-fold. First, we aimed at testing the presence of OM in multiplication and division by eliminating the presence of the correct outcome amongst the response alternatives. By encouraging participants to approximate even in the symbolic notation we aimed at increasing sensitivity to detect any systematic biases during multiplication and division with Arabic digits. Second, engaging participants in both an OM task and a Posner paradigm allowed us to test whether potential OM biases actually correlate with attentional measures. According to the above theoretical accounts of the OM we can break this question down into four aspects. Do attentional parameters correlate with (a) a heuristic according to which multiplication leads to larger outcomes and division to smaller outcomes, (b) flawed decompression, (c) competing spatial biases by the operands, the results or the outcome, or (d) attentional shifts along the mental number line? According to the heuristics account and the compression-decompression approach, no correlation with attentional measures would be predicted. Among the spatial accounts, only the attentional explanation predicts a correlation between attentional parameters and OM effect. No such correlation is predicted by the competing spatial biases account.

Methods

Participants

Participants (n = 18; 19-74 years-old, M = 36.3, SD = 19; 14 female, 4 male) were recruited from the general population (both student and non-student) using a departmental database. Written informed consent was

obtained from all participants prior to participation. Students received course credit; no other reimbursement was given.

Procedure and Materials

The study consisted of two experiments; a calculation task involving symbolic and non-symbolic multiplication and division problems, and a variant of the Posner task to test different aspects of visuo-spatial attention (orienting/selection and reorienting/executive attention).

Calculation Task

The calculation task was created and presented using OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). Participants were given written instructions and then performed 24 (12 symbolic, 12 non-symbolic; intermixed) practice calculation trials. Symbolic (Arabic digits) and non-symbolic (dot-arrays) multiplication and division problems were presented horizontally (separated by 'x' or '÷') to reduce working memory demands, for 2000ms, followed by seven response choices presented in a circle until a response was made or for a maximum of 6000ms (Figure 1). Compared to previous studies (Katz & Knops, 2014), we reduced presentation time of the operands to minimize propensity of adopting a counting strategy. Responses were made with a mouse-click on the chosen value or dot-array.

Non-symbolic stimuli were created using MATLAB and the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). To prevent that participants relied on non-numeric stimulus features such as density, occupied area or individual dot size we de-correlated quantity from visual parameters (area subtended, average dot-size) in each presented set of response alternatives using the method described by Gebuis and Reynvoet (2011). This method resulted in trial-specific response sets (7 dot-arrays) with no correlations between quantity, average dot size or area subtended ($-.2 < r < .2$ respectively).

To catch random responding, symbolic and non-symbolic control trials (16 ÷ 1, 16 x 1) were intermixed with calculation trials. Participants whose performance deviated more than 3 *SD* from the group mean or with symbolic control problem accuracy below 50% (chance = 14.3%) were excluded from the study.

To control for response choice magnitude effects, the same result values (with a random jitter in symbolic trials) or quantities (non-symbolic trials) were presented for multiplication and division; this resulted in different operands for multiplication versus division. We created a geometric series of 11 values ranging from 1/3 to three times the correct value (Knops, Viarouge, et al., 2009; Knops et al., 2013) for non-symbolic notation. For problems presented in symbolic notation the geometric series spanned values ranging from ½ times to two times the correct value. Arithmetic problems with symbolic range are reported in Appendix A. Because previous findings suggested that presenting the exact correct value may have made responses too accurate to detect symbolic response bias (Katz & Knops, 2014), symbolic response choices were jittered so that the correct answer was never presented (Knops, Viarouge, et al., 2009; Knops et al., 2013). To achieve this, all results were jittered either up or down by a random value which fell within plus or minus half the numerical interval between the correct result and the first deviant above or below it. The random value was drawn from a flat distribution on a logarithmic scale with a mean value of zero and was fixed for a given trial.

Participants completed 576 calculation trials (144 per condition) and 120 control trials (16 ÷ 1, 16 x 1; 30 per condition). To prevent the correct result always being the median value (*i.e.* 4^{th} response value rank of 7

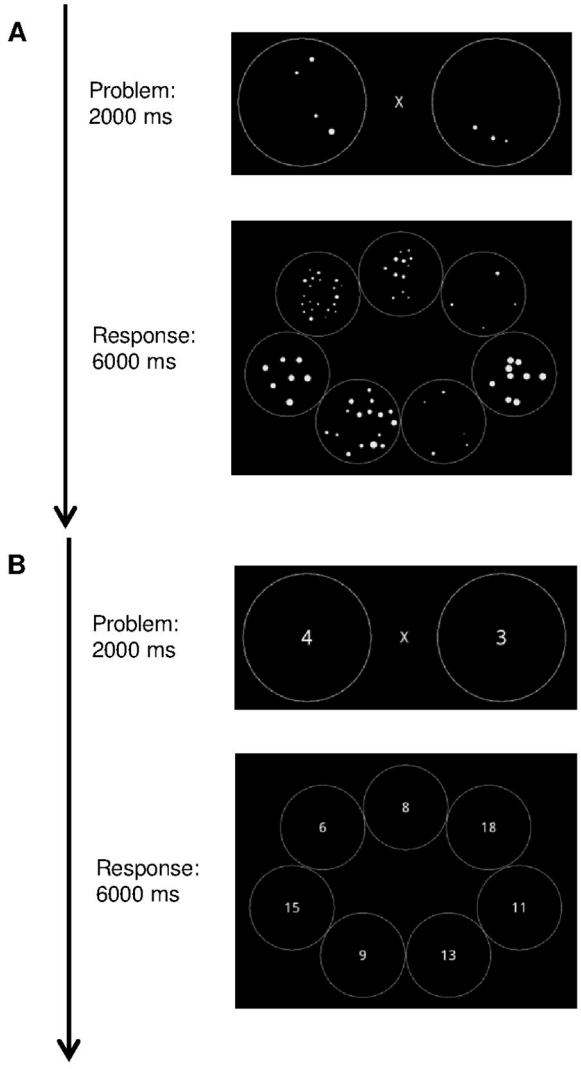

Figure 1. Calculation task. Symbolic and non-symbolic multiplication and division problems were presented in random order. Participants selected the responses with a mouse. (A) Non-symbolic multiplication. A set of 7 dot- arrays were created so that neither area subtended nor average dot-size correlated with quantity. (B) Symbolic multiplication. Response choices were jittered so that the exact answer was never shown.

choices), we varied the range of response alternatives and presented only seven out of the eleven response alternatives for each problem. The seven response alternatives corresponded to the smallest (low), largest (high) or middle (middle) range of response alternatives (Figure A.1). This means that the response alternative which is closest to the correct outcome and upon which the response alternatives were centred changed its ordinal rank for the different ranges. That is, for the smallest range, it fell close to the upper end (rank six out of seven), while it was located close to the lowest end for the highest range (rank two out of seven). In calculation trials, each problem was presented six times (two per range). At the end of the calculation task, participants were asked to describe how they solved the problems ("Please describe how you solved the problems. Which strategies did you use to decide for one of the response alternatives?"). Responses were typed in a blank document and no word limit was given. Strategy and word count were extracted from the responses.

Attention Task

Attention was assessed using an endogenous Posner cueing task (Posner, 1980), created and presented with MATLAB and the Psychophysics Toolbox (PTB) extension (Brainard, 1997; Pelli, 1997). Participants were seated approximately 60cm from the display and instructed to fixate on a red fixation dot (0.6°), which was presented at the center of the screen, flanked by a dark grey square box on either side (boxes: 3.3°, border width: 0.2°; see Figure 2). A yellow left, right, or double-ended (neutral cue) arrow (length: 2°) was presented in the center of the screen. A white target circle (2°) was flashed inside the box for 10ms, either on the side indicated by the arrow (valid cue), the opposite side (invalid cue) or randomly (neutral cue). Distance between fixation and target area border was 7.3°. The fixation sign without cue was presented for a variable duration between 800 and 1000 ms (*M* = 903.45ms, *SD* = 58.54ms). Cues were presented for a random duration between 1200 and 1800 ms (*M* = 1487ms, *SD* = 174.72ms). Participants were instructed to indicate the appearance of a target via button press (space bar) as quickly as possible. Participants were given a maximum of 1200ms to make a response. Participants completed 120 trials (84 valid (70%), 18 invalid (15%), 18 neutral (15%)).

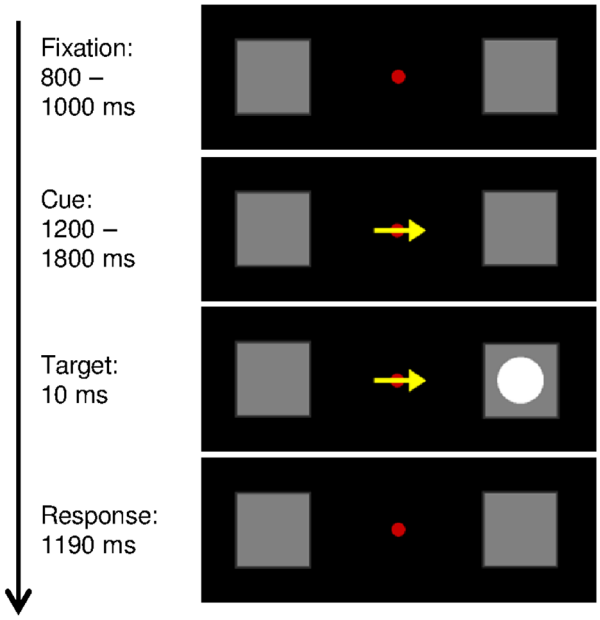

Figure 2. Attentional cueing task. An arrow appeared that either pointed in the direction of the subsequent target (valid cue, shown), the opposite direction (invalid cue) or in both directions (double-headed arrow, neutral cue). Participants were instructed to press the space bar as quickly as possible when they saw the target (white disk) that appeared after a variable SOA (1200 ms – 1800 ms).

Data Preparation and Analysis

Calculation task — Symbolic catch trials were always 16 multiplied or divided by 1. Although responses were jittered, so that '16' was never presented, if participants were paying attention and following the task directions, they should have been able to choose the value closest to 16 most of the time. Some degree of inaccuracy (i.e. < 100%) was expected since response choices were jittered. Therefore, we first eliminated subjects with accuracy less than 50% correct or greater than 3 *SD* from the group mean. This cut-off eliminated one subject whose accuracy (40%) deviated more than 3 SD from the group mean. In the remaining subjects (*n* = 17), the average accuracy was between 93% and 70% (*M* = 85%, *SD* = 0.36).

Next, because multiplication and division problems were not presented in separate blocks (Katz & Knops, 2014) and subjects might occasionally perform the wrong operation (*e.g.* multiplication instead of division), data was trimmed to exclude trials where the difference between the $\log_{10}$ of the chosen value and the $\log_{10}$ of the correct value was more than 3 SD from the subject's mean. When considering all conditions together, this excluded 40 calculation trials (0.4% of calculation responses, all non-symbolic). This was less than previous studies using this method (*e.g.* 1.8%; Knops, Viarouge, & Dehaene, 2009), possibly due to the simultaneous presentation of operands and the use of operation symbols (x, ÷) rather than letters (*e.g.* 'A' for addition & 'S' for subtraction; Knops, Thirion, et al., 2009). This likely minimized operation errors. Because a previous study (Katz & Knops, 2014) indicated greater variance for non-symbolic than symbolic trials, we decided to calculate each subject's mean separately for symbolic and non-symbolic notations and to exclude responses beyond mean plus/minus 3SD. This excluded 100 calculation trials (1% of calculation responses, all symbolic). Additionally, in 116 calculation trials (95 symbolic) and 7 catch trials (6 symbolic), no response was made within 6000ms (maximum response duration). The number of these timeout trials ranged from 0 to 37 trials per subject.

Attention task — In the Posner task, we first eliminated responses that were faster than 200ms, because these likely reflect premature responses. This eliminated 146 trials (7.3% of responses). The number of responses faster than 200ms per subject ranged from 0 to 37 responses (M = 8.6, SD = 10.4). We then eliminated responses where the RT was more than 3 SD from the subject's mean. This eliminated an additional 37 trials (2% of valid responses; 1-4 trials per subject, M = 2.2, SD = 1).

We computed validity effect (RT invalid minus RT valid), benefit (RT neutral minus RT valid) and cost (RT invalid minus RT neutral) as indices for orienting/selection and re-orienting, respectively.

Results

We first analyzed the distribution of responses to exclude the possibility that participants responded randomly. If participants responded non-randomly, then range and rank should have a significant effect on response choice (Katz & Knops, 2014). Indeed, we found that rank and range interacted for both multiplication and division in both symbolic and non-symbolic notation. We report these analyses in Appendix B.

Linear Increase of Response Value With Correct Value

Previous studies found that behavior was well described by Weber's law, suggesting a logarithmic compression of the underlying representation. This also appeared to be true in the present data (see Figure A.2). Transforming data log-scale also better meets the prerequisites of ANOVA, stipulating fixed variance (Katz & Knops, 2014). There was a linear increase of response value with correct value in both formats and operations (Table 1).

Table 1

Linear Increase of Response Value With Correct Value

	Linear scale			Log scale		
	t	Slope	95% CI	*t*	Slope	95% CI
Multiplication						
Symbolic	123.1**	0.96	0.94, 0.97	169.0**	0.98	0.97, 1.00
Non-symbolic	40.7**	1.24	1.18, 1.30	50.4**	1.05	1.01, 1.09
Division						
Symbolic	129.1**	0.96	0.95, 0.97	148.0**	0.98	0.97, 0.99
Non-symbolic	37.0**	0.95	0.90, 1.00	47.9**	0.92	0.89, 0.96

**Bonferroni corrected for multiple comparisons, $p < .013$.

Operational Momentum Effect

To investigate operational momentum, we entered the response bias, defined as the difference between the log chosen and the log correct values, into an ANOVA comprising the factors notation (symbolic, non-symbolic) and operation (multiplication, division). There was a significant interaction between operation and notation on response bias ($F(1,16) = 6.80$, $p = .018$). Since this makes interpretation of putative main effects difficult we no longer followed up on them. Therefore, simple main effects analysis was used to test the effect of operation, separately for symbolic and non-symbolic notations. For symbolic problems, operation did not have a significant effect on response bias ($F < 1$; Figure 3). However, symbolic multiplication (M = -0.007, 95% CI [-0.012, -0.002]), but not division (M = -0.006, 95% CI [-0.014, 0.001]), was significantly underestimated (i.e. mean response bias < 0; Figure 3).

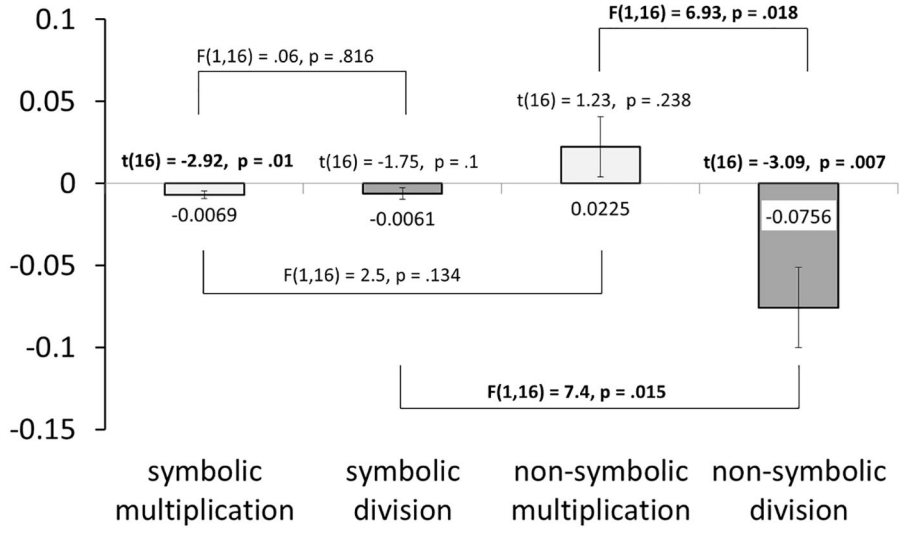

Figure 3. Response bias in symbolic and non-symbolic calculation. F-values represent simple main effects of operation (upper half) and notation (lower half). T-values represent one sample t-tests against a test value of zero. For symbolic problems (left), operation did not have a significant effect on response bias. The response bias for symbolic division (dark grey) was not significantly different than multiplication (light grey). However, symbolic multiplication problems were significantly underestimated. For non-symbolic problems (right), there was a significant effect of operation on response bias. Non-symbolic division (dark grey) problems were underestimated relative to multiplication (light grey). Only division, which was underestimated, showed a response bias significantly different from zero.

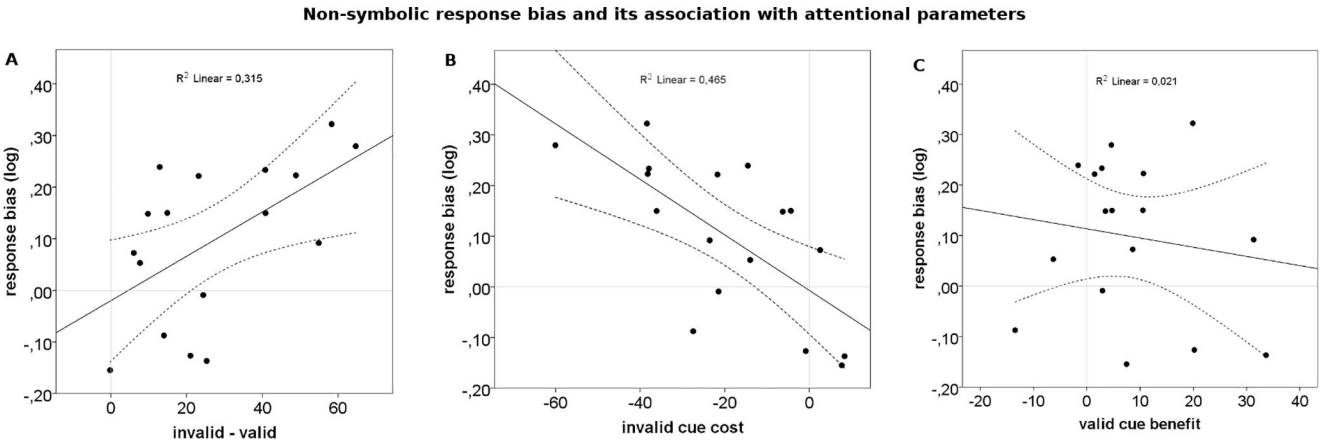

Figure 4. OM bias (log-scale) as a function of attentional parameters (A) validity effect (invalid – valid), (B) cost (neutral – invalid) and (C) benefit (neutral – valid) from the Posner task. Each data point corresponds to one subject.

For non-symbolic problems, operation had a significant effect on response bias ($F(1, 16) = 6.93$, $p < .05$, Bonferroni corrected; Figure 3). The mean response bias was positive for multiplication ($M = .023$, 95% CI [-0.016, 0.062]) and negative for division ($M = -0.076$, 95% CI [-0.127, -0.024]); difference = 0.098, 95% CI [0.019, 0.177]. Whereas we observed a significant underestimation in non-symbolic division ($t(16) = -3.09$, $p = .007$) we found no significant response bias in non-symbolic multiplication ($t(16) = 1.23$, $p = .238$; Figure 3).

Notation did not have a significant effect on response bias for multiplication ($F(1,16) = 2.50$, $p = .134$), but did for division ($F(1,16) = 7.40$, $p = .015$). The response bias was more pronounced (more negative) for non-symbolic than symbolic division ($t(16) = -2.72$, $p = .015$).

In sum, we replicated the results from Katz and Knops (2014), showing a significant OM-effect for non-symbolic multiplication and division but no OM-effect for symbolic operations, despite having encouraged approximate calculation by omitting the correct response from the symbolic response alternatives.

Can Attentional Orienting and Reorienting Predict Response Bias?

It has been put forward that the OM effect reflects the consequences of an attention-induced spatial displacement along the mental number line during the process of approximating the outcome of an arithmetic problem (Knops, Thirion, et al., 2009; Knops, Viarouge, & Dehaene, 2009; Knops, Zitzmann, & McCrink, 2013). Here, we tested the straight-forward hypothesis of an association between OM and spatial attention as measured in the Posner paradigm.

We first tested whether the relative OM bias, defined as the difference between operation-specific OM bias ((log correct minus log chosen Multiplication) minus (log correct minus log chosen Division)) correlated with the validity effect (valid cue minus invalid cue).

The validity effect was consistently larger than zero (i.e., faster response for valid than invalid) for all participants (M 27.5 ms; $t(16) = 5.64$, $p < .001$). To determine whether the validity effect could predict the difference in response bias between multiplication and division, each subject's validity effect was used as a predictor for the mean log-scale relative response bias (multiplication response bias minus division response bias) in a linear regression model. We restricted our analyses to the non-symbolic notation since symbolic OM

bias was not significant. The advantage of a valid cue compared to an invalid cue significantly predicted non-symbolic relative response bias ($F(1,16) = 6.90$, $p = .019$. It accounted for 31.5% of the variability in response bias and had a large effect size ($\bar{R}^2 = .269$).

To further examine the attentional mechanisms potentially driving the OM effect in non-symbolic notation, we separately examined the effect of attentional orienting (benefit of valid cue compared to neutral cue; $M = 8.3$ ms, $t(16) = 2.79$, $p = .013$) and reorienting (cost of invalid cue relative to neutral cue; $M = -19.2$ ms, $t(16) = -4.13$, $p = .001$)) on response bias using linear regression. The adjusted R^2 value ($\bar{R}^2$) was used to determine effect sizes using the cutoffs: small = .01 or 1%, medium = .1 or 10%, and large = .25 or 25% (Vacha-Haase & Thompson, 2004). Invalid cue cost significantly predicted relative response bias in non-symbolic problems; $F(1,16) = 13.02$, $p = .002$, $\bar{R}^2 = .464$. Valid cue benefit could not predict response bias difference in non-symbolic problems; $F(1,16) = 0.23$, $p = .638$, $\bar{R}^2 = -.044$. Results are shown in Figure 4. The correlation between cost and response bias was significantly more negative than the correlation between benefit and response bias (Hotelling's $t(13) = 2.15$; Steiger's $Z = 1.87$, $p < .05$ (one-tailed)).

All correlations remained by-and-large unchanged after partialing out age as a potential confound (r(validity effect, response bias) = .613, $p = .012$; r(benefit, response bias) = -.088, $p = .745$; r(cost, response bias) = -.704, $p = .002$).

These results suggest that in non-symbolic problems, attentional shifts, most likely re-orienting but not orienting, largely account for the difference in response bias between multiplication and division.

Discussion

In this study we examined two questions. First, we explored whether the previously reported absence of an OM effect in symbolic multiplication and division (Katz & Knops, 2014) may have been due to the presentation of the correct outcome among the response alternatives which may have triggered verbally mediated fact retrieval and hence lowered the impact of visuo-spatial processes. In the present study, we found no OM effect in symbolic notation even though we encouraged approximate calculation by presenting exclusively incorrect symbolic response alternatives. In contrast, a significant OM effect was observed for non-symbolic notation, replicating Katz and Knops (2014). Second, we explored the underlying mechanisms of the OM effect by testing the association of the OM effect with visuo-spatial attention as measured by a Posner paradigm. We found that re-orienting attention after the presentation of an invalid cue to the location of the target significantly correlated with the extent to which participants over- and underestimated the outcomes of non-symbolic multiplication and division problems.

The absence of a significant OM in symbolic multiplication and division in the current study is in line with previous results and implies that symbolic multiplication and division strongly rely on verbally mediated fact retrieval which is less prone to cognitive biases such as the OM. While a recent study described how the compression of the mental number line biases arithmetic fact retrieval (Didino, Knops, Vespignani, & Kornpetpanee, 2015), these spatial biases may be too subtle for the current paradigm.

We demonstrated a significant correlation between non-symbolic OM and measure of reorienting attention after invalid cues in the Posner paradigm. No correlation was observed between OM and orienting attention after

valid cueing. This is somewhat unexpected since the attention account of OM holds that attentional shifts propel participants too far along the mental number line. Hence, the benefit from a valid cue should correlate positively with OM. Yet, we did not observe any correlation between OM and cue benefit, a measure of attentional orienting. This might be due to a reduced variability in participants' performance in our Posner task which had slightly longer SOAs compared to the classical Posner paradigm. Variability may have been particularly reduced for the benefit measures as compared to cost measures, which, all else being equal, lowers correlation coefficients.

Reorienting is not a unitary process but has traditionally been subdivided into disengaging, shifting and reengaging attention to the new location (Posner & Petersen, 1990). In support of disengagement as a separate attentional mechanisms involved in reorienting, recent ERP studies revealed circumscribed posterior components linked with disengagement in the absence of attentional shifts (P4pc, Toffanin, de Jong, & Johnson, 2011), and separate from attentional selection (reversed N2pc, Eimer & Kiss, 2008). This more complex process may be subject to greater amount of variability across participants which, in turn, allows for higher correlation coefficients. Since attentional shifts and engaging are involved in both orienting and reorienting, the absence of a correlation between orienting and OM may as well imply that those processes that are unique to reorienting are at the heart of this association, namely disengaging. Future studies may use EEG to investigate this hypothesis.

Another question that arises from the current findings concerns the direction of the observed correlation. Why is the correlation between OM and reorienting positive, meaning that people with a large OM effect, i.e. larger deviations from the correct outcome exhibit larger costs for invalid cue both compared to neutral cues and valid cues? According to the above theoretical accounts of the OM we can break this question down into four aspects. Why would attentional reorienting correlate with (a) a heuristic according to which multiplication leads to larger outcomes and division to smaller outcomes, (b) flawed decompression, (c) competing spatial biases by the operands, the results or the outcome, or (d) attentional shifts along the mental number line?

According to the first account, no correlation with attentional measures would have been predicted. If any, there would have been a prediction that the expectation of larger or smaller outcomes creates attentional shifts to the left or right which might generate a coarse approximation of the result that can be used to check if a given outcome is plausible or not. However, what renders this hypothesis rather unlikely to account for the given results is that no gradation beyond a coarse "more or less" expectation is predicted.

Similarly, according to the compression-decompression approach OM results from flawed compression-decompression mechanisms. No attentional mechanism is involved in this process. According to this account, OM would scale with the size of both operands and results. However, previous research found that OM increases with increasing outcome (Knops, Viarouge, et al., 2009) but no association with operand magnitude was observed. In pointing tasks, OM was strongest when the second operand was zero, clearly speaking against any association between second operand size and OM (Pinhas & Fischer, 2008). Although the compression of the MNL is thought to be logarithmic, it is also subject to interindividual variance. A recent study tested whether the degree of compression of the approximate number system can actually serve to predict OM (Knops et al., 2014). However, the combination of crucial approximate number system parameters such as compression (as measured by the amount of underestimation in non-symbolic estimation) and precision (as measured by the Weber fraction) in a psychophysical model was not successful in predicting OM. While overall

biases in addition and subtraction involving non-symbolic quantities were well predicted by the interindividual variability of the parameters describing the approximate number system in the model, the operation-specific OM was not. Finally, it is hard to see how this would conceptually relate to the costs of reorienting attention to an invalidly cued target. Further evidence against a heuristic-based account comes from the study of Magethis and colleagues (Marghetis et al., 2014) who found systematic biases in mouse pointer trajectories when participants indicated the correct outcome for addition and subtraction problems. Exploiting the fact that on-task tracking of mouse trajectories provides a time-resolved window on cognition, Margethis and colleagues found that the time course of this spatial deviation sequentially reflected the serial impact of first operand, operator and second operand. The authors conclude that neither a heuristic-based nor a compression account would predict this pattern of results, which is in line with a spatial account of the OM (Marghetis et al., 2014). Together, this suggests that the OM is most likely not fully accounted for by flawed compression-decompression mechanisms and suggests the origin in parameters outside the ANS.

With respect to the spatial accounts of the OM, a clear prediction comes from the attentional shift hypothesis which predicts a clear association between attentional parameters and OM. The current results partially confirmed this by the correlation between OM and reorienting, providing further evidence for a role of spatial attention during approximate arithmetic and, more specifically, for the idea that the OM results from attentional mechanisms. Larger reorienting costs may reflect highly efficient orienting mechanisms that need to be overruled after invalid cueing. These results are also in line with the finding that symbolic addition problems are solved faster when the second operand is presented on the right compared to left-sided presentation (Mathieu, Gourjon, Couderc, Thevenot, & Prado, 2016). For subtraction an analog advantage for left-sided operands was observed. Crucially, these authors also failed to find a benefit for lateralized operands in multiplication problems, highlighting the different cognitive processes that contribute to multiplication on the one side and addition or subtraction on the other. Recent results corroborate a tight link between the ocular movement system and OM (Klein et al., 2014). Relative to the first fixation, participants subsequently moved their eyes to the right for additions and to the left for subtractions, paralleling previous fMRI results (Knops, Thirion, et al., 2009). This dynamic process of adjusting fixation during the course of arithmetic processing may in part have been specific to the task which required indicating the outcome by pointing to the respective location on a number line. It is also conceivable, however, that approximate calculation is mediated by the dynamic updating within a coordinate transformation system in parietal cortex.

The spatial competition account, in contrast, does not predict any correlation between attentional shifts and OM. According to this account OM is largest when competition between spatial positions of operands and the results on the MNL are minimal. The strong OM in zero problems where only the first operand and the operational sign induce a spatial bias is explained by the absence of a spatial bias induced by zero (Pinhas & Fischer, 2008; Pinhas et al., 2015). Zero is either not represented on the MNL or triggers rule-based procedures. While we cannot test this prediction in the current experiment, we may interpret reorienting in terms of spatial competition between a cued position and the appearance of a target at an uncued position. The spatial competition account would predict that less competition is associated with larger OM. However, we observe the opposite pattern of results in the current study where larger amount of spatial competition is associated with larger OM. Hence, under the premise of interpreting the reorienting effect in terms of spatial competition our results provide evidence against this account. It should be noted, however, that the spatial competition account was initially proposed in the context of addition and subtraction. It is unclear whether it would also hold for multiplication and division where the split between operands and results is much larger.

Limitations

As this study provides only correlational evidence, no causal inference can be drawn. Further experimental work is required to elucidate the neurocognitive mechanisms underlying the OM effect. The absence of significant correlations between operational momentum bias and attentional measures may in part be due to the long SOAs in our version of the Posner task. Future studies could increase variability by using shorter SOAs, which might favor finding higher correlations. The fact that we related accuracy-based operational momentum bias with speed–based measures of the attentional capacities might further have reduced our statistical power. The limited sample size may raise concerns about (a) stability and reliability of the data and (b) to what extent the observed correlations were due to the increased variability in our sample. We checked whether our results were driven by some outliers by separately excluding all possible combinations of 1, 2 or 3 participants from the sample. We found single participants to have only a minor impact on the correlation pattern. Even when excluding the two participants with the most extreme value pairings, correlations by and large remained significant or marginally significant. To protect against the possibility that age was a confounding variable that drives our results, we recalculated the main findings after partialing out age and found the major findings of the study unchanged. Finally, while one may be concerned about the role of counting in computing the approximate results, we would reason that counting does play a major role in explaining our results. These additional analyses are can be found in Appendix C. Albeit OM and attention appear to be functional related on a conceptual level, future studies could use more comparable parameters for measuring these concepts.

Conclusions

To sum up, we failed to observe an OM effect in symbolic multiplication and division. This is in line with previous findings (Katz & Knops, 2014; Mathieu et al., 2016) and suggests that verbally mediated retrieval of arithmetic facts from long-term memory is less prone to spatial biases, yet not immune (Cavdaroglu & Knops, 2016; Didino et al., 2015). In contrast, we found a significant OM effect in non-symbolic notation, reflecting that participants systematically overestimated results of multiplication problems while underestimating results of division problems. The non-symbolic OM effect correlated with attentional parameters measured in a Posner paradigm. By differentially analyzing benefit and cost measures we found this correlation to mainly results from the costs of reorienting attention after invalid cues. While the exact mechanisms driving this correlation remain elusive, these results provide further evidence for the attentional shift hypothesis.

Funding

This work was funded by a grant (KN 959/2) from Deutsche Forschungsgemeinschaft (DFG, German Research Council) to André Knops in support of Curren Katz and Hannes Hoesterey.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgments

The authors have no support to report.

References

Bächtold, D., Baumuller, M., & Brugger, P. (1998). Stimulus-response compatibility in representational space. *Neuropsychologia, 36*(8), 731-735. doi:10.1016/S0028-3932(98)00002-5

Beck, J. M., Latham, P. E., & Pouget, A. (2011). Marginalization in neural circuits with divisive normalization. *The Journal of Neuroscience, 31*(43), 15310-15319. doi:10.1523/JNEUROSCI.1706-11.2011

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436. doi:10.1163/156856897X00357

Campbell, J. I., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*(2), 299-315. doi:10.1037/0096-3445.130.2.299

Cavdaroglu, S., & Knops, A. (2016). Mental subtraction and multiplication recruit both phonological and visuospatial resources: Evidence from a symmetric dual-task design. *Psychological Research, 80*, 608-624. doi:10.1007/s00426-015-0667-8

Chen, Q., & Verguts, T. (2012). Spatial intuition in elementary arithmetic: A neurocomputational account. *PLoS One, 7*(2), Article e31180. doi:10.1371/journal.pone.0031180

Dehaene, S. (2001). Subtracting pigeons: Logarithmic or linear? *Psychological Science, 12*(3), 244-246. doi:10.1111/1467-9280.00343

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384-398. doi:10.1016/j.neuron.2007.10.004

Didino, D., Knops, A., Vespignani, F., & Kornpetpanee, S. (2015). Asymmetric activation spreading in the multiplication associative network due to asymmetric overlap between numerosities semantic representations? *Cognition, 141*, 1-8. doi:10.1016/j.cognition.2015.04.002

Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 20*(8), 1423-1433. doi:10.1162/jocn.2008.20099

Fischer, M. H., & Knops, A. (2014). Attentional cueing in numerical cognition. *Frontiers in Psychology, 5*, Article 1381. doi:10.3389/fpsyg.2014.01381

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods, 43*(4), 981-986. doi:10.3758/s13428-011-0097-5

Hartmann, M., Mast, F. W., & Fischer, M. H. (2015). Spatial biases during mental arithmetic: Evidence from eye movements on a blank screen. *Frontiers in Psychology, 6*, Article 12. doi:10.3389/fpsyg.2015.00012

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*(6), 435-448. doi:10.1038/nrn1684

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*(5), 822-851. doi:10.3758/BF03196775

Hubbard, T. L. (2015). The varieties of momentum-like experience. *Psychological Bulletin, 141*(6), 1081-1119. doi:10.1037/bul0000016

Katz, C., & Knops, A. (2014). Operational momentum in multiplication and division? *PLoS One, 9*(8), Article e104777. doi:10.1371/journal.pone.0104777

Klein, E., Huber, S., Nuerk, H.-C., & Moeller, K. (2014). Operational momentum affects eye fixation behaviour. *Quarterly Journal of Experimental Psychology, 67*(8), 1614-1625. doi:10.1080/17470218.2014.902976

Knops, A., Dehaene, S., Berteletti, I., & Zorzi, M. (2014). Can approximate mental calculation account for operational momentum in addition and subtraction? *Quarterly Journal of Experimental Psychology, 67*(8), 1541-1556. doi:10.1080/17470218.2014.890234

Knops, A., Thirion, B., Hubbard, E. M., Michel, V., & Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science, 324*(5934), 1583-1585. doi:10.1126/science.1171599

Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect. *Attention, Perception & Psychophysics, 71*(4), 803-821. doi:10.3758/APP.71.4.803

Knops, A., Zitzmann, S., & McCrink, K. (2013). Examining the presence and determinants of operational momentum in childhood. *Frontiers in Psychology, 4*, Article 325. doi:10.3389/fpsyg.2013.00325

Lindemann, O., & Tira, M. D. (2011). Operational momentum in numerosity production judgments of multi-digit number problems. *The Journal of Psychology, 219*(1), 50-57. doi:10.1027/2151-2604/a000046

Marghetis, T., Núñez, R., & Bergen, B. K. (2014). Doing arithmetic by hand: Hand movements during exact arithmetic reveal systematic, dynamic spatial processing. *Quarterly Journal of Experimental Psychology, 67*(8), 1579-1596. doi:10.1080/17470218.2014.897359

Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., & Prado, J. (2016). Running the number line: Rapid shifts of attention in single-digit arithmetic. *Cognition, 146*, 229-239. doi:10.1016/j.cognition.2015.10.002

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics, 69*(8), 1324-1333. doi:10.3758/BF03192949

McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology, 103*(4), 400-408. doi:10.1016/j.jecp.2009.01.013

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience, 32*, 185-208. doi:10.1146/annurev.neuro.051508.135550

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437-442. doi:10.1163/156856897X00366

Pinhas, M., & Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition, 109*(3), 408-415. doi:10.1016/j.cognition.2008.09.003

Pinhas, M., Shaki, S., & Fischer, M. H. (2014). Heed the signs: Operation signs have spatial associations. *Quarterly Journal of Experimental Psychology, 67*(8), 1527-1540. doi:10.1080/17470218.2014.892516

Pinhas, M., Shaki, S., & Fischer, M. H. (2015). Addition goes where the big numbers are: Evidence for a reversed operational momentum effect. *Psychonomic Bulletin & Review, 22*(4), 993-1000. doi:10.3758/s13423-014-0786-z

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*(1), 3-25. doi:10.1080/00335558008248231

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42. doi:10.1146/annurev.ne.13.030190.000325

Pouget, A., Deneve, S., & Duhamel, J.-R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience, 3*(9), 741-747. doi:10.1038/nrn914

Toffanin, P., de Jong, R., & Johnson, A. (2011). The P4pc: An electrophysiological marker of attentional disengagement? *International Journal of Psychophysiology, 81*(2), 72-81. doi:10.1016/j.ijpsycho.2011.05.010

Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology, 51*(4), 473-481. doi:10.1037/0022-0167.51.4.473

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review, 19*(6), 1057-1064. doi:10.3758/s13423-012-0295-x

Wiemers, M., Bekkering, H., & Lindemann, O. (2014). Spatial interferences in mental arithmetic: Evidence from the motion-arithmetic compatibility effect. *Quarterly Journal of Experimental Psychology, 67*(8), 1557-1570. doi:10.1080/17470218.2014.889180

Appendices

Appendix A: Arithmetic Problems and Response Alternatives

The problems we used were identical to the ones in Katz and Knops (2014). The responses for symbolic and non-symbolic problems were 0.5 × correct outcome < correct outcome < 2 × correct outcome and 0.33 × correct outcome < correct outcome < 3 × correct outcome, respectively. Eleven bins within the range of the response alternatives were created to form of a geometric series. The symbolic problems and the response range can be found in Table A.1.

Table A.1

Arithmetic Problems (Multiplication and Division), Including Response Alternatives

Operands		Response alternatives										
1st	2nd	1	2	3	4	5	6	7	8	9	10	11
Multiplication												
4	3	6	7	8	9	10	12	14	16	18	21	24
6	3	9	10	12	14	16	18	21	24	27	31	36
6	4	12	14	16	18	21	24	28	32	36	42	48
7	3	11	13	14	16	19	21	25	28	32	37	42
7	4	14	16	18	21	24	28	32	37	42	49	56
7	6	21	24	28	32	37	42	48	55	64	73	84
8	3	12	14	16	18	21	24	28	32	36	42	48
8	4	16	18	21	24	28	32	37	42	49	56	64
8	6	24	28	32	36	42	48	55	63	73	84	96
9	3	14	16	18	21	24	27	31	36	41	47	54
9	4	18	21	24	27	31	36	41	48	55	63	72
9	6	27	31	36	41	47	54	62	71	82	94	108
12	3	18	21	24	27	31	36	41	48	55	63	72
12	4	24	28	32	36	42	48	55	63	73	84	96
13	4	26	30	34	39	45	52	60	69	79	91	104
13	6	39	45	51	59	68	78	90	103	118	136	156
14	3	21	24	28	32	37	42	48	55	64	73	84
14	6	42	48	55	64	73	84	96	111	127	146	168
16	3	24	28	32	36	42	48	55	63	73	84	96
16	4	32	37	42	49	56	64	74	84	97	111	128
17	3	26	30	34	39	45	51	59	68	78	89	102
17	4	34	39	45	52	59	68	78	90	103	118	136
19	3	29	33	38	44	50	57	66	76	87	99	114
19	4	38	44	50	58	66	76	87	100	115	132	152
Division												
36	2	9	10	12	14	16	18	21	24	27	31	36
48	2	12	14	16	18	21	24	28	32	36	42	48
48	4	6	7	8	9	10	12	14	16	18	21	24
54	2	14	16	18	20	24	27	31	36	41	47	54
63	3	11	12	14	16	18	21	24	28	32	37	42
96	2	24	28	32	36	42	48	55	63	73	84	96
96	4	12	14	16	18	21	24	28	32	36	42	48
108	3	18	21	24	27	31	36	41	48	55	63	72
112	4	14	16	18	21	24	28	32	37	42	49	56
126	3	21	24	28	32	37	42	48	55	64	73	84

Operands		Response alternatives										
1st	2nd	1	2	3	4	5	6	7	8	9	10	11
128	2	32	37	42	49	56	64	74	84	97	111	128
128	4	16	18	21	24	28	32	37	42	49	56	64
136	2	34	39	45	52	59	68	78	90	103	118	136
144	3	24	28	32	36	42	48	55	63	73	84	96
144	4	18	21	24	27	31	36	41	48	55	63	72
152	2	38	44	50	58	66	76	87	100	115	132	152
153	3	26	29	34	39	44	51	59	67	77	89	102
156	2	26	30	34	39	45	52	60	69	79	91	104
156	3	26	30	34	39	45	52	60	69	79	91	104
162	3	27	31	36	41	47	54	62	71	82	94	108
168	2	42	48	55	64	73	84	96	111	127	146	168
168	4	21	24	28	32	37	42	48	55	64	73	84
171	3	29	33	38	43	50	57	65	75	86	99	114
192	4	24	28	32	36	42	48	55	63	73	84	96

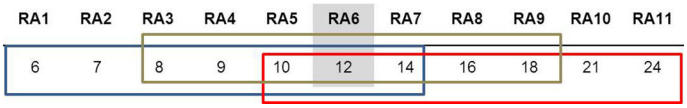

Figure A.1. Exemplary depiction of how we varied the ordinal rank of the correct outcome (gray column) within the low (blue), middle (brown), and high (red) range of response alternatives (RA 1 to RA 11). Example depicts the outcomes for the problem 3 × 4 in the symbolic notation.

Appendix B: Analysis of Interaction Between Range and Rank to Ensure Participants Were not Responding Randomly

Symbolic Multiplication

There was a significant interaction between range and rank ($F(12,192) = 143.315$, $p < .001$, $\varepsilon = .143$). Simple main effects analysis revealed a significant effect of rank on response percentage for all response ranges (low, medium, high). For all response ranges, post-hoc pairwise comparisons confirmed that the correct choice was selected significantly more often than all other choices (Table A.2).

Table A.2

Simple Main Effect of Rank on Response Percentage

				Sphericity[a]	
Range	***F*(6,96)**	***p***	**Partial η^2**	**χ^2**	**ε**
Symbolic multiplication					
Low, 5th correct**	168.847	<.001	.913	157.0	.230
Med., 4th correct**	131.614	<.001	.892	132.2	.243
High, 3rd correct**	207.360	<.001	.928	183.4	.275
Non-symbolic multiplication					
Low, 5th correct**	7.209	.001	.311	39.5	.474
Med., 4th correct	2.628	.058	.141	35.1	.522
High, 3rd correct*	3.084	.038	.162	47.3	.484
Symbolic division					
Low, 5th correct**	109.958	<.001	.873	118.6	.266
Med., 4th correct	142.816	<.001	.889	118.6	.283
High, 3rd correct	116.792	<.001	.880	127.7	.288
Non-symbolic division					
Low, 5th correct**	3.043	.046	.160	59.4	.433
Med., 4th correct*	4.916	.012	.235	75.8	.350
High, 3rd correct**	14.402	<.001	.474	83.0	.355

[a]Greenhouse-Geisser corrected for violations of sphericity as measured by Mauchley's Test of Sphericity.

*Significant simple main effect only.

**At least 1 significant post-hoc pairwise comparison (Bonferroni-corrected $p < .05$).

Symbolic Division

There was a significant interaction between range and rank ($F(12,192) = 121.256$, $p < .001$, $\varepsilon = .133$). Similar to symbolic multiplication, simple main effects analysis revealed a significant effect of rank on response percentage for all response ranges and post-hoc pairwise comparisons confirmed that the correct choice was selected significantly more often than all other choices (Table A.3).

Table A.3

Estimated Marginal Means, Post-Hoc Pairwise Comparisons

	1	2	3	4	5	6	7
Symbolic multiplication							
L	$0.1^{3,4,5,6,7}$	$0.1^{3,4,5,6,7}$	$0.3^{1,2,4,5}$	$1.2^{1,2,3,5,7}$	$5.6^{1,2,3,4,6,7}$	$1.0^{1,2,5}$	$0.4^{1,2,4,5}$
M	$0.2^{2,3,4}$	$0.4^{3,4,5,7}$	$1.2^{1,2,4,6,7}$	$5.2^{1,2,3,5,6,7}$	$1.1^{1,2,4,6,7}$	$0.2^{3,4,5}$	$0.1^{2,3,4,5}$
H	$0.4^{2,3,4,6,7}$	$1.3^{1,3,5,6,7}$	$5.4^{1,2,4,5,6,7}$	$1.0^{1,3,5,6,7}$	$0.1^{2,3,4}$	$0.0^{1,2,3,4}$	$0.0^{1,2,3,4}$
Symbolic division							
L	$0.1^{4,5,6}$	$0.2^{4,5,6}$	$0.3^{4,5,6}$	$1.2^{1,2,3,5,7}$	$5.0^{1,2,3,4,6,7}$	$1.1^{1,2,3,5,7}$	$0.4^{4,5,6}$
M	$0.3^{3,4,5}$	$0.2^{3,4,5}$	$1.2^{1,2,4,6,7}$	$5.0^{1,2,3,5,6,7}$	$1.0^{1,2,4,6,7}$	$0.3^{3,4,5,7}$	$0.1^{3,4,5,6}$
H	$0.4^{2,3}$	$1.4^{1,3,5,6,7}$	$5.0^{1,2,4,5,6,7}$	$1.0^{3,5,6,7}$	$0.3^{2,3,4}$	$0.1^{2,3,4}$	$0.0^{2,3,4}$
Non-symbolic multiplication							
L	1.0^{5}	$1.0^{4,5,6}$	$1.0^{5,6}$	1.2^{2}	$1.5^{1,2,3}$	$1.8^{2,3}$	1.6
M	1.0	1.0	1.1	1.6	1.4	1.5	1.0
H	1.2	1.3	1.6	1.3	1.4	1.1	1.0
Non-symbolic division							
L	1.0	1.4	1.5^{6}	1.4^{6}	1.4^{6}	$0.8^{3,4,5}$	1.0
M	1.6	1.7	1.5	1.3	1.0	0.8	0.6
H	$2.4^{6,7}$	$2.1^{3,4,5,6,7}$	$1.3^{2,6,7}$	$1.0^{2,6}$	$0.8^{2,6}$	$0.4^{1,2,3,4,5}$	$0.5^{1,2,3}$

Note. L = Low-5th correct. M = Medium-4th correct. H = High-3rd correct.

^{1,2,3,4,5,6,7} Significantly different (Bonferroni-corrected) from rank ^{n}

Non-Symbolic Multiplication

There was a significant interaction between range and rank ($F(12,192) = 6.940$, $p < .001$, $\varepsilon = .566$). Simple main effects analysis revealed a significant effect of rank on response percentage for trials where the low high response range was presented, but not for the medium range. When the medium and high ranges of response choices were presented, there were no significant differences in response percentages between any of the seven response choices (Table A.2). Therefore, random responding cannot be excluded in medium and high range non-symbolic multiplication trials. In the low range, the 6th choice (too large) was selected significantly more than the 2nd or 3rd choice, the 5th choice (correct choice) was selected significantly more often that the 1st, 2nd or 3rd and the 4th choice significantly more than the 2nd choice (Table A.3). Thus, the percentage of responses at the 4th, 5th and 6th choice were not significantly different for low range trials.

Non-Symbolic Division

There was a significant interaction between range and rank ($F(12,192) = 10.337$, $p < .001$, $\varepsilon = .414$). Simple main effects analysis revealed a significant effect of rank on response percentage for all response ranges. However, after correction for multiple comparisons, post-hoc pairwise comparisons revealed significant differences between the seven response choices only for low and high range trials (Table A.2). When the medium range of response choices was presented, there were no significant differences in response percentages between any of the seven response choices (Table A.2). This indicates that although random responding cannot be excluded in medium range non-symbolic division trials, it can be in both low and high range trials. In low range trials, the 3rd (too small), 4th (too small) and 5th (correct) choices were selected significantly more than the 6th choice (too large), but there was no significant difference between the 1st through 5th choice (Table A.3). In high range trials, the smallest choices (1st, 2nd & correct, 3rd) were selected significantly more often than the two largest choices (6th & 7th) and the 6th was selected less often than all of the smaller choices (1st-5th). The 2nd choice was selected more often than all of the larger choices (3rd-7th), including the correct choice (Table A.2). This indicates significant pattern of choosing one smaller (2nd choice) than the correct choice when the 3rd choice was correct (high range), a possibly random response pattern when the middle choice (4th) was correct (medium range) and a non-random response pattern driven by a decreased likelihood of choosing the 6th choice than the 3rd, 4th and 5th when the 5th was correct (low range).

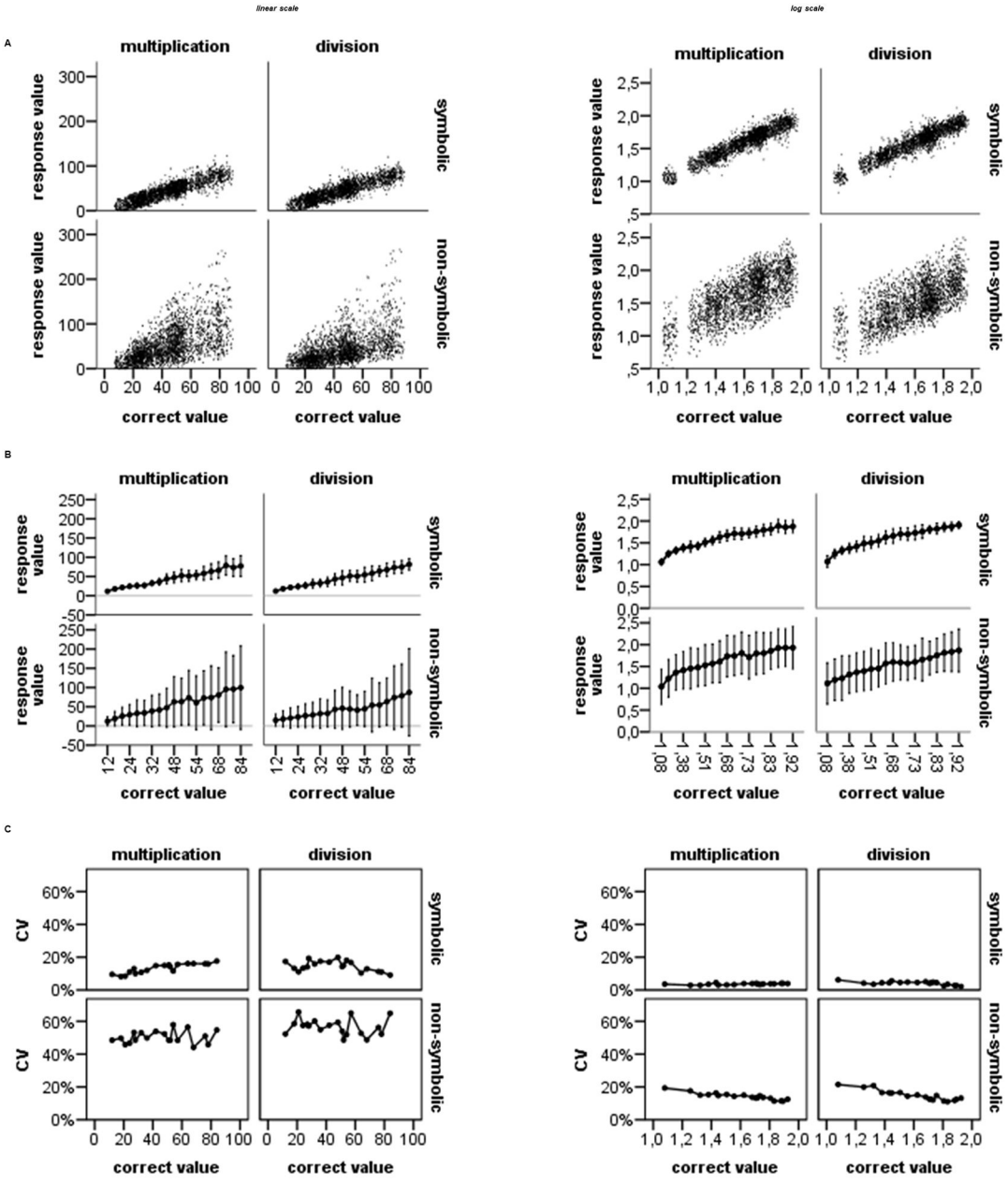

Figure A.2. Response value as a function of correct value on linear and log-scale non-aggregated data. (A) The non-aggregated response value was plotted as a function of correct value on the linear (left) and log (right) scale. Each point represents one case and number of cases is indicated by dot density (*i.e.* darker color). (B) As predicted by Weber's law, linear response value and variability (error bars represent ± 2 SD) (left) increased as a function of correct value. This could clearly be seen in non-symbolic problems. When the $\log_{10}$ of the response and correct value was used, response value increased as a function of correct value, but variability remained constant. (C) The dispersion of the response choices, measured by the mean-centered coefficient of variation (CV), was more constant when the log-scale data was used (right).

Appendix C: Analyses Concerning Reliability and Stability of the Observed Results

In the following we report the results of additional analyses (1) to verify the stability and reliability of the data, as well as the (2) impact of age, a (3) Bayesian analysis of the reported null effects, and (4) why we think counting does not play a major role in the current experiment.

1. Are the Observed Correlations Due to Outliers?

We checked the effect of separately excluding all possible combinations of 1, 2 or 3 participants from the sample on our main findings, that is the correlation between

(1) the validity effect (RT valid – RT invalid) and the OM bias (OM_multiplication – OM_division)

Excluding 1 participant: When excluding the two participants (#13 and #16) marked in red in the Figure A.3, correlation was r = .485 and significance drops to a marginal significance of p = .0569. These two participants were aged 25 and 29, hence not in the extreme age range of our sample (19-74).

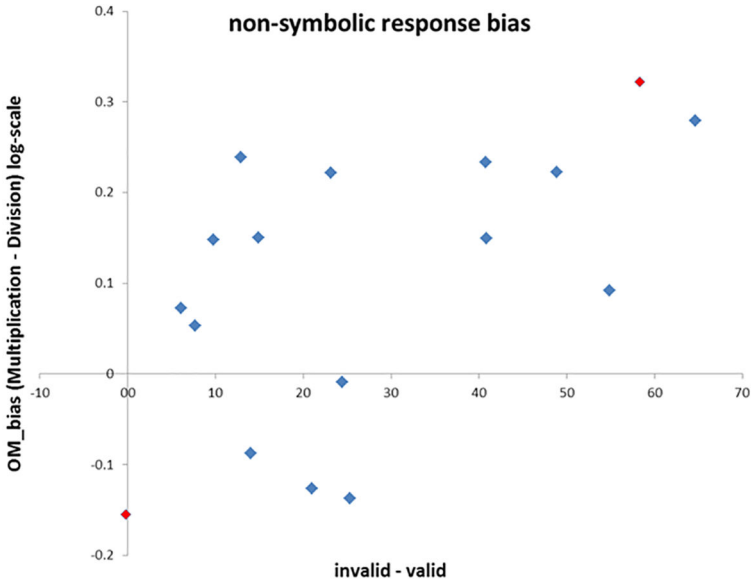

Figure A.3. Scatter plot of the validity effect against the OM bias. Each dot represents one participant.

When excluding more than one participant, 32 out of 136 combinations of two participants (23.5%), or 229 out of 680 combinations of three participants were not significant (33.7%).

(2) the benefit and the OM bias

Excluding one participant did not change the fundamental pattern of results. All correlations remained significant (p > .05).

When excluding more than one participant, 0 out of 136 combinations of two participants (0%), or 11 out of 680 combinations of three participants were not significant (1.62%).

(3) the cost and the OM bias

Only when excluding one participant (#17) the correlation (r = -. 489) was marginally significant only (p = .055).

When excluding more than one participant, 20 out of 136 combinations of two participants (14.7%), or 183 out of 680 combinations of three participants were not significant (26.9%).

2. Are Our Results Due to Some Confounding Impact of Age?

Partialing out age did not significantly impact the correlation between OM bias and validity effect (r_{age}(bias, validity effect) = .613, (p = .0115) or the correlation between cost and OM bias (r_{age}(bias, cost) = .70, p = .002), both of which were significant before partialing out the influence of age, too. The partial correlation between benefit and OM bias corrected for age (r_{age}(bias, benefit) = -.09, p = .75) was not significant, thus remained unchanged, too.

Together with the above results we think that these results imply rather stable data even given the small sample size. Age was not a major factor driving our results. On the contrary, our convenience sample increases the generalizability of the data to the population since we do not – as the majority of studies in the field of experimental psychology – artificially restrict our sample to university students in their early twenties. Further, it should be mentioned that we chose a rather conservative procedure when excluding one participant before engaging in our analyses based on the participant's low performance.

3. Bayesian Analysis of the Correlation Between Benefit and OM Bias.

For the non-significant correlation between benefit and OM bias, a Bayes factor BF10 = 0.214 indicates evidence in favor of the null hypothesis. It is 1/0.214 = 4.673 more likely that the data occurred under the H0 than the H1 (Wetzels & Wagenmakers, 2012).

4. Does Counting Play a Major Role in Explaining Our Results?

For the following reasons we consider counting as an unlikely strategy to account for the results. The majority of operands cannot be counted during two seconds. Assuming a counting rate of ~250 ms per item, participants may have counted up to ~8 items. This allowed counting only those problems where the sum of the operands would be around 8. This was the case for four problems only. What makes it even more unlikely that participants counted is that the results screen contained seven dot patterns, clearly exceeding the time limit to count all dots within only 6 seconds. Mean reaction time for non-symbolic problems was 2.595 second (SD = 1.129), showing that the vast majority of responses was provided within less than 3.7 seconds (M + 1 SD). Even if participants were able to count one of the operands (e.g. where it was 2, 3 or 4), this means that participants had to divide or multiply an approximate second quantity by/with an exact value, again leading to approximate values. Last, we think it is very unlikely that participants engaged in tedious counting routine over a period of approximately 1 hour.

NeuroImage 116 (2015) 135–148

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception

Seda Cavdaroglu *, Curren Katz, André Knops

Humboldt Universität zu Berlin, Faculty of Life Sciences, Department of Psychology, Rudower Chaussee 18, 12489 Berlin, Germany

ARTICLE INFO

Article history:
Received 23 December 2014
Accepted 7 April 2015
Available online 14 April 2015

Keywords:
Numerical cognition
fMRI
Multivariate pattern analysis
Number sense
Supramodal numerosity

ABSTRACT

It has been widely debated whether the parietal cortex stores an abstract representation of numerosity that is activated for Arabic digits as well as for non-symbolic stimuli in a sensory modality independent fashion. Some studies suggest that numerical information in time-invariant (simultaneous) symbolic and non-symbolic visual stimuli is represented in the parietal cortex. In humans, whether the same representation is activated for time-variant (sequential) stimuli and for stimuli coming from different modalities has not been determined. To investigate this idea, we measured the brain activation of healthy adults performing estimation and/or comparison of sequential visual (series of dots) and auditory (series of beeps) numerosities. Our experimental design allowed us to separate numerosity estimation from comparison and response related factors. The BOLD response in the parietal cortex increased only when participants were engaged in the comparison of two consecutive numerosities that required a response. Using multivariate pattern analysis we trained a classifier to decode numerosity in various regions of interest (ROI). We failed to find any parietal ROI where the classifier could decode numerosities during the estimation phase. Rather, when participants were not engaged in comparison we were able to decode numerosity in an auditory cortex ROI for auditory stimuli and in a visual cortex ROI for visual stimuli. On the other hand, during the response period the classifier successfully decoded numerosity information in a parietal ROI for both visual and auditory numerosities. These results were further confirmed by support vector regression. In sum, our study does not support the involvement of the parietal cortex during estimation of sequential numerosity in the absence of an active task with a response requirement.

Introduction

The ability to comprehend and differentiate quantities is an essential cognitive capacity. We use this ability for simple automated acts such as calculating the grip aperture before we grasp a cup as well as when we are performing mathematical operations. Certain developmental or genetic disabilities can render people incapable of performing even very simple calculations or understanding quantities in general (Chu et al., 2013; Mazzocco, 2009; McCloskey, 1992). Therefore, over the last few decades a considerable amount of research has focused on the neural basis of numerical cognition. One of the most influential ideas in the field has been the triple-code model (TCM) of numerical cognition suggested by Dehaene (1992) and Dehaene and Cohen (1995). According to a recent update of the TCM (Arsalidou and Taylor, 2011), numerical information is represented by three interacting but distinct codes, each associated with separate cortical structures. The Arabic number code is used for multi-digit arithmetic operations. Visually presented Arabic digits are associated with activity in bilateral fusiform and lingual regions. The verbal number code is used for memorized arithmetic problems that are phonologically coded, such as single-digit multiplication and addition. Memorized arithmetic facts are associated with activity in predominantly left-hemisphere perisylvian language areas and the left angular gyrus. The abstract magnitude code is used for non-verbal quantity and magnitude understanding. Tasks requiring magnitude processing are associated with activity in the horizontal aspect of the intraparietal sulcus (hIPS). Although they can function independently in certain tasks, these three components are thought to interact with each other for more complex numerical operations (Arsalidou and Taylor, 2011). Interestingly, while the temporal cortex and articulatory loop are involved in learned aspects of arithmetic (e.g., the acquisition of the Arabic digit system and memorization of simple calculation facts; see Arsalidou and Taylor, 2011 for a review), the parietal cortex (specifically the intraparietal sulcus-IPS) is thought to contain an innate mechanism that humans share with other species, often referred to as 'number sense' (Nieder and Dehaene, 2009). The number sense is thought to enable us to comprehend quantities in an abstract fashion, independent of notation (e.g., symbolic or non-symbolic), presentation format (e.g., simultaneous or sequential) or sensory modality (e.g., visual or auditory; Dehaene et al., 2004).

A number of studies have examined the role of the IPS in representing numerical information presented in different notations (symbolic or non-symbolic). Eger et al. (2009) used multivariate pattern analysis (MVPA) to

* Corresponding author at: Rudower Chaussee 18, Room 1.222, 12489 Berlin, Germany.
E-mail address: seda.cavdaroglu@hu-berlin.de (S. Cavdaroglu).

http://dx.doi.org/10.1016/j.neuroimage.2015.04.019
1053-8119/

decode symbolic (Arabic digits) and non-symbolic (dot-arrays) numerosities in a parietal ROI during numerical comparison. Interestingly, the classifier trained with Arabic digits generalized to non-symbolic numerosities, although the reverse generalization (from non-symbolic numerosities to Arabic digits) was not observed, supporting only partial notation independence. Piazza et al. (2007) observed a similar asymmetry. Exploiting fMRI adaptation, they found that after adaptation to non-symbolic numerosities, deviant digits led to a strong recovery of signal in the left parietal cortex. Conversely, after adaptation to Arabic digits, non-symbolic deviants did not lead to a recovery of signal in the left parietal cortex. In the right parietal cortex, they found symmetric recovery-effects across presentation notations (Piazza et al., 2007). These results suggest complete notation independence for the right parietal cortex and partial for the left. In contrast, Cohen Kadosh et al. (2007) reported symmetric recovery effects in the left parietal cortex, but none in the right. In short, both studies (Cohen Kadosh et al., 2007; Piazza et al., 2007) support hemispheric asymmetry of notation independence, but with opposite lateralization. On the other hand, using dot-arrays, an earlier fMRI study failed to find a numerosity specific representation in the parietal cortex (Shuman and Kanwisher, 2004). Specifically, Shuman and Kanwisher (2004) demonstrated that 1) a numerical comparison task did not induce higher BOLD response in the parietal cortex compared to a non-numerical comparison task, 2) there were no adaptation effects for numerosity repetition in the parietal cortex, and 3) the difficulty-related BOLD increase in the parietal cortex was not higher for numerical tasks compared to non-numerical tasks. Furthermore, Cohen Kadosh et al. (2011) found greater recovery in parietal BOLD response for a notation change (e.g., dots to digits) compared to a magnitude change. Two recent studies also found qualitatively different parietal representations of numerosity in dot arrays and Arabic digits (Bulthé et al., 2014; Lyons et al., 2014). Taken together, the role of parietal cortices in representing symbolic and non-symbolic numerical information remains controversial.

The idea of an abstract number system ('number sense') stipulates the same mechanism for the representation of numerical information that was extracted from simultaneous (dot-array) and sequential (series of dots) stimuli. In other words, an abstract number sense should be format-independent. Results from neurophysiology support format-independency. Nieder et al. (2006) found neurons in primate region VIP that selectively responded to numerosities from one to four, whether presented simultaneously (dot-array) or sequentially (series of dots). However, during the sample period only 2 out of 228 recorded neurons (~1%) were tuned to the same numerosity for both simultaneous and sequential stimuli (e.g., responding maximally to both a series of three dots and an array of three dots). In humans, the neural underpinnings of simultaneous versus sequential visual numerosities have not been investigated in detail.

Although a number of studies suggest at least partial notation and format independence, the sensory modality independence of numerical magnitude remains unclear. Eger et al. (2003) found overlapping BOLD response in the parietal cortex during auditory and visual symbolic number tasks. Piazza et al. (2006) found overlapping BOLD response in the right IPS during estimation of auditory (beeps) and visual (dots) sequential non-symbolic stimuli. However, overlapping BOLD response does not necessarily mean that the auditory and visual numerical information converge onto the same neural circuitry. In the macaque brain, Nieder (2012) found neurons that coded numerosities from one to four in a sensory-modality independent (i.e., supramodal) fashion. In humans, although a supramodal number sense is thought to reside in the hIPS (Dehaene et al., 2004), the neural underpinning of supramodal numerosity representation has also not been investigated in detail.

Given the lack of unanimous evidence for supramodal numerosity representation in the human parietal cortex, we aimed at studying the neural representation of auditory and visual numerosities using an event-related fMRI paradigm. We conducted multivariate analyses using machine-learning methods (support vector classification-SVC and support vector regression-SVR) that allowed us to investigate whether common representation for auditory and visual numerosities existed in selected regions of the brain.

Moreover, we designed our paradigm such that numerosity estimation and comparison could be separated. Similar bilateral IPS regions are activated for numerical tasks and response-selection (Eliassen et al., 2003; Göbel et al., 2004; Schumacher and Jiang, 2003). Interestingly, Göbel et al. (2004) found no increase in BOLD response in IPS during a numerical comparison task after they controlled for response-selection and reaction time. To overcome a potential confusion between task-related processes and numerosity estimation, we presented auditory and visual numerosities (5, 7, 11, 16) either as series of beeps or dots and only asked participants to make a comparison on 20% of trials (see the Materials and methods section). This design enabled us to keep participants attentive during the whole experiment, assess performance, separate numerosity estimation from task related processing, and optimize the number of estimation-only trials.

Materials and methods

Participants

14 healthy right-handed participants underwent fMRI scanning after giving written informed consent (4 males; mean age = 26.3 ± 6.29 years). They were recruited using a Humboldt University database. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric illnesses. The study was approved by the Berlin Center for Advanced Neuroimaging (BCAN, Nr. 112 and 117), and the Ethical committee of Humboldt Universität zu Berlin. Participants were reimbursed 24 € for their participation.

Stimuli and procedure

Participants engaged in a non-symbolic numerosity-processing task. The non-symbolic numerosities were presented either visually (series of dots) or auditorily (series of beeps). Four numerosities (5, 7, 11, and 16) outside the subitizing range were chosen. They had approximately equal distances from each other on logarithmic scale. Non-numerical sensory features of stimulus sequences were balanced using four different stimulus sets. Single dot/beep duration and total duration increased with numerosity in set 1 and decreased with numerosity in set 4. The interval between single dots/beeps (ISI) increased with numerosity in set 2 and decreased with numerosity in set 3. Frequency (numerosity divided by total duration) increased with numerosity in sets 3 and 4 and decreased with numerosity in sets 1 and 2. This way, we ensured that participants could not rely on a single sensory cue (i.e., duration, frequency, or ISI) to extract numerosity information. To prevent counting, the majority of individual beeps and dots lasted less than 270 ms. Only in set 4 did we use dot/beep durations longer than 270 ms as well. Otherwise it was not possible to have a set of trials where total duration decreased with numerosity. This threshold is consistent with previous studies showing that participants cannot rely on verbal strategies (e.g., counting) within that period (e.g., Piazza et al., 2006; Tokita and Ishiguchi, 2011). We introduced random jitters within the series of dots/beeps to prevent numerosity perception based on periodicity. The length of the jitter depended on the single dot/beep duration for that trial. It was calculated such that after the subtraction of the jitter, the duration of that dot/beep was 40 ms (i.e., jitter = [dot / beep duration − 40 ms]). This way, we made sure that 1) by the subtraction of the jitter, the single dot/beep did not become incomprehensible (i.e., too short to be perceived) and 2) when the duration of a single dot/beep was longer than 270 ms, subjects could not reliably use a counting strategy because they would miss stimuli that were too fast to count. The number of jitters inserted also increased with numerosity to keep periodicity constant. Supplementary Table S1 reports the average, minimum and maximum single dot/beep durations and total stimulus duration for each set and numerosity.

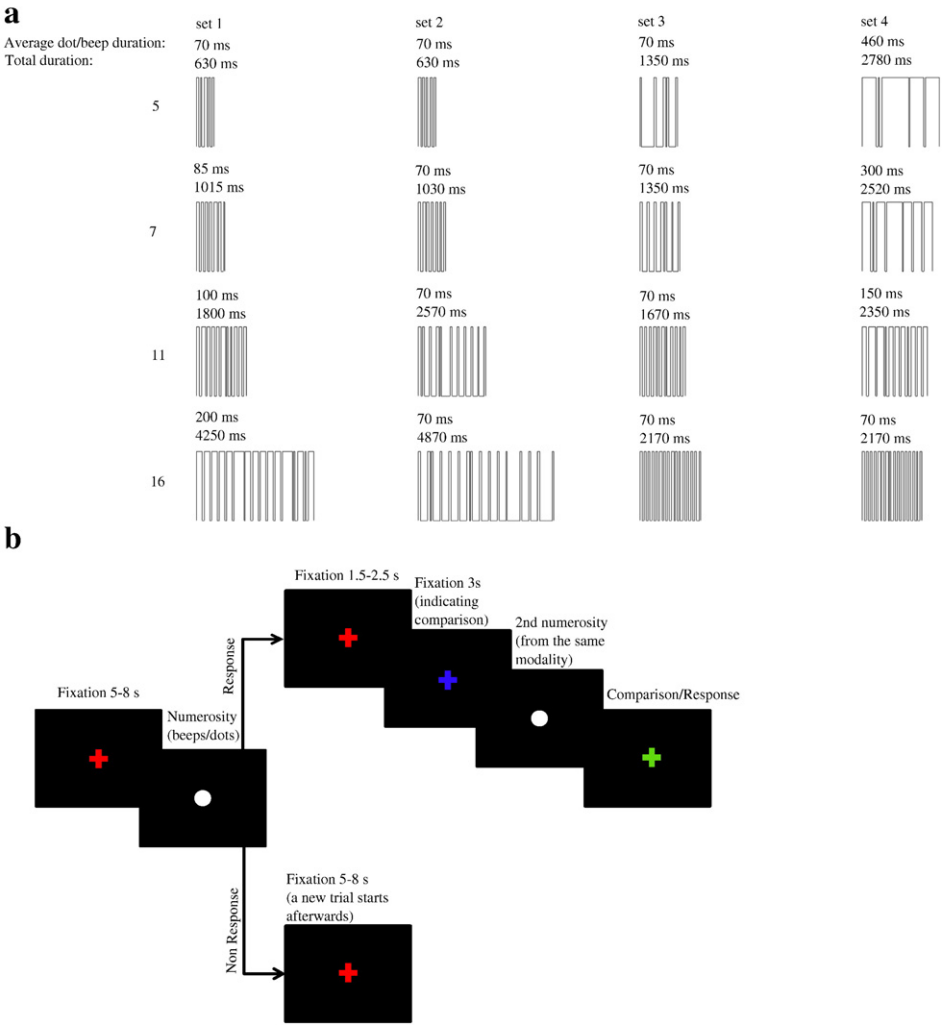

Fig. 1. a) Schematic illustration of timing properties of numerosities used in the experiment. Total duration, single pulse duration, ISI and frequency of the stimuli has been varied in 4 different stimulus sets. b) Schematic depiction of a response (top) and non-response trial (bottom). In response trials (top), after the visual (series of dots) or auditory (series of beeps) presentation of the first numerosity, the color of the fixation point changed to blue, indicating that a comparison of the previous numerosity with the upcoming one would be required. In non-response trials (bottom), after the presentation of the first numerosity, color of the fixation point remained red for 5–8 s after which a new trial started.

Fig. 1a graphically illustrates the time series for the numerosities in each set. Identical timing parameters were used for auditory and visual numerosities.

There were two different task conditions. In 20% of the trials, two beep or dot sequences were presented and participants had to decide which of the two successive stimulus sequences (from the same sensory modality) was numerically larger and respond using a button box (response trials), whereas in the remaining 80% of the trials, only one sequence was presented and no response was required (non-response trials). Participants held the button box using two hands throughout the experiment and responded with the left thumb for the left button press and with the right thumb for the right button press. They were instructed to press the left button if they thought the first numerosity was larger and the right button if they thought the second numerosity was larger. The order of response and non-response trials as well as visual and auditory trials were randomized (i.e., a complete randomization was employed for all the conditions within a block). When viewing the first numerosity, participants did not know whether there would be a second numerosity for comparison. A change in the color of the fixation point from red to blue between 1.5 and 2.5 s after the first numerosity indicated whether or not participants had to make a numerical comparison with the upcoming numerosity. This allowed us to separate response-related processes from mere numerosity perception, keep participants attentive during the whole task, and assess estimation accuracy. The duration of the fixation point between trials ranged from 5 to 8 s and was determined randomly but balanced within each block and across blocks and participants. The blue fixation point after the first numerosity in response trials lasted for 2 s and after that, the second numerosity (probe numerosity) from the same modality was presented. The red fixation point after the first numerosity in non-response trials lasted for 5-8 s, after which a new trial began. The probe numerosity was 25% smaller or larger than the first numerosity (e.g., 16 vs. 20) and the amount of smaller and larger decisions was balanced within participants in each block. The experiment consisted of 8 blocks each containing 40 trials (32 non-response trials and 8 response trials) making a total of 320 trials. Each block took approximately 8 min and the session lasted around 64 min in total. Each numerosity was presented 4 times per block in each modality (auditory and visual) and was drawn from a different stimulus set each time. The stimuli were presented with Matlab using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Fig. 1b illustrates the trial structure and the design of the experiment.

Psychophysical task

A subgroup of participants (10 participants) was invited for a psychophysical session outside the scanner. In this session participants compared sequentially presented auditory or visual numerosities (see

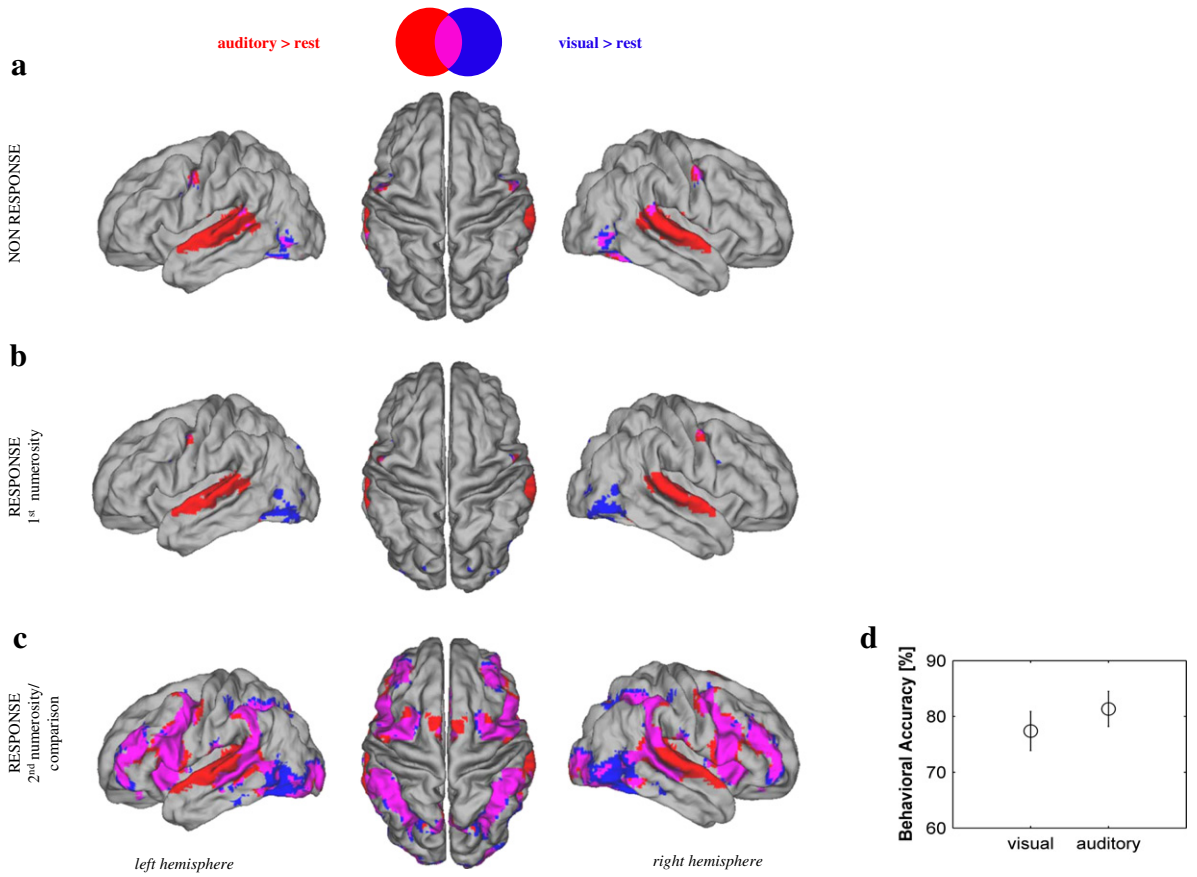

Fig. 2. a) The parametric modulation of BOLD response with increasing numerosity for the numerosity-GLM, for auditory (red) and visual numerosities (blue). Auditory and frontal cortices were activated for auditory numerosities and V5 and frontal cortex were activated for visual numerosities (FDR corrected at $p < .05$ on cluster level). b) The parametric modulation of the BOLD response from the first numerosity (in response trials) > rest contrast, for auditory (red) and visual (blue) modality. c) The parametric modulation of the BOLD response from the probe numerosity (second numerosity in response trials) > rest contrast, for auditory (red) and visual (blue) modality. In addition to sensory cortices, parietal and frontal areas exhibit significant increase in the BOLD response (FDR corrected at $p < .05$ on cluster level). d) The accuracy of participants in comparing the numerosities in response trials; separately for auditory and visual modalities.

Supplementary Table S2 for details). The psychophysical task served as another control to check that participants were actually doing numerosity comparison and not relying on non-numerical features. The details of this task are given in the Supplementary materials section.

Table 1
Peak coordinates and anatomical labels of brain regions in which the BOLD response increases parametrically with numerosity in non-response trials.

Modality	Brain region	Z-score	Peak coordinates (MNI)		
			x	y	z
Auditory	R. superior temporal gyrus	6.37	52	−15	3
AN > R		5.73	70	−20	8
		5.59	55	−35	8
	L. superior temporal gyrus	6.07	−48	−17	3
	L. inferior frontal gyrus	6.00	−48	−25	10
	L. superior temporal gyrus	5.55	−51	−22	5
	R. precentral gyrus	4.57	55	1	48
Visual	R. inferior occipital gyrus	4.73	47	−67	−5
VN > R		4.51	42	−67	−13
	R. inferior temporal gyrus	4.34	45	−50	−13
	L. inferior temporal gyrus	4.49	−43	−65	−8
	L. inferior occipital gyrus	4.46	−51	−77	−5
	L. middle occipital gyrus	4.46	−46	−70	3
	R. precentral gyrus	4.22	52	3	45
	L. precentral gyrus	4.13	−48	−7	45
		3.90	−53	1	43

R — right hemisphere, L — left hemisphere; FDR corrected at $p < .05$ on cluster level.

Localizer task

To independently determine ROIs for multivariate analysis, a short (7.8 min) functional localizer (adapted from Pinel et al., 2007) was created using OpenSesame (Mathôt et al., 2012) and presented after the numerical task. The reading, calculation and memory conditions were presented using an optimized rapid event related design. A description of the optimization is given in the Supplementary materials section. Ten simple sentences ("Bears are fond of salmon and honey"), subtraction problem sentences ("Calculate eleven minus five") (translated from Pinel et al., 2007) and novel date recall sentences ("The date of New Year's Eve is ____"), were intermixed with ten rest periods, for a total of 40 trials. In all three conditions, subjects were instructed to silently read the sentences and mentally generate an answer (subtraction and date recall), although no response was made. The words were presented on four consecutive screens (250 ms/screen), with one word or phrase (e.g., "the rain", "New Year's Eve", "are fond of") per screen, followed by a blank screen (screens 1–3, 100 ms; screen 4, 900 ms). In the rest condition, a blank screen with a central fixation dot was presented. During the rest condition, participants were told to look at the screen with their eyes open. All trials were followed by a temporally jittered fixation dot.

The saccades, object grasping, and color naming blocks were presented using an optimized epoch design. Black and white illustrations of graspable objects (e.g., pencil, corkscrew; courtesy of Philippe Pinel), multidirectional (360°) saccade targets and color photographs of houses with different roof colors, were presented. In object grasping trials (7200 ms/trial), three objects were presented on three

Table 2
Peak coordinates and anatomical labels of brain regions in which the BOLD response increases during the presentation of the second numerosity in response trials.

Modality and contrast	Brain region	Z-score	Peak coordinates (MNI)		
			x	y	z
Auditory	L. putamen	6.48	−23	8	8
RAN > R	L. superior temporal gyrus	6.06	−46	−20	0
	R. superior temporal gyrus	6.00	65	−37	23
	L. SMA	6.16	−8	11	45
		5.67	−3	3	60
	R. middle cingulate cortex	5.77	7	16	43
	L. cerebellum	5.34	−31	−57	−30
		5.13	−36	−72	−25
	R. cerebellum	5.19	35	−60	−23
	R. inferior parietal lobule	5.04	45	−40	48
	R. angular gyrus	4.81	37	−55	48
	R. supramarginal gyrus	4.69	55	−35	43
	R. middle frontal gyrus	4.78	42	41	28
		4.51	50	41	23
	L. middle occipital gyrus	4.59	−28	−67	28
	R. inferior temporal gyrus	4.56	55	−57	−13
Visual	R. insula lobe	6.89	35	21	3
RVN > R	L. insula lobe	6.01	−28	18	−8
	R. cerebellum	5.70	42	−62	−25
	R. middle cingulate cortex	6.50	5	16	45
	L. SMA	6.18	−8	11	45
	R. SMA	5.98	10	23	48
	R. angular gyrus	6.06	37	−55	48
	L. inferior parietal lobule	5.83	−48	−35	48
	R. precuneus	5.62	10	−70	43
	L. cingulate gyrus	5.88	−6	−20	30
	BA 23	5.54	5	−22	28
	L. inferior frontal gyrus	5.66	−43	16	28
	L. precentral gyrus	5.37	−38	3	35
	L. middle frontal gyrus	4.83	−33	46	13
	R. middle frontal gyrus	5.61	42	41	28
		4.62	40	53	0
	R. middle orbital gyrus	4.79	27	46	−13

R — right hemisphere, L — left hemisphere, SMA — supplementary motor area; FDR corrected at < .05 on cluster level, k > 10.

consecutive screens (1200 ms/screen), each followed by a blank screen (1200 ms/screen). Subjects were instructed to mentally imagine grasping the objects with their dominant (right) hand. In each 7200 ms trial, three houses with various colored roofs were presented for 1200 ms followed by a blank screen for 1200 ms. Subjects were instructed to silently name the roof color. In saccade trials (6200 ms/trial), three saccades were made by following a saccade target (+). For each saccade, a center target (+) was presented for 200 ms, followed by a 1300 ms saccade target (presented in two 650 ms equally spaced steps from center target), and ending with a 200 ms center target.

Table 3
Peak coordinates of task and localizer based-ROIs used for the multivariate analysis of data.

Contrast	Center of gravity	# voxels	Peak coordinates (MNI)		
			x	y	z
Auditory (aud)	R. Heschls gyrus	77	51	−16	6
	L. superior temporal gyrus	65	−51	−21	7
	R. superior temporal gyrus	47	66	−26	9
Visual (occ)	L. inferior occipital gyrus	116	−44	−69	−5
	R. inferior temporal gyrus	72	45	−68	−8
SPL	R. superior parietal lobule	115	25	−64	53
	L. superior parietal lobule	83	−22	−61	59
hIPS	L. hIP1	162	−38	−46	44
	R. inferior parietal lobule	34	42	−42	44

SPL — superior parietal lobule, hIPS — horizontal segment of intra parietal sulcus, k > 10.

All trials were alternated with jittered fixation trials, with a minimum ISI of 4000 ms.

The localizer was run on 31 participants in total, 14 of which were the participants of the current experiment and 17 (8 males, mean age = 30.1 ± 16.18 years) participated in another experiment. Using the group activity of 31 participants, we determined the regions activated during saccades, grasping, date retrieval and mental subtraction by contrasting each condition with rest (for saccades), reading (for date retrieval and mental subtraction) or houses (for grasping).

fMRI data acquisition

Imaging data were acquired at the Berlin Center for Advanced Neuroimaging (BCAN) with a 3 T Siemens TIM Trio scanner (Siemens, Erlangen) using a 12-channel head coil. Before the experiment, a T1-weighted image (MPRAGE) was collected as high resolution anatomical reference (TR = 1900 ms, TE = 2.52 ms, flip angle = 9°, FOV = 256 mm × 256 mm × 192 mm, resolution = 1 mm). After anatomical scanning, a magnetic field mapping sequence was run to correct for magnetic field inhomogeneities (TR = 400 ms, TE = 5.19 ms/7.65 ms, flip angle = 60°, FOV = 192 mm × 192 mm, resolution = 3 mm, slice gap = 25%, slices = 33). $T2^*$ -weighted gradient-echo echo-planar images were collected during the experiment (TR = 2500 ms, TE = 25 ms, flip angle = 82°, FoV = 190 mm × 190 mm, resolution = 2.5 mm, slices = 42 slices with a 20% distance factor; interleaved acquisition order). Finally, $T2^*$-weighted gradient-echo echo-planar images were collected during the localizer task (TR = 2000 ms, TE = 30 ms, flip angle = 78°, FoV = 192 mm × 192 mm, resolution = 3 mm, slices = 33 with a 25% distance; descending acquisition order).

fMRI data analysis

Images were analyzed using Statistical Parametric Mapping software (SPM8; Wellcome Trust Centre for Neuroimaging; http://www.fil.ion.ucl.ac.uk/spm/). In the univariate analysis of task-related fMRI data, we applied a standard mass-univariate general linear model (GLM). Functional images were first reoriented to the anatomical reference and then corrected for inhomogeneities in the magnetic field. Subsequent preprocessing included slice-timing correction – the middle image in the time series was taken as basis – spatial realignment and unwarping, co-registration to the unwarped mean image, segmentation, normalization to standard Montreal Neurological Institute (MNI) space and smoothing (FWMH = 5 × 5 × 5 mm). After preprocessing, a GLM model based on numerosity was defined using a canonical hemodynamic response function and its first temporal derivative. The numerosity model included a regressor for each numerosity, separately for visual and auditory modalities and response and non-response trials. In response trials, the first and second numerosities were also modeled separately. Thus, the numerosity-GLM had 24 regressors in total along with 6 movement parameters from preprocessing to capture signal variations due to head motion. The event-related numerosity regressors were locked to the onset of the numerosity presentation. The unmodeled part of the data (that corresponded to fixation periods between trials) was used as rest when defining the contrasts in GLMs.

For all multivariate analysis (SVC and SVR), the unsmoothed images of the preprocessed data were used instead of smoothed images to preserve the maximal amount of spatial information. To train the classifier, we extracted one beta estimate per block for each level of the stimuli from the GLM. The multi-class classifier was trained using support vector machines based on LIBSVM 3.12 (Chang and Lin, 2011). We performed grid search to optimize the regularization parameter (C = [0.01, 1, 10, 100, 1000]). A one-block-out cross validation scheme was employed, where one block was left as the test data and the remaining seven blocks were used to train the classifier. The left-out block was iterated over all eight blocks and an average decoding accuracy estimate was obtained at the end. Additionally, SVR was done by fitting a step function (corresponding

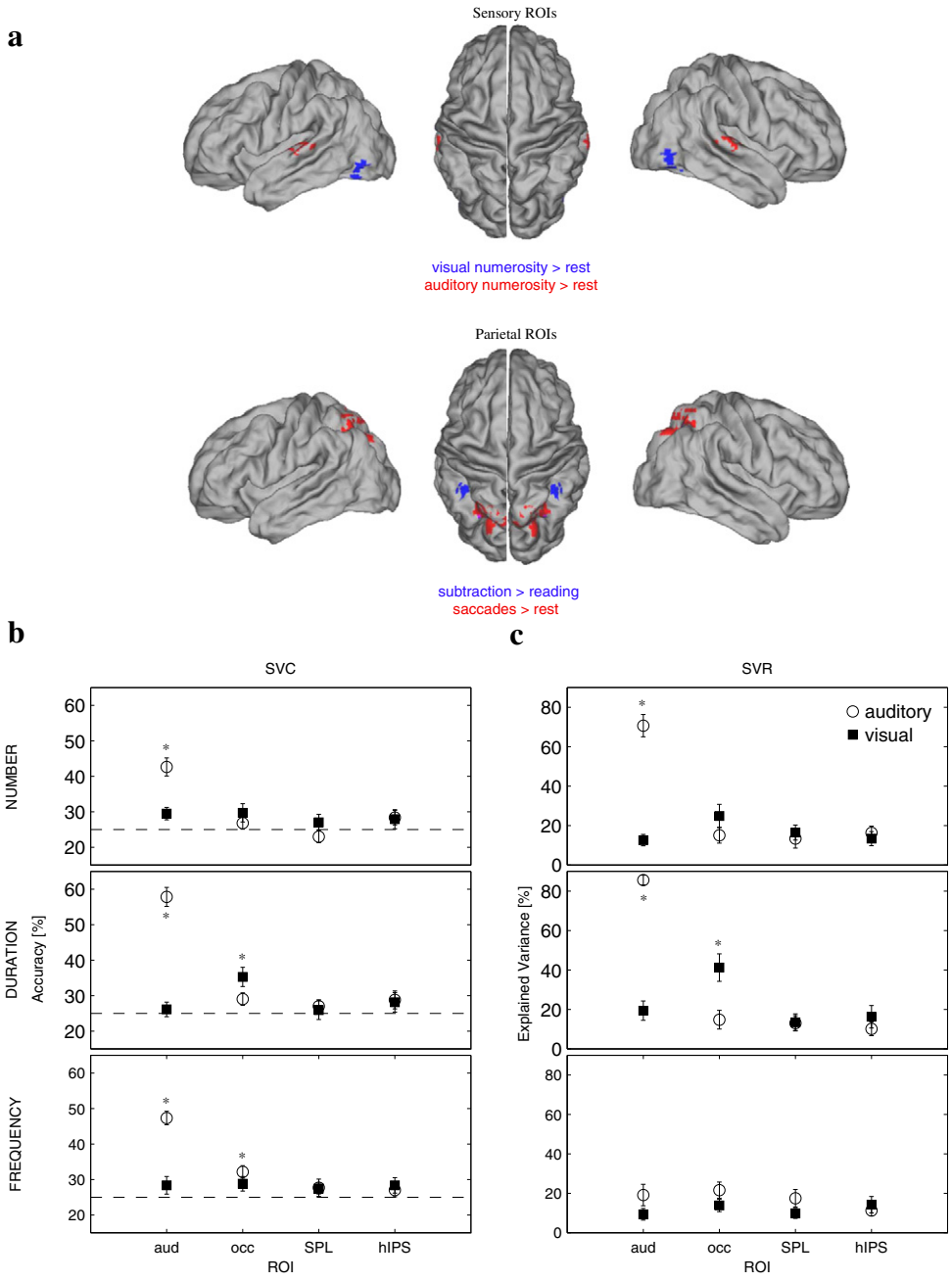

Fig. 3. a) Depiction of four main ROIs used in multivariate analysis. Results for b) support vector classification (SVC) and c) support vector regression (SVR) for numerosity, duration and frequency models. * = $p < .05$, Bonferroni corrected.

to the logarithm of numerosities after mean correction) to the patterns evoked by the four numerosities (separately for auditory and visual modalities), in line with Eger et al. (2009). SVC results were tested for statistical significance using a Student's t-test of the average multi-class classification accuracies against chance level (25%) over all participants. SVR results were tested for statistical significance using a t-test of the correlation coefficients against zero.

Analysis of localizer data and ROI extraction

Preprocessing of the localizer data was identical to the functional data besides the reference slice used in slice-timing correction (first image), the kernel size used during smoothing (FWHM = 6 × 6 × 6 mm) and the order of slice-timing correction and spatial alignment (here, spatial alignment and unwarping was performed before slice-timing correction). After preprocessing, the localizer task was modeled by a canonical hemodynamic response function and its first temporal derivative and a GLM was defined that included a regressor for each condition (namely houses, objects, dates, reading, subtraction, saccades, fixation) and 6 motion parameters from preprocessing to capture signal variations due to head motion. Two of the main ROIs used in the multivariate analysis were defined from the group-level activity in the localizer task. To identify calculation-related parietal ROIs, the subtraction > reading contrast was masked with the hIPS (Jülich Atlas; Eickhoff et al., 2005, 2006, 2007). To identify saccade-related parietal ROIs, the saccades > rest contrast was masked with superior parietal lobule (SPL; WFU_PickAtlas; Maldjian

Table 4
Summary of ROIs used for the exploratory analysis.

Mask	Method	# voxels	Auditory				Visual			
			SVC		SVR		SVC		SVR	
			Accuracy	p	Explained variance	p	Accuracy	p	Explained variance	p
hIPS	Jülich atlas	906	25.45	.80	14.94	.95	27.46	.38	24.72	.37
	Jülich atlas and subtraction > rest (localizer)	187	25.89	.59	13.37	.79	26.12	.71	14.74	.83
hIPS1	Jülich atlas	397	24.33	.66	13.41	.74	26.56	.57	20.39	.84
	Jülich atlas and subtraction > rest (localizer)	29	24.11	.72	18.48	.31	22.99	.37	13.90	.15
hIPS2	Jülich atlas	198	26.34	.63	14.97	.91	24.78	.94	12.28	.32
hIPS3	Jülich atlas	334	28.79	.16	16.68	.49	25.67	.84	24.93	.72
	Jülich atlas and subtraction > rest (localizer)	16	29.24	.06	10.79	.23	24.78	.92	8.12	.23
rhIPS	Jülich atlas	363	24.33	.71	15.95	.97	25.45	.82	32.15	.34
	Jülich atlas and subtraction > rest (localizer)	52	26.79	.29	11.72	.72	21.43	.11	9.56	.17
rhIPS1	Jülich atlas	130	18.75	< .01	18.42	.46	25.67	.73	15.63	.59
rhIPS2	Jülich atlas	69	24.33	.77	15.76	.73	24.55	.86	10.84	.69
rhIPS3	Jülich atlas	173	27.46	.39	19.55	.52	25.89	.73	25.36	.85
lhIPS	Jülich atlas	543	26.79	.19	14.48	.95	24.78	.92	13.73	.86
	Jülich atlas and subtraction > rest (localizer)	135	27.01	.23	9.87	.93	26.34	.65	14.41	.59
lhIPS1	Jülich atlas	267	28.35	.05	16.45	.79	22.54	.27	14.00	.85
	Jülich atlas and subtraction > rest (localizer)	29	24.11	.72	18.48	.31	22.99	.37	13.09	.15
lhIPS2	Jülich atlas	129	28.57	.12	14.69	.83	23.44	.60	14.08	.99
	Jülich atlas and subtraction > rest (localizer)	17	29.24	.11	14.96	.34	26.34	.66	15.78	.36
lhIPS3	Jülich atlas	159	26.12	.71	22.41	.32	26.34	.70	16.02	.37
V5	Jülich atlas and visual numerosity > rest (functional)	350	22.09	.33	8.91	.98	29.02	.13	**19.11**	**.02**
Grasping	Objects > houses (localizer) and parietal cortex (WFU_Pickatlas)	128	24.33	.79	11.58	.69	23.88	.40	30.77	.24
F-contrast	F-contrast (functional) and parietal cortex (WFU_Pickatlas)	1000	27.68	.29	21.02	.34	25.67	.79	13.81	.45
Occipital	Houses > rest (localizer) and occipital cortex (WFU_Pickatlas)	200	**30.58**	**.02**	20.97	.11	**30.58**	**.03**	**29.47**	**.01**
Probe Par.	Response trials > rest and parietal cortex (WFU_Pickatlas)	1000	28.35	.11	18.16	.02	25	1	20.77	.70
Frontal	Auditory and visual > rest and frontal cortex (WFU_Pickatlas)	65	28.13	.19	**18.98**	**.04**	24.33	.74	15.59	.44

rhIPS — right horizontal intraparietal sulcus, lhIPS — left horizontal intraparietal sulcus, bold print indicates $p < .05$.

et al., 2003, 2004). The two other ROIs used for the main multivariate analysis were created from the functional activity during the task by contrasting all numerosities against rest, separately for auditory and visual modalities. Note that these contrasts are orthogonal to the decoding dimension (i.e., numerosity) in that they do not differentially take into account numerosity. In other words, they are not parametric

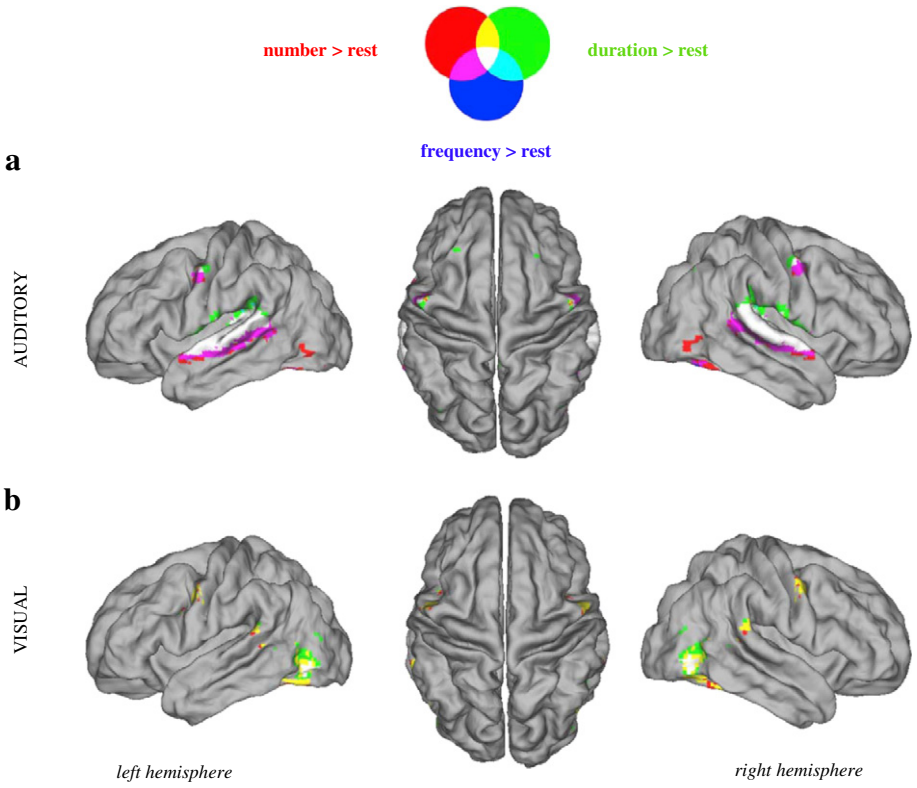

Fig. 4. The parametric modulation of the BOLD response with increasing numerosity (red), duration (green) and frequency (blue) for a) auditory and b) visual stimuli during non-response trials (FDR corrected at $p < .05$ on cluster level).

Table 5
Peak coordinates and anatomical labels of brain regions in which the BOLD response increases parametrically with duration in non-response trials.

Modality and contrast	Brain regions	Z-score	Peak coordinates (MNI)		
			x	y	z
Auditory	R. superior temporal gyrus	6.39	52	−17	3
AD > R		5.89	65	−32	5
		5.79	67	−20	8
	L. superior temporal gyrus	6.14	−48	−25	10
		5.94	−48	−17	3
		5.48	−51	−22	5
Visual	R. inferior temporal gyrus	4.53	47	−67	−3
VD > R		4.29	47	−50	−13
	R. middle temporal gyrus	4.11	40	−67	8
	L. middle occipital gyrus	4.40	−46	−70	3
	L. inferior occipital gyrus	4.05	−41	−67	−5
	Sub-gyral	4.04	−43	−55	−5
	L. middle occipital gyrus	4.32	−23	−90	10
	R. precentral gyrus	4.11	52	3	40
	L. fusiform gyrus	3.88	−41	−55	−20
	L. precentral gyrus	3.84	−48	−5	45

R — right hemisphere, L — left hemisphere; auditory contrast FDR corrected at $p < .05$ while visual is uncorrected at $p < .001$, $k > 10$.

contrasts that would detect changes of BOLD as a function of numerosity. The activity during auditory numerosity presentation was masked with the temporal cortex (WFU_PickAtlas). The activity during visual numerosity presentation was masked with occipital cortex (WFU_PickAtlas). A significance level of $p < .001$ (uncorrected) was used for each contrast and the 200 most active voxels in each of these ROIs were used for multivariate analysis (Fig. 3a). The decision about voxel size was based on the ROI with the smallest voxel size (SPL).

Exploratory analysis

Extra ROIs

The multivariate analysis was run on additional ROIs for the exploratory analysis because we were conservative about the total voxel size during the selection of four main ROIs. That is, the size of the main ROIs was based on the ROI with smallest number of voxels (SPL, 200 voxels) to make sure that decoding performance is not biased by the number of features (voxels). We chose extra ROIs mainly from regions that were found to have relevance to numerical tasks in previous studies. Specifically, we chose the parietal and frontal ROIs, as these areas have been repeatedly activated in previous numerical studies (Ansari et al., 2006; Arsalidou and Taylor, 2011; Bulthé et al., 2014; Cohen Kadosh and Walsh, 2009; Cohen Kadosh et al., 2007, 2011; Damarla and Just, 2013; Eger et al., 2003, 2009; Harvey et al., 2013; Hubbard et al., 2005; Nieder et al., 2006; Piazza et al., 2006, 2007). We chose a grasping related ROI because embodied cognition suggests that we encode information using the sensory and motor representations that were employed during acquisition of that information; hence, the representation of numerosities might share a neural basis with hand movements (Andres et al., 2012; Fischer, 2012). Also, neurons in the anterior IPS have been found to code for the number of arm movements in monkeys (Sawamura et al., 2002) and to be susceptible to numerical–spatial interference effects (Koten et al., 2011). Finally, we chose V5 because this area showed task related activity and an ROI in occipital cortex to test the role of sensory cortices in numerosity perception. These extra ROIs were created from activity from the localizer or functional tasks, anatomical labeling using the Jülich atlas (Eickhoff et al., 2005, 2006, 2007) or WFU_PickAtlas (Maldjian et al., 2003, 2004), or task-related activity masked with anatomy (Table 4).

Table 6
Peak coordinates and anatomical labels of brain regions in which the BOLD response increases parametrically with frequency in non-response trials.

Modality	Brain region	Z-score	Peak coordinates (MNI)		
			x	y	z
Auditory	R. superior temporal gyrus	6.32	52	−15	3
AF > R		5.92	70	−20	8
		5.51	62	−27	10
	L. superior temporal gyrus	5.97	−48	−17	3
		5.82	−48	−25	10
		5.57	−61	−22	5
Visual	R. inferior temporal gyrus	4.80	47	−67	−3
VF > R	R. inferior occipital gyrus	4.17	42	−65	−13
	R. middle temporal gyrus	3.95	47	−57	8
	L. inferior occipital gyrus	4.67	−51	−77	−5
		4.33	−41	−67	−5
	L. middle occipital gyrus	4.45	−46	−70	3
	R. superior temporal gyrus	4.12	65	−35	18
	R. precentral gyrus	4.00	52	3	40
	L. superior temporal gyrus	3.75	−46	−40	23
	L. precentral gyrus	3.72	−53	−2	43
	L. postcentral gyrus	3.13	−53	−7	50

R — right hemisphere, L — left hemisphere; FDR corrected at $p < .05$ on cluster level, $k > 10$.

Non-numerical stimulus dimensions

To see to what extent the observed BOLD response and decoding results in the numerosity model were due to sensory features, we created two additional GLM models that were based on the duration and the frequency of the stimuli rather than numerosity. For the duration model, we re-clustered the trials into four categories C1–C4 (C1: duration (dur) < 1.5 s; C2: 1.5 s < dur < 2.5 s; C3: 2.5 s < dur < 3.5 s; C4: dur > 3.5 s) based on the total duration of numerosity presentation in non-response trials. Each duration category was modeled by a different regressor, separately for each modality. The first and second numerosities in response-trials were also modeled separately with one regressor for each modality. Onset of stimulus presentation was used to define the regressors. Hence, in the duration GLM, there were 12 regressors along with 6 movement parameters from preprocessing to capture signal variation due to head motion.

For the frequency model, we re-clustered the trials into four categories C1–C4 (C1: frequency (freq) < 3.42 Hz; C2: 3.42 Hz < freq < 4.6 Hz; C3: 4.6 Hz < freq < 5.6 Hz; C4: freq > 5.6 Hz) based on frequency of stimulus presentation in non-response trials. Each frequency category was modeled by a different regressor, separately for each modality. The first and second numerosities in response-trials were also modeled separately with one regressor for each modality. Onset of stimulus presentation was used to define the regressors. Hence, in frequency-GLM, there were 12 regressors along with 6 movement parameters from preprocessing to capture signal variation due to head motion.

SVC and SVR were applied to the duration and frequency models in almost the same way as the numerosity model. In duration and frequency models, SVR was done by fitting a linear step function (rather than logarithmic) to the patterns evoked by the four stimulus categories.

Results

Behavioral results (fMRI task)

One participant performed significantly below chance level (33% accuracy in visual response-trials and 28% accuracy in auditory response-trials) and was excluded from behavioral analysis. Therefore, all subsequent behavioral analyses were based on the data from 13 participants. The overall mean accuracy (standard deviation) was 77.57% (11.54%) for visual trials and 81.35% (10.92%) for auditory trials. The error rates were analyzed after arc-sine transformation. There was no significant difference in accuracy between visual and auditory response trials ($t(12) = 1.14$, $p = .28$) and accuracy was significantly above chance for both ($t(12) = 5.87$, $p < .001$ for visual; $t(12) = 7.25$, $p < .001$ for auditory). To explore whether responses were based on non-numerical features of the stimuli (duration and frequency) rather

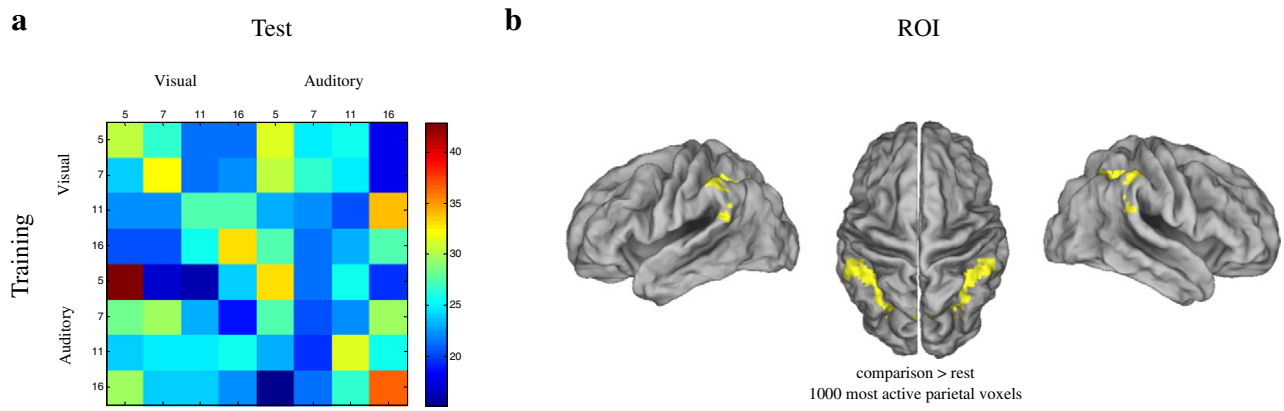

Fig. 5. a) Confusion matrix for visual and auditory response trials, within-modality and cross-modality decoding combined. b) The ROI used in this analysis that covers 1000 most active parietal voxels during response phase.

than numerosity, we divided response-trials according to congruency between numerical and non-numerical features. When the total duration was congruent with numerosity, the accuracy was 73.28% (15.13%) for visual and 87.22% (13.28%) for auditory trials. When the duration was incongruent with numerosity, the accuracy was 91.47% (10.94%) for visual and 71.53% (17.65%) for auditory trials. When the frequency was congruent with numerosity, the accuracy was 72.69% (15.42%) for visual and 88.81% (12.59%) for auditory trials. And finally, when the frequency was incongruent with numerosity, the accuracy was 82.39% (12.67%) for visual and 76.30% (12.26%) for auditory trials. If responses were based on non-numerical features (duration or frequency), we would expect below-chance performance on incongruent trials (e.g., when the larger numerosity had shorter duration or smaller frequency). However, on incongruent trials participants performed significantly above chance for both visual and auditory numerosity comparison, ($t(12) = 235.71$, $p < .001$ for incongruent duration in auditory modality; $t(12) = 359.89$, $p < .001$ for incongruent frequency in auditory modality; $t(12) = 419.10$, $p < .001$ for incongruent duration in visual modality; $t(12) = 359.79$, $p < .001$ for incongruent frequency in visual modality). Interestingly, in 2 (modality: visual vs. auditory) × 2 (sensory feature: duration vs. frequency) × 2 (congruency: congruent vs. incongruent) repeated measures ANOVA, there was an interaction between modality (auditory or visual) and congruency (congruent or incongruent; $F(1,12) = 16.88$, $p = .001$, $\eta_p^2 = .58$), between modality and sensory feature (duration or frequency; $F(1,12) = 23.49$, $p < .001$, $\eta_p^2 = .66$), and between modality, sensory feature and congruency ($F(1,12) = 6.15$, $p = .03$, $\eta_p^2 = .34$). This may imply that the extent to which participants relied on frequency or duration during numerosity estimation was differentially influenced by modality. Yet, overall these results do not indicate a consistent use of non-numerical stimulus dimensions throughout the entire task, but rather suggest that participants mainly based their responses on numerosity estimates.

Behavioral results (psychophysical task)

In the psychophysical task, accuracy was high for both modalities. When plotted against log numerosity, accuracy followed the well-known sigmoid function with no apparent differences between modalities (mean = 81.34% for auditory and 81.11% for visual; $t(10) = 0.13$, $p = .89$; Supplementary Fig. S1). The psychometric parameters (point of subjective equality-PSE and Weber fraction-WF) did not differ between visual and auditory modalities (mean PSE = 9.92 for auditory and 10.12 for visual, $t(10) = 1.03$, $p = .33$; mean WF = .1267 for auditory and .1258 for visual, $t(10) = 0.35$, $p = .73$). The performance of subjects in the psychophysical task further indicates that subjects were able to estimate numerosity, without depending on a single sensory feature.

Post-hoc analysis of dot and gap durations

Due to the refresh rate of the projector (60 Hz), there might have been a mismatch between the planned and experimentally achieved timing in visual stimuli. In order to check the degree of this mismatch, durations of each dot as well as gaps between dots were analyzed post-hoc. Specifically, the correlation between planned and experimentally achieved mean dot/gap duration was calculated for each numerosity. We observed that the correlation was highly significant for each numerosity both for single-dot and gap durations ($r > .9$, $p < .001$ for all). We further analyzed the minimum experimental gap between two consecutive dots and observed that the shortest experimental gap between two consecutive dots was never <50 ms for numerosities 5,7, and 11 and it got shorter than that only for numerosity 16 for 16 of 448 trials. However, even in those 16 cases two subsequent dots were separated by at least one blank frame (i.e., 16.67 ms). Despite an increased variability, behavioral results did not reveal any statistically significant performance differences between modalities. Neither did the mean overall duration of the stimulus sets differ between modalities nor did we observe any merging of subsequent dot presentations even at fastest stimulus presentation rate. Taken together, limited timing resolution of the visual stimuli was not suited to compromise the overall results of the present study.

fMRI results

We included all 14 participants in the analysis of the fMRI data. Results did not change when we excluded the participant with low accuracy from the multivariate analyses.

Numerosity activation

The numerosity-related BOLD response was captured by parametrically contrasting auditory numerosity against rest (AN > R) and visual numerosity against rest (VN > R) in non-response trials. Response trials were analyzed, separately for both modalities, by parametrically contrasting the second auditory numerosity against rest (RAN > R) and the second visual numerosity against rest (RVN > R). When viewing the second numerosity, participants knew that they would have to compare it to the first numerosity. Interestingly, pure numerosity perception (non-response trials and 1st numerosity in response trials; Figs. 2a, b, Table 1), did not increase the BOLD response in the parietal cortex. Parietal activity was observed only during the comparison in response trials (Fig. 2c, Table 2; FDR corrected at $p < .05$ on cluster level). In non-response trials, sensory areas (auditory cortex for auditory numerosities, occipital cortex for visual numerosities), along with a small locus in

frontal cortex (BA6), showed prominent increase of the BOLD response (Fig. 2a, Table 1).

Decoding numerosity

We ran multivariate pattern analysis on four ROIs obtained from either the localizer activity or task-related activity (see the Materials and methods section) to investigate whether these carried numerosity selective activation. The ROIs covered the auditory cortex (200 most active voxels during auditory numerosity presentation in non-response trials (all numerosities > rest)), occipital cortex (200 most active voxels during visual numerosity presentation in non-response trials (all numerosities > rest)), SPL (200 most active voxels during saccade formation in the localizer task) and hIPS (200 most active parietal voxels during subtraction in the localizer task, see Table 3 and Fig. 3a for details).

We conducted a two-way repeated-measures (ROI × modality) ANOVA on accuracy values of classification. There was a significant main effect of ROI ($F(1,13) = 13.14$, $p < .001$, $\eta_p^2 = .50$) and a significant interaction between ROI and modality ($F(1,13) = 4.51$, $p = .008$, $\eta_p^2 = .26$). Post-hoc tests revealed that the decoding accuracy for auditory numerosity was significantly above chance (>25%) in the auditory cortex ROI (42.63%, $t(13) = 1.05$, $p < .001$; Bonferroni corrected). We failed to decode visual or auditory numerosity in parietal ROIs (all ps > .05). In none of the ROIs we were able to decode numerosity from both modalities, including parietal ROIs. Therefore, no attempt was made to generalize numerosity decoding from one modality to the other.

As numerosity is not merely categorically organized but also inherently ordinal in nature, we applied SVR to test for linearly ordered decoding performance. SVR results were in line with the SVC results. A two-way repeated measures (ROI × modality) ANOVA on explained variance [%] values revealed a main effect of ROI ($F(13) = 15.19$, $p < .001$, $\eta_p^2 = .54$) and modality ($F(13) = 15.07$, $p = .002$, $\eta_p^2 = .54$) and an interaction between modality and ROI ($F(13) = 30.97$, $p < .001$, $\eta_p^2 = .70$). Post-hoc tests demonstrated that the highest amount of variance explained by numerosity was in the auditory cortex for auditory stimuli (70.67%, $t(13) = 20.07$, $p < .001$; Bonferroni corrected; Fig. 3c).

Exploratory results

Decoding numerosity in additional ROIs

The four main ROIs reported above yielded no hint of numerosity-related encoding in the parietal cortex, possibly because we were conservative in determining ROI size. Therefore, we explored twenty-five additional ROIs (Table 4). These additional ROIs varied in size and were based on localizer/task activity, anatomical regions (Jülich atlas or WFU_PickAtlas), or a combination of both. Only one of the additional ROIs, in occipital cortex, yielded partially consistent above-chance numerosity classification in non-response trials (Table 4; houses > rest). None of the other ROIs, including parietal ROIs, yielded better-than-chance classification of numerosity in non-response trials.

Duration and frequency activation

The increase in the BOLD response that correlated with duration was captured by parametrically contrasting auditory duration against rest (AD > R) and visual duration against rest (VD > R). Sensory areas and a circumscribed locus in the frontal cortex (BA6) showed significant increase (Fig. 4, Table 5). The increase in the BOLD response that correlated with frequency was captured by parametrically contrasting auditory frequency against rest (AF > R) and visual frequency against rest (VF > R). Again, sensory areas and the locus in the frontal cortex (BA6) showed significant increase (Fig. 4, Table 6). The BOLD response increase for duration and frequency mostly overlapped (Fig. 4).

Decoding duration and frequency

Because duration and frequency can provide a proxy for numerosity (numerosity = duration × frequency), we also conducted multivariate analyses (SVC and SVR) on duration and frequency models. These analyses helped determine whether regions that code numerosity also code sensory features (i.e., the specificity of numerosity coding).

We conducted a two-way repeated-measures (ROI × modality) ANOVA on accuracy values of classification separately for duration and frequency models. In the duration model, there was a significant main effect of ROI ($F(1,13) = 17.10$, $p < .001$, $\eta_p^2 = .57$) and modality ($F(1.13) = 10.16$, $p = .007$, $\eta_p^2 = .44$) and a significant interaction between ROI and modality ($F(1,13) = 31.86$, $p < .001$, $\eta_p^2 = .71$). Post-hoc tests revealed that the decoding accuracy for duration was significantly above chance (>25%) in the auditory cortex ROI for auditory stimuli (57.81%, $t(13) = 12.26$, $p < .001$; Bonferroni corrected) and in the visual cortex ROI for visual stimuli (35.27%, $t(13) = 3.79$, $p = .002$; Bonferroni corrected). In the frequency model, there was again a significant main effect of ROI ($F(1,13) = 13.57$, $p < .001$, $\eta_p^2 = .51$) and modality ($F(1.13) = 10.29$, $p = .007$, $\eta_p^2 = .44$) and a significant interaction between ROI and modality ($F(1,13) = 11.33$, $p < .001$, $\eta_p^2 = .47$). Post-hoc tests revealed that the decoding accuracy for frequency was significantly above chance (>25%) in auditory and visual cortex ROIs for auditory stimuli (auditory cortex: 47.32%, $t(13) = 11.71$, $p < .001$; visual cortex: 32.14%, $t(13) = 4.09$, $p = .001$; Bonferroni corrected). The decoding accuracy was not significant in the visual cortex ROI for visual stimuli (28.79%, $t(13) = 1.85$, $p = .08$; Bonferroni corrected). However, when we increased the size of the visual cortex ROI to 350 voxels, the decoding accuracy in the visual cortex for visual stimuli also reached significance (31.69%, $t(13) = 3.24$, $p = .007$). We failed to decode duration or frequency in parietal ROIs for both modalities (all ps > .05; Fig. 3b).

We continued with SVR to confirm the SVC results. In the duration model, a two-way repeated measures (ROI × modality) ANOVA on explained variance [%] values revealed a main effect of ROI ($F(13) = 35.57$, $p < .001$, $\eta_p^2 = .73$) and modality ($F(13) = 4.87$, $p = .046$, $\eta_p^2 = .27$) and an interaction between modality and ROI ($F(13) = 36.51$, $p < .001$, $\eta_p^2 = .74$). Post-hoc tests demonstrated that the highest amount of variance explained by duration was in the auditory cortex for auditory stimuli (70.67%, $t(13) = 20.07$, $p < .001$; Bonferroni corrected; Fig. 3c) which was the only value that reached significance. In the frequency model, two-way repeated measures (ROI × modality) ANOVA on explained variance [%] values revealed no main effects or interactions (all ps > .05).

Decoding numerosity in response trials

Because the parietal activation was present only in response trials, we applied SVC and SVR during the response/decision phase. We modeled the response/decision phase as an event after the presentation of the second numerosity. This allowed us to test whether numerosity coding in the parietal cortex is linked to engagement in an active task (e.g., comparison), including planning and selection of a manual response. Because the second numerosity in response trials was always 25% smaller or larger than the first numerosity (half of the time larger, other half smaller), we took the average of both numerosities in response trials. This resulted in the same numerosities as we used in non-response trials (i.e., 5, 7, 11, and 16).

Interestingly, despite the lower number of response trials, we were able to decode the average numerosity for both modalities in a parietal ROI (activity in response trials masked with Jülich atlas defined bilateral parietal cortex, 1000 most active voxels, Fig. 5b). The decoding accuracy for visual numerosities was 30.80% ($t(13) = 2.77$, $p = .02$) and the accuracy for auditory numerosities was 30.36% ($t(13) = 2.92$, $p = .01$).

To test the modality specificity of numerosity-selective activation patterns, we checked if the classifier trained in a given modality generalized to the other. Generalization was significant only when the classifier was trained on auditory numerosities and tested on visual numerosities (from auditory to visual, mean accuracy = 29.90%, $t(13) = 2.41$, $p = .03$; from visual to auditory, mean accuracy = 26.56%, $t(13) = .92$, $p = .37$). However, the confusion matrix suggests that in the present

data the successful generalization was mainly driven by numerosity 5, limiting the interpretability of this finding (Fig. 5a).

Discussion

Over the last decade, it has been advocated that the IPS stores an abstract representation of numerosity in a presentation notation (i.e., Arabic digits, dot arrays etc.), format (i.e., simultaneous or serial), and sensory-modality independent fashion (Cohen Kadosh et al., 2007; Dormal et al., 2010; Eger et al., 2009, 2003; Nieder, 2012; Nieder et al., 2006; Piazza et al., 2007, 2006). However, in investigating the abstract nature of the numerical magnitude representation, most studies focused on demonstrating notation independence of simultaneously presented numerical information (e.g., sets of dots or visual Arabic digits and number words). How numerosity information is extracted over time from a sequentially presented series of events from different sensory modalities has not been studied in humans in great detail. In the current study, we addressed these issues by investigating the behavioral characteristics and neural correlates of non-symbolic numerosities being sequentially presented in auditory or visual modality. By using a paradigm that allowed us to disentangle task-related processes from numerosity representation, we were able to investigate the representation of sequential numerosity information both with and without a numerical comparison task.

On a behavioral level, we observed that comparing sequentially presented numerosities follows well-known psychophysical characteristics such as Weber's law. Therefore, behavioral response patterns suggest that participants were able to solve the task. On a neuro-functional level, we observed increases in the BOLD response in frontal and parietal regions, including numerosity-related areas around the IPS when participants were engaged in an active task (comparison of two numerosities). In contrast, when participants viewed auditory and visual numerosities without active response engagement, we did not observe numerosity-modulated increase in the BOLD response in numerosity-related areas.

We employed SVC and SVR in four main ROIs (sensory: auditory and visual cortex; parietal: SPL and hIPS) using The Decoding Toolbox (Hebart et al., 2015) to further substantiate the contribution of sensory and numerosity-related brain regions to the estimation of sequentially presented numerosities. None of the multivariate algorithms could classify numerosity in parietal ROIs. In contrast, we were able to classify numerosity information in the relevant sensory areas (i.e., auditory cortex for auditory numerosities and visual cortex for visual numerosities). To further explore the role of the parietal cortex in numerosity processing, we analyzed additional partitions of the parietal cortex that were either defined by the activity in a functional localizer, the experimental task, or the atlas-based anatomical regions. Yet, even this exploratory analysis did not yield any region that had numerosity-modulated increase in the BOLD response that could be successfully classified in both modalities. Finally, because the parietal activation was present only during comparison, we tried to decode average numerosity during the response phase. We were able to decode numerosity in a parietal ROI for both modalities. Yet, when testing generalization between modalities, there was only weak evidence for modality independence.

One might argue that the present (null-) results are due to a lack of statistical power. We determined the number of participants and trials per condition based on previous studies that successfully used the parietal cortex activation to decode the number of dots in dot-arrays (Bulthé et al., 2014; Eger et al., 2009, 2013; Knops et al., 2014). We calculated the mean effect size d_z (Lakens, 2013; Rosenthal, 1991) from these studies and conducted a power analysis to determine the expected statistical power in the current study (see Supplementary Table S3). Assuming a mean observed effect size of $d_z = 1.51$ with alpha = .05 and a sample size of N = 14 participants, G*power software (Faul et al., 2009) estimated statistical power (1 − beta error probability) as power = .99989. Hence the present sample size appeared sufficiently large (i.e., statistical power was sufficiently high) to detect numerosity coding in parietal cortex using MVPA. This conclusion is further corroborated by significant decoding results in response trials on otherwise unchanged stimuli.

Role of the parietal cortex

Based on human neuroimaging, primate neurophysiology, and developmental neuropsychology studies, the TCM assumes that the parietal cortex hosts an abstract semantic representation of numerosities (Ansari et al., 2006; Dehaene et al., 2004; Fias et al., 2001; Nieder, 2012; Nieder et al., 2006; Piazza et al., 2006, 2007). This representation has been claimed to be notation (e.g., Arabic digit or dot-arrays), format (i.e., simultaneous or sequential presentation), and modality (e.g., visual or spoken number words) independent (i.e., abstract; Dehaene et al., 2004; Eger et al., 2003, 2009; Piazza et al., 2007). However, recent neuroimaging studies have challenged the idea of a universally abstract magnitude representation. While Eger et al. (2009) were able to decode symbolic (Arabic digits) and non-symbolic (dot-arrays) numerosities in a parietal ROI, the classifier only generalized from Arabic digits to dot arrays. This asymmetric generalization was attributed to different tuning widths of non-symbolic and symbolic numerosities (Eger et al., 2009). Consistent with notation-specific representations, Damarla and Just (2013) found that different populations of parietal neurons encoded numerosity in non-symbolic (a set of objects, e.g., five tomatoes) and digit-object (an Arabic digit (e.g., 5) in combination with an object (e.g., a tomato)) notations. Moreover, two recent studies pointed out qualitatively different parietal representations of numerosity in dot arrays and Arabic digits (Bulthé et al., 2014; Lyons et al., 2014). Finally, an earlier study failed to attribute the observed parietal activations in numerical tasks to magnitude processing (Shuman and Kanwisher, 2004). Our results may be interpreted as further evidence questioning the abstract characteristic of the magnitude representation in the parietal cortex as postulated in the TCM.

It has been suggested that the IPS activity during numerical tasks might reflect a response-selection process rather than mapping of numerosity onto a parietal number module (Göbel et al., 2004). To overcome this confound, recent studies used an adaptation paradigm to examine the neural response to rapid numerosity changes. Because participants were passively exposed to dot-arrays or Arabic digits, response-selection did not interfere with results (Ansari et al., 2006; Cohen Kadosh et al., 2007, 2011; Piazza et al., 2007). Interestingly, although they arrived at different conclusions about abstract number representation, all reported parietal involvement during passive viewing. A recent study using non-symbolic object arrays reported also a topographic organization of numerosity in the parietal cortex (Harvey et al., 2013). However, these results might also be reflecting sensory topography rather than numerosity (Gebuis et al., 2014).

Considering the engagement of the parietal cortex even in passive viewing of dot-arrays, the lack of parietal involvement in non-response trials of our task could be related to the presentation format (sequential). In the light of recent results pointing to the important role of priority maps during visual scene encoding (Bisley and Goldberg, 2003; Franconeri et al., 2013; Roggeman et al., 2010), one may also wonder whether the parietal functional architecture is suited to sequential stimuli. Parietal regions exhibit a map-like architecture and guide visual perception by integrating bottom-up salience information, top-down strategic goals, and reward experiences into a single map that indicates the relative priority of individual items in a visual scene (Bisley and Goldberg, 2003; Franconeri et al., 2013; Roggeman et al., 2010). Inter-item competition for resources limit performance across cognitive domains such as enumeration and visual short term memory (Knops et al., 2014). Therefore, repeatedly presenting stimuli at one location might not sufficiently activate a visual system built to integrate item information over space rather than over time. Albeit tentative, this speculation is in line with the fact that all previous human fMRI

studies reporting parietal activation in the absence of response used dot-arrays or symbolic number information (e.g., Arabic digits or number words; Ansari et al., 2006; Cohen Kadosh et al., 2007; Harvey et al., 2013; Piazza et al., 2007). Furthermore, two previous studies also reported decreased parietal involvement when the same visual stimuli were shown sequentially compared to when they are shown simultaneously, again suggesting a spatial role for the parietal cortex (Shafritz et al., 2002; Xu and Chun, 2006).

At this point, the role of the parietal cortex in estimation of numerosity information over space and time remains elusive. Given the inconsistent results on notation-independence, what actually drives the parietal activation during symbolic and non-symbolic numerosity perception tasks should be further investigated. Our experiment suggests a role for the parietal cortex only during the comparison of the numerical information. The parietal cortex is known to be involved in finger movements (Baker et al., 1999; Sirigu et al., 1996; Taira et al., 1990). Because participants responded manually, it is not surprising to see parietal activation during response trials. However, the spread of activation could be reflecting response-related processing (e.g., comparison, response selection, and motor preparation) as well as task switching costs (i.e., switching from passive viewing to an active comparison task; Deiber et al., 1996; Hazeltine et al., 2000; Sohn et al., 2000; Yeung et al., 2006). With the current experiment, it is not possible to disentangle which component of response trials is responsible for which part of the observed pattern of activations. Future studies should examine whether the presentation format, the absence of an active task or response accounts for the present results. Our results add to the accumulating evidence that fails to corroborate the abstract nature of magnitude representation in the parietal cortex, potentially warranting an update of the TCM (Bulthé et al., 2014; Cohen Kadosh et al., 2011; Damarla and Just, 2013; Shuman and Kanwisher, 2004).

Role of sensory cortices

Although it has been suggested that numerosity is an abstract feature that is neurally represented in parietal rather than primary sensory cortices, we obtained the best classification results in sensory cortices (auditory cortex (main ROI) for auditory numerosities and visual cortex (exploratory ROI) for visual numerosities). The extent of the BOLD response and classifier accuracy were higher for auditory numerosities in the auditory cortex than for visual numerosities in the visual cortex. This might result from the greater similarity of the visual stimuli to the baseline rest condition. Although in the visual domain the stimuli were delivered by temporally alternating the fixation dot with a blank screen, in the auditory domain the stimuli were delivered by adding sound to the continuously visible fixation dot. We used multivariate analysis (SVC and SVR) to investigate whether numerosity or non-numerical features (duration and frequency) drove the high decoding performance in sensory cortices. Interestingly, classifier accuracy was also high for duration and frequency in the sensory ROIs. Although the main ROI in the visual cortex did not reach significance for frequency, when the ROI size was increased, both SVC and SVR reached significance.

We have two hypotheses concerning the role of sensory cortices. The first hypothesis is that sensory cortices represent numerosity independently. During sequential presentation, numerosity can be estimated by combining total duration and frequency (numerosity = duration × frequency). Therefore, the information in auditory and visual cortices might be sufficient to deduce numerosity and sensory cortices might be engaged in representing numerosity independently from the previously assumed parietal contribution. While the majority of fMRI studies in the field concentrated on the role of parietal cortex in numerosity representation, recent studies reported successful decoding of the number of dots in a set in the visual cortex as well (Bulthé et al., 2014; Eger et al., 2009). The second hypothesis is that sensory cortices are maintaining non-numerical sensory information for future parietal processing. We cannot exclude the possibility that the sensory features (duration and frequency) confounded our results as the decoding accuracy in sensory cortices was higher for duration and frequency than for numerosity. Therefore, the sensory cortices might simply be maintaining sensory information for future comparison. In this case, the actual numerosity estimation may be performed in parietal circuits. Although estimating the numerosity only during comparison does not benefit participants in our task, we cannot rule it out. Given successful decoding of numerosity in sensory cortices both for sequential stimuli and dot-arrays, the exact role of sensory cortices in numerosity estimation should be further investigated.

Role of the frontal cortex

For auditory and visual numerosities, we found prominent increase in the BOLD response in a frontal region (BA 6), both during response and non-response trials. BA 6 is known to be involved in response preparation and inhibition as well as working memory maintenance (Fassbender et al., 2011; Simmonds et al., 2009; Mars et al., 2008; Owen et al., 2005; Suskauer et al., 2010). The lack of activation in the parietal cortex during non-response trials might bring up the idea that participants actually did not keep the information in working memory in non-response trials. However, the parietal activation was absent for the first numerosity in response trials as well, which participants must have kept in working memory considering the high task performance. As demonstrated by multivariate analysis, this frontal activity was not specific to numerical information. Taken together, the non-specific frontal activity seen in non-response trials might reflect response-inhibition and/or working memory related activity. In contrast to Eger et al. (2009), Nieder (2012) and Nieder et al. (2006), we did not find any frontal ROI that coded numerosity irrespective of modality.

Conclusion

In sum, it seems reasonable to assume that time-variant (i.e., sequential) numerical information is processed differently than time-invariant (i.e., simultaneous) numerical information. Given the significant classification of numerosities in the parietal cortex during the comparison/response phase, it is also possible that the numerosity information in sequential stimuli is extracted only during an active task. This is in contrast to Arabic digits and dot-arrays, where the numerosity information is claimed to be extracted even in the absence of a numerical task (Cohen Kadosh and Henik, 2006; Cohen Kadosh, 2008; Dormal et al., 2006; Fias et al., 2001). The role of the presentation format, task and response requirement in parietal activation should be further investigated. In addition, the role of sensory cortices seems to be underestimated. Given the significant decoding accuracy of numerosity in sensory cortices, in our experiment as well as others, it seems reasonable to investigate the exact contribution of sensory information and sensory cortices to numerosity estimation. Future research should also elaborate on when (and why) an abstract parietal representation of numerosity could be beneficial.

Acknowledgments

This work was supported by a grant of Deutsche Forschungsgemeinschaft/German Research Council (DFG) to André Knops (KN 959/2-1) in support of Seda Cavdaroglu.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.04.019.

References

Andres, M., Michaux, N., Pesenti, M., 2012. Common substrate for mental arithmetic and finger representation in the parietal cortex. NeuroImage 62, 1520–1528. http://dx.doi.org/10.1016/j.neuroimage.2012.05.047.

Ansari, D., Dhital, B., Siong, S.C., 2006. Parametric effects of numerical distance on the intraparietal sulcus during passive viewing of rapid numerosity changes. Brain Res. 1067, 181–188. http://dx.doi.org/10.1016/j.brainres.2005.10.083.

Arsalidou, M., Taylor, M.J., 2011. Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. NeuroImage 54, 2382–2393. http://dx.doi.org/10.1016/j.neuroimage.2010.10.009.

Baker, J.T., Donoghue, J.P., Sanes, J.N., 1999. Gaze direction modulates finger movement activation patterns in human cerebral cortex. J. Neurosci. 19, 10044–10052.

Bisley, J., Goldberg, M., 2003. Neuronal activity in the lateral intraparietal area and spatial attention. Science 80–86.

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436. http://dx.doi.org/10.1163/156856897X00357.

Bulthé, J., De Smedt, B., Op de Beeck, H.P., 2014. Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. NeuroImage 87, 311–322. http://dx.doi.org/10.1016/j.neuroimage.2013.10.049.

Chang, C., Lin, C., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 1–39.

Chu, F.W., Vanmarle, K., Geary, D.C., 2013. Quantitative deficits of preschool children at risk for mathematical learning disability. Front. Psychol. 4, 195. http://dx.doi.org/10.3389/fpsyg.2013.00195.

Cohen Kadosh, R., 2008. The laterality effect: myth or truth? Conscious. Cogn. 17, 350–354. http://dx.doi.org/10.1016/j.concog.2007.03.008.

Cohen Kadosh, R., Henik, A., 2006. A common representation for semantic and physical properties. Z. Exp. Psychol. 53, 87–94. http://dx.doi.org/10.1027/1618-3169.53.2.87.

Cohen Kadosh, R., Walsh, V., 2009. Numerical representation in the parietal lobes: abstract or not abstract? Behav. Brain Sci. 32, 313–328. http://dx.doi.org/10.1017/S0140525X09990938 (discussion 328–73).

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., Goebel, R., 2007. Notation-dependent and -independent representations of numbers in the parietal lobes. Neuron 53, 307–314. http://dx.doi.org/10.1016/j.neuron.2006.12.025.

Cohen Kadosh, R., Bahrami, B., Walsh, V., Butterworth, B., Popescu, T., Price, C.J., 2011. Specialization in the human brain: the case of numbers. Front. Hum. Neurosci. 5, 62. http://dx.doi.org/10.3389/fnhum.2011.00062.

Damarla, S.R., Just, M.A., 2013. Decoding the representation of numerical values from brain activation patterns. Hum. Brain Mapp. 34, 2624–2634. http://dx.doi.org/10.1002/hbm.22087.

Dehaene, S., 1992. Varieties of numerical abilities. Cognition 44, 1–42.

Dehaene, S., Cohen, L., 1995. Towards an anatomical and functional model of number processing. Math. Cogn. 1, 83–120.

Dehaene, S., Molko, N., Cohen, L., Wilson, A.J., 2004. Arithmetic and the brain. Curr. Opin. Neurobiol. 14, 218–224. http://dx.doi.org/10.1016/j.conb.2004.03.008.

Deiber, M.P., Ibañez, V., Sadato, N., Hallett, M., 1996. Cerebral structures participating in motor preparation in humans: a positron emission tomography study. J. Neurophysiol. 75, 233–247 (doi:8822554).

Dormal, V., Seron, X., Pesenti, M., 2006. Numerosity–duration interference: a Stroop experiment. Acta Psychol. (Amst) 121, 109–124. http://dx.doi.org/10.1016/j.actpsy.2005.06.003.

Dormal, V., Andres, M., Dormal, G., Pesenti, M., 2010. Mode-dependent and mode-independent representations of numerosity in the right intraparietal sulcus. NeuroImage 52, 1677–1686. http://dx.doi.org/10.1016/j.neuroimage.2010.04.254.

Eger, E., Sterzer, P., Russ, M.O., Giraud, A.-L., Kleinschmidt, A., 2003. A supramodal number representation in human intraparietal cortex. Neuron 37, 719–726. http://dx.doi.org/10.1016/S0896-6273(03)00036-9.

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., Kleinschmidt, A., 2009. Deciphering cortical number coding from human brain activity patterns. Curr. Biol. 19, 1608–1615. http://dx.doi.org/10.1016/j.cub.2009.08.047.

Eger, E., Pinel, P., Dehaene, S., Kleinschmidt, A., 2013. Spatially invariant coding of numerical information in functionally defined subregions of human parietal cortex. Cereb. Cortex http://dx.doi.org/10.1093/cercor/bht323.

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. NeuroImage 25, 1325–1335. http://dx.doi.org/10.1016/j.neuroimage.2004.12.034.

Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K., 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. NeuroImage 32, 570–582. http://dx.doi.org/10.1016/j.neuroimage.2006.04.204.

Eickhoff, S.B., Paus, T., Caspers, S., Grosbras, M.-H., Evans, A.C., Zilles, K., Amunts, K., 2007. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. NeuroImage 36, 511–521. http://dx.doi.org/10.1016/j.neuroimage.2007.03.060.

Eliassen, J., Souza, T., Sanes, J., 2003. Experience-dependent activation patterns in human brain during visual–motor associative learning. J. Neurosci. 23, 10540–10547.

Fassbender, C., Schweitzer, J.B., Cortes, C.R., Tagamets, M.A., Windsor, T.A., Reeves, G.M., Gullapalli, R., 2011. Working memory in attention deficit/hyperactivity disorder is characterized by a lack of specialization of brain function. PLoS One 6, e27240. http://dx.doi.org/10.1371/journal.pone.0027240.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.-G., 2009. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 41, 1149–1160. http://dx.doi.org/10.3758/BRM.41.4.1149.

Fias, W., Lauwereyns, J., Lammertyn, J., 2001. Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. Cogn. Brain Res. 12, 415–423. http://dx.doi.org/10.1016/S0926-6410(01)00078-7.

Fischer, M.H., 2012. A hierarchical view of grounded, embodied, and situated numerical cognition. Cogn. Process. 13, 161–164. http://dx.doi.org/10.1007/s10339-012-0477-5.

Franconeri, S.L., Alvarez, G.A., Cavanagh, P., 2013. Flexible cognitive resources: competitive content maps for attention and memory. Trends Cogn. Sci. 17, 134–141. http://dx.doi.org/10.1016/j.tics.2013.01.010.

Gebuis, T., Gevers, W., Cohen Kadosh, R., 2014. Topographic representation of high-level cognition: numerosity or sensory processing? Trends Cogn. Sci. 18, 1–3. http://dx.doi.org/10.1016/j.tics.2013.10.002.

Göbel, S.M., Johansen-Berg, H., Behrens, T., Rushworth, M.F.S., 2004. Response-selection-related parietal activation during number comparison. J. Cogn. Neurosci. 16, 1536–1551. http://dx.doi.org/10.1162/0898929042568442.

Harvey, B.M., Klein, B.P., Petridou, N., Dumoulin, S.O., 2013. Topographic representation of numerosity in the human parietal. Cortex 496, 1123–1125.

Hazeltine, E., Poldrack, R., Gabrieli, J.D.E., 2000. Neural activation during response competition. J. Cogn. Neurosci. 12, 118–129. http://dx.doi.org/10.1162/089892900563984.

Hebart, M.N., Görgen, K., Haynes, J.-D., 2015. The Decoding Toolbox (TDT): A Versatile Software Package for Multivariate Analyses of Functional Imaging Data 8. pp. 1–18 http://dx.doi.org/10.3389/fninf.2014.00088.

Hubbard, E.M., Piazza, M., Pinel, P., Dehaene, S., 2005. Interactions between number and space in parietal cortex. Nat. Rev. Neurosci. 6, 435–448. http://dx.doi.org/10.1038/nrn1684.

Kleiner, M., Brainard, D.H., Pelli, D.G., 2007. Thirtieth European conference on visual perception. Perception, p. 14.

Knops, A., Piazza, M., Sengupta, R., Eger, E., Melcher, D., 2014. A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. J. Neurosci. 34, 9857–9866. http://dx.doi.org/10.1523/JNEUROSCI.2758-13.2014.

Koten, J.W., Lonnemann, J., Willmes, K., Knops, A., 2011. Micro and macro pattern analyses of fMRI data support both early and late interaction of numerical and spatial information. Front. Hum. Neurosci. 5, 1–12. http://dx.doi.org/10.3389/fnhum.2011.00115.

Lakens, D., 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front. Psychol. 4, 1–12. http://dx.doi.org/10.3389/fpsyg.2013.00863.

Lyons, I.M., Ansari, D., Beilock, S.L., 2014. Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain. Hum. Brain Mapp. http://dx.doi.org/10.1002/hbm.22641 (00, n/a–n/a).

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. NeuroImage 19, 1233–1239. http://dx.doi.org/10.1016/S1053-8119(03)00169-1.

Maldjian, J.A., Laurienti, P.J., Burdette, J.H., 2004. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. NeuroImage 21, 450–455. http://dx.doi.org/10.1016/j.neuroimage.2003.09.032.

Mars, R.B., Coles, M.G.H., Hulstijn, W., Toni, I., 2008. Delay-related cerebral activity and motor preparation. Cortex 44, 507–520. http://dx.doi.org/10.1016/j.cortex.2007.07.002.

Mathôt, S., Schreij, D., Theeuwes, J., 2012. OpenSesame: an open-source, graphical experiment builder for the social sciences. Behav. Res. Methods 44, 314–324. http://dx.doi.org/10.3758/s13428-011-0168-7.

Mazzocco, M.M.M., 2009. Mathematical learning disability in girls with Turner syndrome: a challenge to defining MLD and its subtypes. Dev. Disabil. Res. Rev. 15, 35–44. http://dx.doi.org/10.1002/ddrr.50.

McCloskey, M., 1992. Cognitive mechanisms in numerical processing: evidence from acquired dyscalculia. Cognition 44, 107–157. http://dx.doi.org/10.1016/0010-0277(92)90052-J.

Nieder, A., 2012. Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. Proc. Natl. Acad. Sci. U. S. A. 109, 11860–11865. http://dx.doi.org/10.1073/pnas.1204580109.

Nieder, A., Dehaene, S., 2009. Representation of number in the brain. Annu. Rev. Neurosci. 32, 185–208. http://dx.doi.org/10.1146/annurev.neuro.051508.135550.

Nieder, A., Diester, I., Tudusciuc, O., 2006. Temporal and spatial enumeration processes in the primate parietal cortex. Science 313, 1431–1435. http://dx.doi.org/10.1126/science.1130308.

Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E., 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. 25, 46–59. http://dx.doi.org/10.1002/hbm.20131.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–442.

Piazza, M., Mechelli, A., Price, C.J., Butterworth, B., 2006. Exact and approximate judgements of visual and auditory numerosity: an fMRI study. Brain Res. 1106, 177–188. http://dx.doi.org/10.1016/j.brainres.2006.05.104.

Piazza, M., Pinel, P., Le Bihan, D., Dehaene, S., 2007. A magnitude code common to numerosities and number symbols in human intraparietal cortex. Neuron 53, 293–305. http://dx.doi.org/10.1016/j.neuron.2006.11.022.

Pinel, P., Thirion, B., Meriaux, S., 2007. Fast reproducible identification and large-scale databasing of individual functional cognitive networks. BMC 8, 91. http://dx.doi.org/10.1186/1471-2202-8-91.

Roggeman, C., Fias, W., Verguts, T., 2010. Salience maps in parietal cortex: imaging and computational modeling. NeuroImage 52, 1005–1014. http://dx.doi.org/10.1016/j.neuroimage.2010.01.060.

Rosenthal, R., 1991. Meta-analysis Procedures for Social Research.

Sawamura, H., Shima, K., Tanji, J., 2002. Numerical representation for action in the parietal cortex of the monkey. Nature 415, 918–922. http://dx.doi.org/10.1038/415918a.

Schumacher, E.H., Jiang, Y., 2003. Neural mechanisms for response selection: representation specific or modality independent? J. Cogn. Neurosci. 15, 1077–1079. http://dx.doi.org/10.1162/089892903322598058.

Shafritz, K.M., Gore, J.C., Marois, R., 2002. The role of the parietal cortex in visual feature binding. Proc. Natl. Acad. Sci. U. S. A. 99, 10917–10922. http://dx.doi.org/10.1073/pnas.152694799.

Shuman, M., Kanwisher, N., 2004. Numerical magnitude in the human parietal lobe: tests of representational generality and domain specificity. Neuron 44, 557–569. http://dx.doi.org/10.1016/j.neuron.2004.10.008.

Simmonds, D.J., Pekar, J.J., Mostofsky, S.H., 2008. Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. Neuropsychologia 46, 224–232. http://dx.doi.org/10.1016/j.neuropsychologia.2007.07.015.

Sirigu, A., Duhamel, J.R., Cohen, L., Pillon, B., Dubois, B., Agid, Y., 1996. The mental representation of hand movements after parietal cortex damage. Science 273, 1564–1568. http://dx.doi.org/10.1126/science.273.5281.1564.

Sohn, M.H., Ursu, S., Anderson, J.R., Stenger, V.A., Carter, C.S., 2000. The role of prefrontal cortex and posterior parietal cortex in task switching. Proc. Natl. Acad. Sci. U. S. A. 97, 13448–13453. http://dx.doi.org/10.1073/pnas.240460497.

Suskauer, S.J., Simmonds, D.J., Fotedar, S., Blankner, J.G., Pekar, J., Denckla, M.B., Mostofsky, S.H., 2010. NIH public access. J. Cogn. Neurosci. 20, 478–493. http://dx.doi.org/10.1162/jocn.2008.20032.Functional.

Taira, M., Mine, S., Georgopoulos, A.P., Murata, A., Sakata, H., 1990. Parietal cortex neurons of the monkey related to the visual guidance of hand movement. Exp. Brain Res. 83, 29–36. http://dx.doi.org/10.1007/BF00232190.

Tokita, M., Ishiguchi, A., 2011. Temporal information affects the performance of numerosity discrimination: behavioral evidence for a shared system for numerosity and temporal processing. Psychon. Bull. Rev. 550–556. http://dx.doi.org/10.3758/s13423-011-0072-2.

Xu, Y., Chun, M.M., 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. Nature 440, 91–95. http://dx.doi.org/10.1038/nature04262.

Yeung, N., Nystrom, L.E., Aronson, J.A., Cohen, J.D., 2006. Between-task competition and cognitive control in task switching. J. Neurosci. 26, 1429–1438. http://dx.doi.org/10.1523/JNEUROSCI.3109-05.2006.

Consciousness and Cognition 31 (2015) 60–72

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Exploring the boundary conditions of unconscious numerical priming effects with continuous flash suppression

G. Hesselmann [a,*], N. Darcy [a], P. Sterzer [a], A. Knops [b]

[a] Visual Perception Laboratory, Department of Psychiatry and Psychotherapy, Campus Charité Mitte, Charité – Universitätsmedizin, 10117 Berlin, Germany
[b] Department of Psychology, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

ARTICLE INFO

Article history:
Received 14 July 2014
Available online 15 November 2014

Keywords:
Continuous flash suppression
Interocular suppression
Priming
Number processing

ABSTRACT

The scope and limits of unconscious processing are a controversial topic of research in experimental psychology. Particularly within the visual domain, a wide range of paradigms have been used to experimentally manipulate perceptual awareness. A recent study reported unconscious numerical processing during continuous flash suppression (CFS), which is a powerful variant of interocular suppression and disrupts the conscious perception of visual stimuli for up to seconds. Since this finding of a distance-dependent priming effect contradicts earlier results showing that interocular suppression abolishes semantic processing, we sought to investigate the boundary conditions of this effect in two experiments. Using statistical analyses and experimental designs that precluded an effect of target numerosity, we found evidence for identity priming, but no conclusive evidence for distance-dependent numerical priming under CFS. Our results suggest that previous conclusions on high-level numerical priming under interocular suppression may have been premature.

1. Introduction

The scope and limits of unconscious perceptual priming effects have been a central and, at the same time, controversial topic of research in experimental psychology and cognitive neuroscience for the last decades (Holender, 1986; Kouider & Dehaene, 2007). In the course of this scientific endeavor, a wide range of paradigms have been used to present stimuli outside of participants' awareness, particularly within the visual domain (Bachmann, Breitmeyer, & Ogmen, 2007; Kim & Blake, 2005). One emerging view is that not all invisible stimuli are equally invisible, since different paradigms suppress the conscious perception of stimuli at different levels of the neurocognitive architecture (Breitmeyer, 2014; Faivre, Berthet, & Kouider, 2012; Fogelson, Kohler, Miller, Granger, & Tse, 2014).

A paradigm that has recently become very popular for the investigation of unconscious visual processing is continuous flash suppression (CFS): high-contrast dynamic patterns shown to one eye disrupt the conscious perception of a low-contrast stimulus shown to the other eye for up to several seconds (Tsuchiya & Koch, 2005). In contrast to binocular rivalry, onset and offset of stimulus suppression can be deterministically controlled by switching the dynamic CFS masks on and off, respectively. Behavioral and neuroimaging studies using CFS have already produced a large but heterogeneous body of evidence regarding the types of visual information that can be processed during this powerful variant of interocular suppression (Gayet, Van der Stigchel, & Paffen, 2014; Hesselmann, 2013; Sterzer, Stein, Ludwig, Rothkirch, & Hesselmann, 2014; Yang, Brascamp, Kang, & Blake, 2014).

* Corresponding author.
E-mail address: guido.hesselmann@charite.de (G. Hesselmann).

http://dx.doi.org/10.1016/j.concog.2014.10.009
1053-8100/

The findings of a recent study (Bahrami et al., 2010) suggested that numerical processing of small quantities (1–3) can escape CFS and lead to robust numerical priming effects in an enumeration task. In a series of three priming experiments (and two control experiments to assess prime invisibility), the authors showed that unconsciously presented non-symbolic and symbolic primes (sets of Gabor patches and Arabic digits, respectively) induced a priming effect for non-symbolic numerosity targets which was linearly dependent on the numerical distance between target (t) and prime (p). Specifically, the priming effect signaled "interference" (i.e., slower RTs relative to a prime-absent baseline) for negative t–p distances (e.g., t: 1, p: 3), and "facilitation" for positive t–p distances (e.g., t: 3, p: 1), while facilitatory priming was relatively small and less robust for zero t–p distance, i.e., numerically congruent trials (e.g., t: 2, p: 2). The same priming function was observed for invisible and visible primes. While intriguing and robust across experiments, this pattern of results merits further investigation due to two reasons. First, the results are in disagreement with earlier interocular suppression studies which have shown that binocular rivalry abolishes visual semantic priming (Cave, Blake, & McNamara, 1998; Zimba & Blake, 1983), and that semantic analysis does not occur in the absence of awareness induced by CFS (Kang, Blake, & Woodman, 2011). Second, the specific shape of the observed priming functions is difficult to reconcile with the results from previous numerical priming studies. For example, it has usually been found that when a target is preceded by a prime number, participants' response latencies decrease with decreasing absolute t–p distance (Dehaene et al., 1998; Koechlin, Naccache, Block, & Dehaene, 1999). This well-established feature of numerosity priming is generally explained by increasing representational overlap between the prime and the target as t–p distance decreases (Van Opstal, Gevers, De Moor, & Verguts, 2008).

Furthermore, the pattern of priming across t–p distances has been shown to depend on the notation of the prime: While V-shaped priming (centered on zero t–p distance) was found for symbolic digit primes, a step-like priming function resulted from trials with non-symbolic dot primes (Roggeman, Verguts, & Fias, 2007). Computational models of number perception often assume two different coding schemes (Dehaene & Changeux, 1993; Verguts & Fias, 2004). V-shaped priming effects are taken as indicative of a place coding scheme. Since numerical magnitude is thought to be represented along a continuum, place coding implies the activation of circumscribed positions (i.e., numbers) along this continuum which in turn leads to the co-activation of neighboring positions (i.e., neighboring numbers). Step-like priming functions are thought to indicate a summation coding scheme. Here, a given numerosity corresponds to the sum of activated nodes along the continuum. That is, a given numerosity activates all nodes up to and including a certain node, thereby leading to step-like priming functions. Surprisingly, Bahrami et al. (2010) reported similar linear priming functions for both non-symbolic and symbolic primes.

Here, we sought to investigate the boundary conditions of the distance-dependent priming effect previously observed under CFS. In the first experiment, we asked whether the linear priming effect, which was originally reported by Bahrami et al. (2010) for primes and targets within the subitizing range (Burr, Turi, & Anobile, 2010; Kaufman & Lord, 1949), generalizes to larger numerosities (>4). It has been suggested that the apperception of small and large numerosities invoke distinct cognitive functions, and more specifically, that subitizing is functionally different from estimation (Revkin, Piazza, Izard, Cohen, & Dehaene, 2008). Experiment 1 closely followed the procedures of the original numerosity priming study (Bahrami et al., 2010) but involved only non-symbolic primes and non-symbolic targets either in a small or large numerosity range.

2. Experiment 1

2.1. Methods and materials

2.1.1. Participants

We determined sample size based on the original study by Bahrami et al. (2010). The authors reported F and p values and degrees of freedom for three independent repeated measures ANOVAs with factor "t–p distance" (Exp. 2: $N = 17$, $F_{4,64} = 7.72$, $p < .0001$; Exp. 3: $N = 13$, $F_{4,48} = 2.6$; $p = .04$; Exp. 4: $N = 16$, $F_{4,60} = 4.50$, $p = .003$). We calculated the associated effect sizes f (Exp. 2: $f = 0.43$; Exp. 3: $f = 0.28$; Exp. 4: $f = 0.34$) assuming a mean correlation between repetitions of 0.5. Using G∗Power 3.1.9 (Faul, Erdfelder, Lang, & Buchner, 2007) we determined that for $f = 0.35$, and $\alpha = 0.05$, a sample size of $N = 14$ was required to achieve a power of 0.90 (actual power: 0.91).

19 observers participated in our experiment, which was conducted with local ethics approval at the Department of Psychiatry and Psychotherapy, Charité-Universitätsmedizin Berlin, Germany. They were recruited from a student pool via email and paid 8 €/h for their participation. Three participants were excluded from further analyses because they showed significant above-chance forced-choice discrimination performance for invisible stimuli in the control experiment (see Section 2.1.10). All remaining 16 participants (9 female, mean age: 22, range: 18–30 years) had normal or corrected-to-normal vision, were naïve to the purpose of the study, and provided informed written consent. GH, ND, and KW collected data.

2.1.2. Apparatus and setup

Participants were seated in a dark environment, the only light coming from the experimental monitor and a second monitor, and viewed the dichoptic images on a 19" CRT monitor (SAMTRON 98PDF; effective screen diagonal: 43.6 cm; refresh rate 60 Hz) via a mirror stereoscope. To stabilize head position the participants placed their heads on a chinrest. The viewing distance from the eyes to the screen (including distances within the mirror system) was 66 cm. All stimuli were generated with PsychToolbox 3 (Brainard, 1997; Pelli, 1997) running under Matlab R2007b (MathWorks Inc., USA).

2.1.3. Stimuli

In the small range condition, prime and target stimuli consisted of 1, 2 or 3 vertically oriented Gabor patches (diameter: 0.79–1.61°; spatial frequency: 1.8 cpd). In the large range condition, prime and target stimuli consisted of 10, 14 or 20 vertically oriented Gabor patches (diameter: 0.35–0.79°; spatial frequency: 3.9 cpd). We aimed to obtain comparable discrimination performance in both numerosity ranges by choosing larger numerical intervals in the large range which followed Weber's law. In total, we created 180 arrays of pseudo-randomly positioned Gabor patches; we controlled non-numeric stimulus properties by using the routines described by Gebuis and Reynvoet (2011), separately for each numerical range. Non-numeric stimulus properties included convex hull (i.e., the shortest contour around a given Gabor set), overall surface of the Gabors (i.e., sum of covered pixels across all Gabors in a given set), density (i.e., convex hull divided by overall surface), average Gabor size, and contour length. Each parameter correlated positively with numerosity in one half of the stimuli and negatively in the other half. Thus, no cue could consistently be used to reliably infer numerosity.

2.1.4. Interocular masking

We used continuous flash suppression (CFS) to render the Gabor prime stimuli invisible (Tsuchiya & Koch, 2005; Tsuchiya, Koch, Gilroy, & Blake, 2006). CFS uses high-contrast dynamic images (masks) flashed to one eye to suppress images presented to the other eye from awareness. The mask images consisted of rectangles and circles of 5 grey levels (ranging from black to white) and sizes ranging from 4% to 18% of the size of the CFS area which measured 8.11°. The rectangles and circles were positioned at random locations on the mask image. 25 of these images were produced and flashed in random order at 10 Hz to the dominant eye (Fig. 1). To minimize afterimages of the prime stimuli, the Gabor targets were superimposed on the CFS masks. A white square frame was presented around the stimuli to promote stable binocular fusion during dichoptic presentation. Throughout each trial, a central red fixation cross was presented (0.26°); during the presentation of the prime stimulus, the fixation cross was shown only to the dominant eye. The dominant eye was determined using a hole-in-card-test (Miles, 1930).

2.1.5. Procedure

Prior to the main experiment, a training session acquainted participants with the different target numerosities. In two small range and large range blocks (15 trials each), participants were instructed to enumerate Gabor targets (presented for 200 ms without CFS), and feedback was provided after each trial. Each block started with a screen indicating the numerosity range of the subsequent block. Participants used the arrow keys on a PC keyboard to indicate the numerosity (left: small; down: medium; right: large), and they were instructed to respond as quickly and as accurately as possible.

In the main experiment, the small range and large range conditions were presented in alternating blocks of 40 trials. The order of blocks was counterbalanced across participants. Each block started with a screen indicating the numerosity range of the subsequent block. The prime stimulus was presented together with the CFS masks for a random duration (range 1000–3000 ms). After the prime stimulus, CFS masks were shown to both eyes for 2000 ms. The interval between prime offset and target onset was random (range 100–200 ms). In prime-absent trials, only the CFS masks were presented. The target was superimposed on CFS masks and presented to both eyes for 200 ms. Participants used the arrow keys to indicate the target numerosity (left: small; down: medium; right: large), and they were instructed to respond as quickly and as accurately as possible. After their speeded response (but only after the 2000 ms of post-prime CFS), participants were asked to provide their subjective visibility rating for the prime stimulus according to the perceptual awareness scale [PAS, (Ramsoy & Overgaard, 2004)]. Participants were made familiar with the levels of the PAS prior to the experiment. On each trial, the four

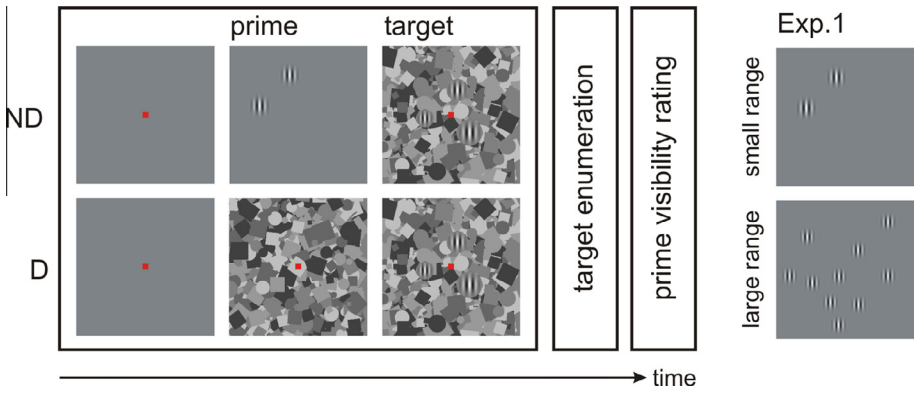

Fig. 1. The continuous flash suppression (CFS) paradigm. Using a mirror stereoscope, Gabor primes were presented to participants' non-dominant eye (ND), while Mondrian masks were flashed to the dominant eye (D) at 10 Hz. In each trial of the main experiment, participants had to enumerate the visible targets as quickly and as accurately as possible, and then rate the visibility of the primes using the 4-point perceptual awareness scale (PAS). Targets were embedded in dynamic CFS masks to minimize prime afterimages. Shown is a trial from the small numerosity range in Experiment 1 (target: 3; prime: 2; *t–p* distance: +1). In Exp. 1, primes and targets were either 1, 2, 3 (small range), or 10, 14, 20 (large range). In Exp. 2, targets were 3 and 4, and primes were 1–6.

levels of the PAS were vertically presented on the screen (“nothing”, “weak glimpse”, “almost clear”, “fully visible”), and participants used the up/down arrow keys to move a cursor and select their rating. The cursor randomly preselected one of the PAS ratings on each trial. By pressing the space bar participants confirmed their rating and started the next trial. For each numerosity range, a total of 240 trials were presented (80 prime-absent trials, 160 prime-present trials).

The awareness control experiment followed immediately after the main experiment. In this experiment, no targets were presented after the presentation of the suppressed primes. Participants were instructed to indicate the numerosity of the primes and provide prime visibility ratings. Participants were instructed to guess prime numerosity in case the primes were invisible. A total of 60 trials were presented. The order of main and control experiment is important because it has been shown previously that participants can learn to better detect the suppressed stimuli during the course of many trials (Ludwig, Sterzer, Kathmann, Franz, & Hesselmann, 2013). The whole experiment lasted approximately 90 min.

2.1.6. Prime contrast

Prior to the main experiment, Gabor prime contrasts were adjusted individually for each participant following a staircase procedure, as follows: After a stimulus presentation that conformed to that of the main experiment, participants had to press a key according to whether the stimulus had been visible or not. Based on this response, the stimulus contrast was decreased or increased following a logarithmic scale in the next trial (1-up-1-down staircase). Each participant completed two staircases of 20 trials for each of the stimulus categories (small range, large range). The stimulus contrast in the main experiment was set to the highest stimulus contrast that the participant always judged as invisible in the pretest. The contrast of the small and large numerosity Gabor displays was adjusted separately. This individual adjustment of the prime contrast was performed to ensure maximal stimulus strength even under full suppression. For the small numerosity range, the resulting contrast was 0.24 ± 0.05. For the large numerosity range, contrast was 0.14 ± 0.03 (mean ± SEM, in arbitrary units of stimulus alpha level).

2.1.7. Exclusion of trials

In the main experiment, only data from trials with correct responses were analyzed. Trials with response time (RT) outliers were determined (i.e., RTs shorter than the first quartile minus 1.5 times the interquartile range, or longer than the third quartile plus 1.5 times the inter-quartile range), separately for the small and large numerosity range, and excluded from all further analysis. To further minimize the effect of outliers, trials with anticipatory responses (RT < 100 ms) were excluded as well (Whelan, 2008).

2.1.8. Data analysis using baselines

We did not follow the baseline approach originally reported by Bahrami et al. (2010) because it does not fully account for the potential effects of target numerosity (Hesselmann & Knops, 2014). Details of this analysis can be found in the Supplement.

2.1.9. Alternative data analysis

As an alternative statistical approach, we first tested for the effects of target numerosity and identity priming using a 2×3 factorial repeated measures ANOVA with factors “prime presence” (2 levels) and “target numerosity” (3 levels), separately for the small and large numerosity range. The ANOVA was based on prime-present conditions with zero t–p distance (small range: “t: 1, p: 1”, “t: 2, p: 2”, and “t: 3, p: 3”; large range: “t: 10, p: 10”, “t: 14, p: 14”, and “t: 20, p: 20”) and all prime-absent conditions (small range: “t: 1, no prime”, “t: 2, no prime”, and “t: 3, no prime”; large range: “t: 10, no prime”, “t: 14, no prime”, and “t: 20, no prime”). Degrees of freedom were Greenhouse-Geisser (ε) corrected to account for possible violations of sphericity (Greenhouse & Geisser, 1959). We report partial eta squared (η_p^2) as measure of effect size (SPSS 13.0 for Windows, SPSS Inc.). In addition, the effects were tested using non-parametric Wilcoxon signed-rank tests.

Next, we conducted the multiple regression analysis reported by Roggeman et al. (2007) to directly test for different shapes of the distance-dependent priming functions. This analysis included prime-present conditions only. We fitted regression equations with three predictors that coded for a step-function, a V-shaped function, and a linear function, respectively. The step-function predictor had a value equal to -1 if prime $\geqslant$ target and a value +1 if prime < target. The V-function predictor had values equal to |target–prime|, i.e., the absolute t–p distance. The linear predictor had values equal to target–prime, i.e., the t–p distance. In addition to these three predictors, an intercept was included in the regression. The regression was run for each participant separately (Lorch & Myers, 1990). We report the mean regression coefficients (β) as well as the 95% confidence intervals (CIs) for each function shape predictor (Cumming, 2014).

2.1.10. Exclusion of participants

In the control experiment, 3AFC forced-choice discrimination performance for Gabors rated as invisible (PAS = 1) was determined for each participant and submitted to binomial tests (chance level: 33%; alpha level: 0.05). Three participants were excluded because they showed significant above-chance performance (49%, 56%, and 47%; all $p < .05$). Note that this procedure based on objective task performance differs from Bahrami et al. (2010) who excluded participants for whom the frequency of trials where the prime was rated subjectively invisible was less than the frequency of one of the other two subjective visibility levels.

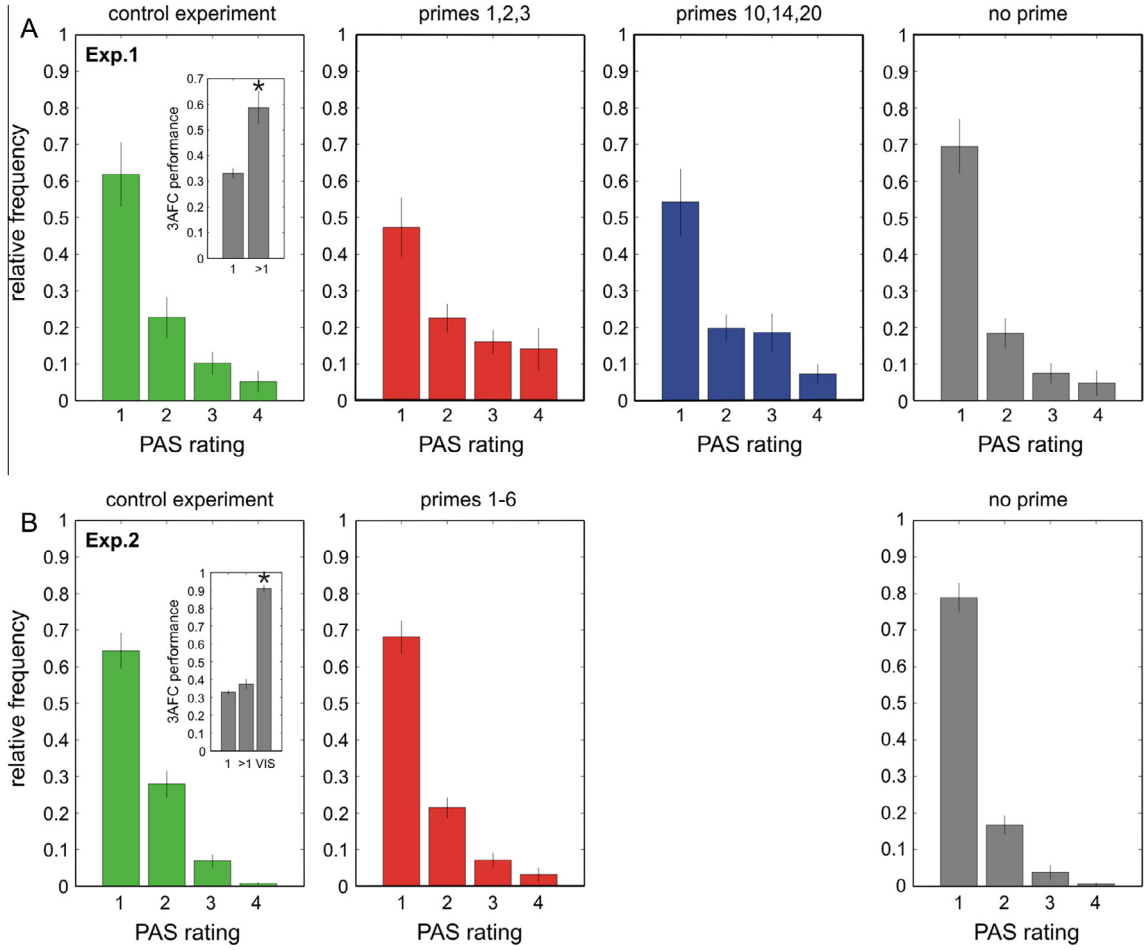

Fig. 2. Subjective prime visibility ratings and objective prime discrimination performance in Experiments 1 and 2. (A) Experiment 1. PAS ratings and 3AFC discrimination performance (inset, $^{*}p < .005$) in the control experiment (green bars), and PAS ratings in the main experiment, separately for the small (primes 1, 2, 3; red bars) and large (primes 10, 14, 20; blue bars) numerosity range and in prime-absent trials (grey bars). (B) Experiment 2. PAS ratings and 3AFC discrimination performance (inset, $^{*}p < .001$) in the control experiment (green bars), and PAS ratings in the main experiment (primes 1–6; red bars) and in prime-absent trials (grey bars). Insets: "1" denotes PAS ratings of 1; ">1" denotes PAS ratings of 2, 3, 4; "VIS" denotes full-visibility trials without CFS. Plotted is the mean ± standard error of the mean. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.2. Results

2.2.1. Prime visibility and prime numerosity discrimination

Fig. 2A summarizes subjective prime visibility ratings and objective prime discrimination performance in Experiment 1. In the control experiment, which was conducted immediately after the main experiment, primes were predominantly rated as invisible (PAS = 1; 61.83%), intermediate (PAS = 2 or 3) and full visibility ratings (PAS = 4) were less frequent. Forced-choice discrimination performance for invisible primes was not significantly above 3AFC chance level (33.22%; $t_{15} = -.06$, $p = .955$; one-sided t-test). By contrast, when primes were rated as visible (PAS > 1), performance was significantly above chance level (58.74%; $t_{13} = 3.67$, $p < .005$; df = 13 because two participants gave less than three "visible" ratings and were excluded from this test). In the main experiment, "invisible" ratings were most frequent, both in the small (47.36%) and large (54.27%) numerosity conditions, as well as in prime-absent trials (69.48%).

2.2.2. Target numerosity discrimination

In the small numerosity range, 3AFC discrimination of target numerosity was at 94.55 ± 1.00%. In the large numerosity range, discrimination performance was 69.97 ± 2.03%.

2.2.3. Small numerosity range: RTs in trials with invisible primes

For the small numerosity range (RT outliers: 5.35%), Fig. 3A summarizes RTs in trials with invisible primes (PAS = 1). RTs are plotted separately for all t–p combinations in prime-present trials, as well as in prime-absent trials ("target only"). T–p combinations that were not part of the experimental design are marked with Xs. The figure shows RT differences between

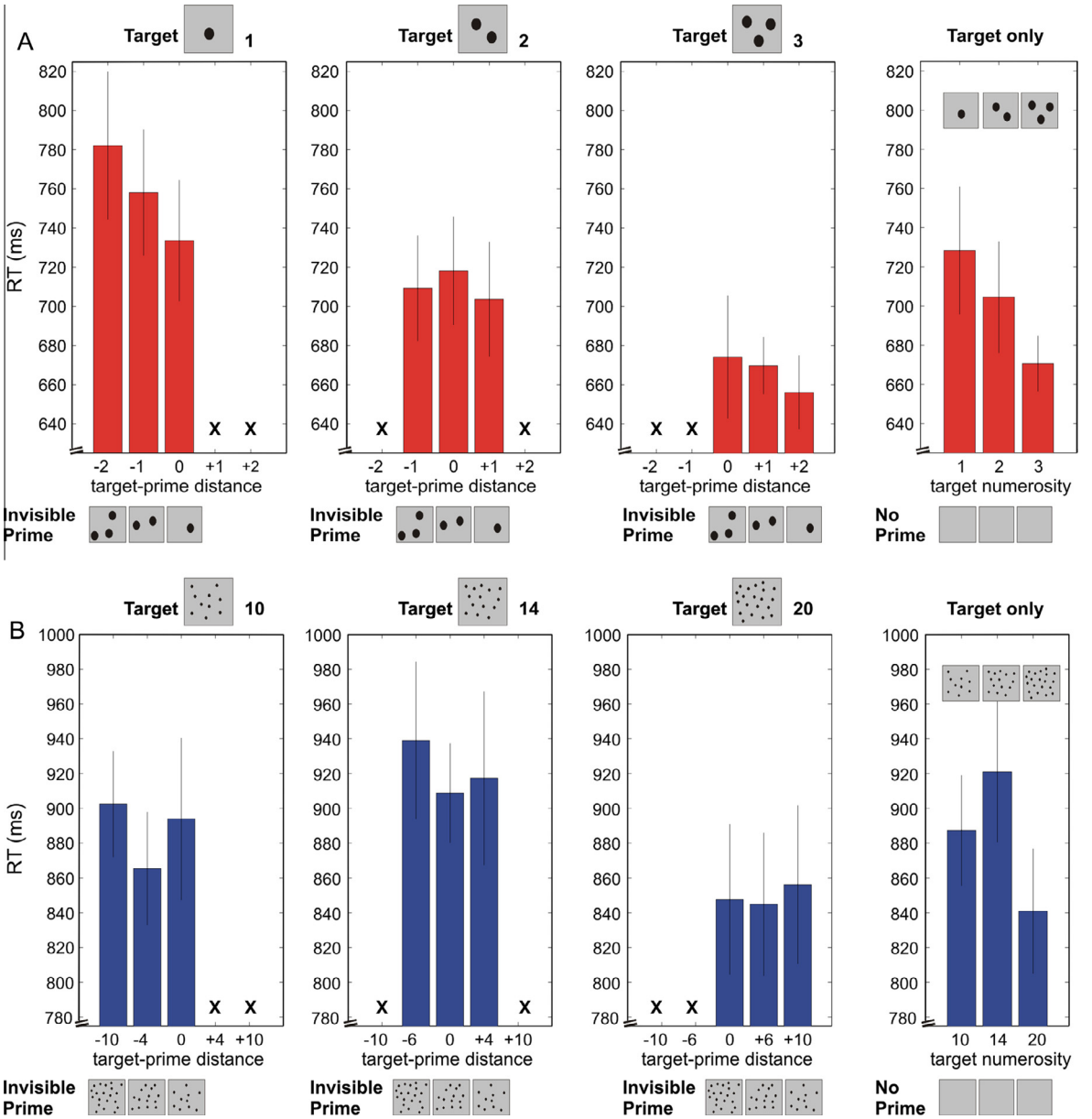

Fig. 3. RTs in Experiment 1. (A) RTs in the small numerosity range (1, 2, 3), plotted separately for all target-prime combinations and prime-absent trials (“target only”). (B) RTs in the large numerosity range (10, 14, 20), plotted separately for all target-prime combinations and prime-absent trials (“target only”). (A and B) Xs denote target-prime combinations which did not exist in the experimental design. Plotted is the mean ± standard error of the mean.

responses to target numerosities 1, 2, and 3, both in trials with invisible primes (mean RTs: 758, 706, and 663 ms) and in prime-absent trials (728, 704, and 671 ms). The figure also suggests that, albeit to a lesser degree, RTs were modulated by *t*–*p* distance, primarily for target numerosity 1.

We first tested for the effects of target numerosity and identity priming. To this end, we performed an ANOVA with the factors “target numerosity” and “prime presence” only on the trials with zero *t*–*p* distance. There was a significant main effect of target numerosity ($F_{2,24} = 4.64$, $p = .020$, $\varepsilon = .99$, $\eta_p^2 = .28$) but no effect of prime presence on RTs ($F_{1,12} = .74$, $p = .405$, $\varepsilon = 1.00$, $\eta_p^2 = .06$) and no interaction ($F_{2,24} = .24$, $p = .743$, $\varepsilon = .81$, $\eta_p^2 = .02$; three participants had less than one trial in one of the conditions and were excluded from this test). Post hoc paired *t*-tests showed a significant difference between target numerosity 1 and 3 ($t_{15} = 2.66$, $p = .018$), but not between 1 and 2 ($t_{15} = 1.50$, $p = .154$) or 2 and 3 ($t_{15} = 1.71$, $p = .109$). Non-parametric statistical tests confirmed this pattern of results (numerosity: 1 versus 3: $Z = 2.53$, $p = .011$; 1 versus 2: $Z = 1.09$, $p = .278$; 2 versus 3: $Z = 1.66$, $p = .100$; prime presence: $Z = .23$, $p = .820$).

We then applied multiple regression analysis to prime-present RTs to test for different shapes of the distance-dependent priming functions (mean $R^2 = .69$). Regression coefficients of the linear (L) function predictor did not deviate from zero (mean

β_L = 12.32, 95% CI [−4.11, 28.76]), nor did we observe any V-shaped (V) or step-like (S) effects (β_V = 6.91, [−7.55, 21.36]; β_S = 0.44, [−16.11, 17.00]). The coefficients of the target numerosity (N) predictor, however, deviated from zero (β_N = −36.01, [−58.90, −13.12]), thus confirming the main effect of target numerosity.

2.2.4. Large numerosity range: RTs in trials with invisible primes

For the large numerosity range (RT outliers: 4.09%), Fig. 3B summarizes RTs in trials with invisible primes (PAS = 1). RTs are plotted separately for all *t–p* combinations in prime-present trials, as well as in prime-absent trials ("target only"). *T–p* combinations that were not part of the experimental design are marked with Xs. The figure suggests RT differences between responses to target numerosities 10, 14, and 20, both in trials with invisible primes (884, 918, and 859 ms) and in prime-absent trials (887, 921, and 841 ms).

The ANOVA yielded no significant main effects (target numerosity: $F_{2,22}$ = 2.29, p = .128, ε = .95, η_p^2 = .17; prime presence: $F_{1,11}$ < .01, p = .999, ε = 1.00, η_p^2 < .01) and no interaction ($F_{2,22}$ = .30, p = .741, ε = .98, η_p^2 = .03; four participants had less than one trial in one of the conditions and were excluded from this test). Non-parametric statistical tests confirmed the absence of significant effects of target numerosity and identity priming (numerosity: 1 versus 3: Z = .72, p = .469; 1 versus 2: Z = 1.14, p = .255; 2 versus 3: Z = 1.81, p = .070; prime presence: Z = .68, p = .496). When multiple regression analysis was applied to prime-present RTs to directly test for different shapes of the priming function (mean R^2 = .53), regression coefficients of the linear function predictor deviated from zero (mean β_L = 6.49, 95% CI [0.34, 12.64]), but none of the other coefficients (β_V = −4.25, [−9.02, 0.51]; β_S = 40.50, [−3.09, 84.08]; β_N = −1.17, [−8.44, 6.10]).

2.3. Discussion

In Experiment 1, we investigated whether the distance-dependent unconscious priming effect observed in the subitizing range for primes rendered invisible by CFS (Bahrami et al., 2010) generalizes to larger numerosities. We did not find conclusive evidence for identity priming or distance-dependent numerical priming effects, neither in the small nor in the large numerosity range. In the small numerosity range, we observed significant target-specific RT differences, which may falsely signal the presence of distance-dependent priming effects, if the data are analysed using a baseline approach (Hesselmann & Knops, 2014).

Why did we observe significant RT differences for small target numerosities (1–3), known to be within the subitizing range? Specifically, why did responses to target numerosity 1 take approx. 100 ms longer than to target numerosity 3? Numerosities within subitizing range usually exhibit a flat reaction time profile (Burr et al., 2010; Knops, Piazza, Sengupta, Eger, & Melcher, 2014; Piazza, Fumarola, Chinello, & Melcher, 2011). We reasoned that the observed RT differences are due to how stimuli were presented in the experimental paradigm used by Bahrami et al. (2010) and by us. In both studies, Gabor targets were superimposed on dynamic CFS masks to avoid prime afterimages, thereby embedding the targets in visual noise. We suggest that this feature of the visual stimulation added a search component to the task: When participants only detected one Gabor patch, they checked whether another one or two patches were to be found; when they immediately detected three patches, they stopped their search, knowing that this was the maximal number of patches. Note that the same point can be made for the large numerosity range in our study, even though these results were less clear-cut. The "visual noise" account is also capable of explaining why the baseline approach applied to the small range data from our Experiment 1 (Hesselmann & Knops, 2014) resulted in a linear priming function of larger effect size (f = 1.22) than the linear priming functions observed by Bahrami and colleagues in their study (average f = 0.35). In our setup, both CFS masks and Gabor targets were grayscale, while Bahrami and colleagues used colored CFS masks and grayscale Gabor targets. It is likely that targets were less salient in our case and further aggravated the target-specific RT differences, thus signaling a larger "priming" effect when effects of target numerosity were not fully accounted for. Further evidence for the "visual noise" account comes from Experiment 1b (see Supplement). Using highly salient targets (pink color) abolished latency differences between target numerosities, thus yielding a response pattern which is in line with previous studies investigating subitizing. Importantly, no linear priming function indicative of distance-dependent priming was observed.

Exploratory analyses of the data from Experiment 1 revealed that in both numerosity conditions subjective prime visibility (i.e., mean PAS ratings) increased with prime numerosity in the main experiment (small range: 1: 1.81, 2: 1.97, 3: 2.12; large range: 10: 1.75, 14: 1.78, 20: 1.84), thus leaving less "invisible" trials for the analysis of *t–p* combinations involving larger prime numerosities. Future studies could overcome this trend, which might result in more variable RTs for smaller *t–p* distances, by determining prime contrasts specifically for each prime numerosity. Furthermore, the significant linear function coefficients for the large numerosity range suggest that increased (partial) visibility might slow target processing independently of *t–p* distance, e.g., by introducing noise into the decision process.

We conclude that using the original design used by Bahrami et al. (2010), the question whether distance-dependent numerical priming exists under CFS cannot be answered with certainty. To overcome the pervasive effect of target numerosity when investigating effects of *t–p* distances from −2 to +2, one must (a) choose targets larger than 2, and (b) increase the range of prime numerosities so that all *t–p* distances are present for all targets. In such a design, effects of target numerosity can be fully accounted for.

In Experiment 2, we therefore investigated the influence of prime numerosities 1–6 on target numerosities 3 and 4. This design allowed us to fully separate distance-dependent priming effects from any target-specific RT effects because all *t–p* distances within the full range from −2 to +2 were available for both targets. However, to keep the experiment as

comparable as possible with the original setup used by Bahrami et al. (2010), non-symbolic targets consisted of Gabor patches superimposed on dynamic CFS masks as in the first experiment.

3. Experiment 2

3.1. Methods and materials

3.1.1. Participants

Due to a lack of further studies on unconscious numerosity processing under CFS [for an exception: (Sklar et al., 2012); see discussion section], determining sample size for Experiment 2 (main effect "*t*–*p* distance") was less straightforward than for Experiment 1 which was based on the design used by Bahrami et al. (2010). While there are previous priming studies with non-symbolic primes and non-symbolic targets ("dot-dot" condition), these either used fully visible primes or primes rendered invisible by backward masking [e.g., (Koechlin et al., 1999; Van Opstal et al., 2008)]. We therefore took a more practical approach in Experiment 2 and simply assumed an effect of medium size ($f = 0.25$). Using G*Power 3.1.9 (Faul et al., 2007) we determined that for $f = 0.25$, and $\alpha = 0.05$, a sample size of $N = 26$ was required to achieve a power of 0.90.

33 new observers participated in this experiment, which was conducted with local ethics approval at the Department of Psychiatry and Psychotherapy, Charité-Universitätsmedizin Berlin, Germany. They were recruited from a student pool via email and paid 8 €/h for their participation. Four participants were excluded because they showed significant above-chance forced-choice discrimination performance for invisible stimuli in the control experiment (48%, 53%, 53%, and 59%; all $p < .05$). One further participant was excluded because he/she did not follow the instructions and responded to the targets only after CFS offset. All remaining 28 participants (20 female, mean age: 24, range: 18–37 years) had normal or corrected-to-normal vision, were naïve to the purpose of the study, and provided informed written consent. ND, LK, KW, and JG collected data.

3.1.2. Methods

Methods were the same as in Experiment 1, except for the following differences. Target stimuli consisted of 3 or 4 vertically oriented Gabor patches. Prime stimuli consisted of 1–6 Gabors (diameter: 0.49–1.14°; spatial frequency: 2.3 cpd; for both targets and primes). In total, we created 3600 arrays of pseudo-randomly positioned Gabor patches (Gebuis & Reynvoet, 2011). Of all 12 possible *t*–*p* combinations, two were not analysed (*t*: 3, *p*: 6; *t*: 4, *p*: 1). A circular space around the fixation (diameter: 0.64°) never contained any prime stimuli. Individually adjusted prime contrast was 0.11 ± 0.03 (in arbitrary units of stimulus alpha level). In the main experiment, a total of 448 trials were presented in blocks of 56 trials (64 prime-absent trials, 384 prime-present trials); participants used two arrow keys on a PC keyboard to indicate the target numerosity (left: 3; right: 4). In the control experiment, a total of 60 trials were presented, and participants used three arrow keys to indicate the prime numerosity (left: 1,2; down: 3,4; right: 5,6). We monitored performance in all control experiments on a participant-by-participant basis in order to decide on her/his exclusion from the sample based on her/his performance in the invisible prime enumeration task. After data collection from eight participants, we added a second control experiment (30 trials) in which primes were presented to both eyes on a gray background without CFS masks. This was done to exclude the scenarios that (a) participants showed low performance in the first control experiment because the control 3AFC task was too difficult, and (b) participants showed low performance in the first control experiment because the individually adjusted prime contrasts were too low. One participant did not complete the second control experiment due to a software problem (but he/she was excluded on the basis of task performance in the first control experiment anyway). The whole experiment lasted approx. 90 min.

3.1.3. Data analysis

We first tested for the effects of target numerosity and identity priming using a 2×2 factorial repeated measures ANOVA with factors "prime presence" (2 levels) and "target numerosity" (2 levels). The ANOVA was based on prime-present conditions with zero *t*–*p* distance ("*t*: 3, *p*: 3", and "*t*: 4, *p*: 4") and the prime-absent conditions ("*t*: 3, no prime", and "*t*: 4, no prime). Degrees of freedom were Greenhouse-Geisser (ε) corrected to account for possible violations of sphericity (Greenhouse & Geisser, 1959). We report partial eta squared (η_p^2) as measure of effect size (SPSS 13.0 for Windows, SPSS Inc.). In addition, the effects were tested using non-parametric Wilcoxon signed-rank tests.

Next, we used multiple regression analysis to test for different shapes of the distance-dependent priming functions, separately for targets 3 and 4. Based on the findings from Experiment 1, we fitted regression equations with only two predictors that coded for a step-function and a V-shaped function, respectively. We report the mean regression coefficients (β) as well as the 95% confidence intervals (CIs) for each function shape predictor (Cumming, 2014).

3.2. Results

3.2.1. Prime visibility and prime numerosity discrimination

Fig. 2B summarizes subjective prime visibility ratings and objective prime discrimination performance in Experiment 2. In the control experiment, primes were predominantly rated as invisible (PAS = 1; 65.95%), intermediate (PAS = 2 or 3) and full visibility ratings (PAS = 4) were less frequent. Forced-choice discrimination performance for invisible primes was not significantly above 3AFC chance level (32.84%; $t_{27} = -.36$, $p = .724$; one-sided *t*-test). When primes were rated as visible

(PAS > 1), performance was higher, but still close to chance level (37.36%; $t_{23} = 1.46$, $p = .079$; four participants gave less than three "visible" ratings and were excluded from this test). In the additional control experiment without CFS, virtually all stimuli (94.44%) were rated as visible (PAS > 1), and discrimination performance for these stimuli was significantly above 3AFC chance level (91.03%; $t_{20} = 25.45$, $p < .001$), suggesting that low performance in the visible trials of the first control experiment was due to partial awareness of the numerosity stimuli (also see Section 4.1). In the main experiment, "invisible" ratings were most frequent in prime-present (68.12%), as well as in prime-absent trials (78.91%).

3.2.2. Target numerosity discrimination

2AFC discrimination of target numerosity was at 92.41 ± 0.97%.

3.2.3. RTs in trials with invisible primes

Of all correct trials, 6.14% were outliers. Fig. 4A summarizes RTs in trials with invisible primes (PAS = 1). In trials without primes ("target only"), RTs to target numerosity 3 were on average 28 ms slower than to target numerosity 4 (709 versus 681 ms). In trials with primes, RTs to target 3 appeared to be more affected by the presentation of invisible primes than RTs to target 4. Specifically, for target 3, RTs followed an approximate V-shape with the fastest RT for *t–p* distance 0 (683 ms) and the slowest RT for *t–p* distance +2 (705 ms). For target 4, RTs showed less distance-dependent variation.

To test for the effects of target numerosity and identity priming, we performed an ANOVA with the factors "target numerosity" and "prime presence" only on trials with zero *t–p* distance. There was no main effect of target numerosity ($F_{1,26} = 2.92$, $p = .099$, $\varepsilon = 1.00$, $\eta_p^2 = .10$), but a significant effect of prime presence on RTs ($F_{1,26} = 6.27$, $p = .019$, $\varepsilon = 1.00$, $\eta_p^2 = .19$) and significant interaction ($F_{1,26} = 10.90$, $p = .003$, $\varepsilon = 1.00$, $\eta_p^2 = .30$; one participant had less than one trial in one of the conditions and was excluded from this test). Non-parametric statistical tests yielded a significant effect of numerosity ($Z = 1.98$, $p = .048$), no significant effect of prime presence ($Z = 1.83$, $p = .067$), but showed a target-specific effect of prime presence (3: $Z = 3.48$, $p < .001$; 4: $Z = .73$, $p = .466$), thus confirming the interaction indicated by the ANOVA.

Fig. 4B plots the results of the multiple regression analysis which tested for different shapes of the priming function (RT against *t–p* distance). Regressions were run separately for targets 3 and 4. For target numerosity 3 (mean $R^2 = .50$), only the regression coefficients of the V-shaped (V) priming function deviated from zero (mean $\beta_V = 7.19$, 95% CI [0.66, 13.72]), while the regression coefficients of the step-like (S) function did not ($\beta_S = -2.46$, [−2.36, 7.27]). For target 4 (mean $R^2 = .57$), none of the regression coefficients deviated from zero ($\beta_V = 0.91$, [−5.15, 6.97]; $\beta_S = -3.32$, [−6.65, 0.02]).

Following the approach by Roggeman et al. (2007), we re-analyzed the regression coefficients for target numerosity 3 without zero *t–p* distance, in order to test whether the V-shaped priming function was mainly determined by identity priming or also reflects distance-related priming (and hence semantic access). After omitting zero *t–p* distance from the analysis (mean $R^2 = .61$), none of the regression coefficients deviated from zero ($\beta_V = 6.40$, [−2.18, 14.99]; $\beta_S = 2.33$, [−2.79, 7.44]). As further explorative analysis, we compared *t–p* distances +1 and +2 for target numerosity 3, but found no significant difference ($t_{26} = 1.64$, $p = .113$; two-sided paired *t*-test).

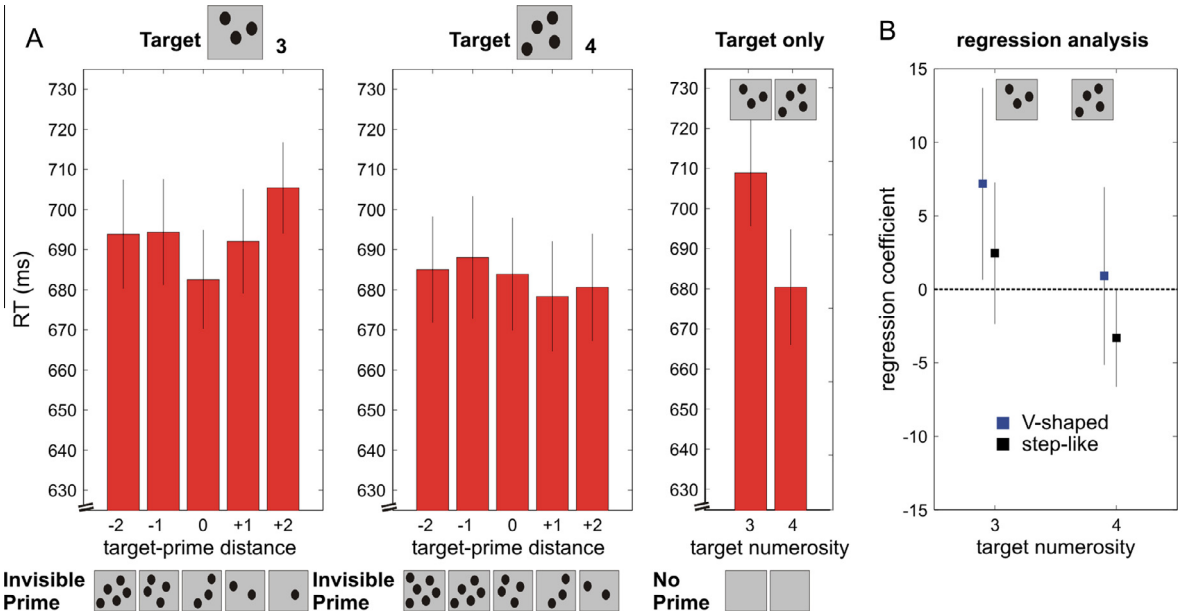

Fig. 4. RTs in Experiment 2. (A) RTs plotted separately for target-prime distances −2, −1, 0, +1, +2 and prime-absent trials ("target only"). Plotted is the mean ± standard error of the mean. (B) Regression coefficients for the predictors describing V-shaped and step-like priming functions, separately for targets 3 and 4. Plotted is the mean and 95% CI.

3.3. Discussion

The optimized experimental design of Experiment 2 resulted in a small V-shaped priming effect, but exclusively for target numerosity 3. When a larger target numerosity was used (target 4), no priming was observed. Thus one might conclude that we found evidence for a distance-dependent numerical priming effect under CFS, which was specific for targets within the upper limits of the subitizing range (Burr et al., 2010; Piazza et al., 2011). However, the priming effect was mainly driven by identity priming, and therefore does not provide conclusive evidence for distance-dependent priming and high-level semantic access under CFS. Further research will be needed to investigate the intriguing possibility of a priming effect limited to targets within the subitizing range. These experiments should also aim to use both symbolic and non-symbolic primes and targets to (a) further investigate the boundary conditions of numerical priming under CFS, and to (b) test the notation-independence of putative numerical priming effects.

Arguably, the priming effect in Experiment 2 has little in common with the effects previously described by Bahrami et al. (2010). The priming effects differ dramatically both in terms of effect size and in terms of the underlying priming function (i.e., RT against *t–p* distance). A V-shaped numerical priming effect, as observed by us, seems to be more in accordance with previous work on number priming (Roggeman et al., 2007; Van Opstal et al., 2008). By contrast, linear priming functions as found by Bahrami and colleagues, have not been reported previously and are not easily compatible with current models of number processing. However, the fact that we observed V-shaped priming with non-symbolic Gabor arrays as primes was indeed unexpected as previous work by Roggeman et al. (2007) showed V-shaped priming for symbolic primes (digits), but step-like priming for non-symbolic primes (dots). The large and overlapping CIs of our regression coefficients also suggest that more data will be needed to better characterize the extent and type of numerical priming which might (or might not) be possible under interocular suppression.

4. General discussion

In two behavioral experiments, we have explored the boundary conditions of distance-dependent numerical priming effects when primes are rendered invisible by interocular suppression (CFS). Using statistical analyses and experimental designs that precluded an effect of target numerosity (Hesselmann & Knops, 2014), we found evidence for identity priming, but no conclusive evidence for distance-dependent numerical priming under CFS. Specifically, the second experiment showed a V-shaped priming effect when the target was within the subitizing range, but this effect was based on identity priming. Our results suggest that previous conclusions on high-level distance-dependent numerical priming under interocular suppression may have been premature. We will briefly discuss some remaining questions as well as implications of our findings.

4.1. Priming effects in trials with visible primes

In both experiments, participants gave "visible" PAS ratings (>1) in a considerable percentage of trials (49.19%, and 31.88%, respectively). In an exploratory analysis, we therefore looked for evidence of priming effects in these trials but found none (data not shown). We propose that the reason for this is as follows. When stimuli "break through" or escape the suppression by CFS, they often do so in parts [e.g., (Hesselmann & Malach, 2011)]. In our paradigm using non-symbolic primes, this would mean that only a few of the total number of Gabors get access to conscious perception. Therefore, any potential distance-dependent priming effects based on visible primes would be abolished. Under full prime visibility (PAS = 4), however, we would expect to observe significant priming effects, given that participants used this rating appropriately, i.e., only in the case when all targets were clearly visible. Unfortunately, there were not sufficient trials for this analysis, since in both experiments only 10.68 and 3.22%, respectively, of primes were rated as PAS = 4.

4.2. Unconscious numerical priming with CFS

Apart from the study by Bahrami et al. (2010), there is only one other CFS study on unconscious processing of numerical information. In three experiments, Sklar et al. (2012) investigated whether the solutions of unconsciously presented single-digit equations with three terms but no result (e.g., "9-3-4=") would prime the enumeration of a visible target number. Target numbers were either congruent with the result of the equation (here: 2) or incongruent. Surprisingly, the authors found a significant congruency effect (~10–20 ms) for invisible subtraction primes but not for addition primes. For addition primes, the authors observed significant congruency effects only when single-digit equations with two terms (e.g., "8+7=") were unconsciously presented and participants had to report whether a subsequently presented visible addition equation with two terms and result (e.g., "9+6=15") was correct or not (both correct and incorrect equations were presented). In case of addition priming, only trials in which the correct solution to the equation was presented were used in the analysis, which showed that participants made slightly fewer mistakes in congruent (3.2%) than in incongruent (4.4%) trials.

While these results seem to suggest that effortful arithmetic equations can be solved unconsciously, the exact nature of the priming effects remains unclear. For example, Sklar et al. (2012) report congruency priming (i.e., numerical identity priming) for subtractions, but do not address the question of whether any distance-dependent priming effects were observed in their data (e.g., was there a larger priming effect for "9-3-4=" than for "9-1-2=", when the target was 3?). Clearly, such

effects should be expected under the assumption of representational overlap between the prime and the target (Van Opstal et al., 2008). Future investigations of these intriguing priming effects should also aim to overcome one methodological shortcoming which is common to many CFS studies but particularly apparent in the study by Sklar et al. (2012). In their three experiments on unconscious arithmetics, a large percentage of participants were excluded based on offline awareness checks involving objective forced-choice tasks and subjective debriefing (Exp. 6: 60%; Exp. 7: 54%; Exp. 9: 41%, including exclusions due to ceiling effects). One could argue that the need to exclude so many participants indicates a rather high level of stimulus visibility prior to exclusion. The participants finally remaining in the sample therefore constitute a biased sample, and it is not clear whether all of them truly failed to perceive the stimulus or were selected due to chance factors.

4.3. Unconscious numerical priming with other techniques

In line with the results of Bahrami et al. (2010), a robust impact of non-consciously perceived numerical information has been demonstrated in healthy participants with a variety of techniques beyond CFS, such as crowding (Huckauf, Knops, Nuerk, & Willmes, 2008) and masking (Dehaene et al., 1998; Naccache & Dehaene, 2001), as well as in patients suffering from hemineglect (Sackur et al., 2008). These results were taken as evidence for an elaborate processing of non-consciously presented numerical information. However, these results have also been discussed critically. For example, Kunde, Kiesel, and Hoffmann (2003) argued that the observed priming effects were rather due to a "match with pre-specified cognitive action-trigger conditions" (p. 223). Importantly, while being based on semantic categorization the assumed action-trigger conditions do not necessarily require a semantic elaboration of the stimulus at hand. Recent neuroimaging results suggest that areas within dorsal stream (i.e., posterior parietal cortex) may be organized in a map-like architecture, representing visual items in a salience/priority map which flexibly adapts its response profile to the task at hand (Knops et al., 2014). Reflecting well-known capacity limits in different cognitive regimes activity in these areas was markedly different for identical numerosities in different task contexts, refuting the idea of a purely stimulus-driven, automatic extraction of numerosity information from the visual scene in dorsal stream. The current results are in line with evidence against an automatic semantic elaboration of numerical information under all circumstances and may help defining the scope and limits of a visual sense of number (Burr & Ross, 2008).

4.4. Privileged access of visual information to dorsal stream processing under CFS?

The study by Bahrami et al. (2010) was motivated by functional magnetic resonance imaging (fMRI) results suggesting that there might a "dorsal CFS bias", i.e., that there is some form of *privileged* access of suppressed information to the dorsal visual stream, and parietal cortex in particular. Specifically, one influential fMRI study showed preserved dorsal processing of potentially action-relevant tool images in the absence of perceptual awareness (Fang & He, 2005). Based on the functional overlap of visuospatial and numerical processing in human parietal cortex, Bahrami et al. (2010) hypothesized that numerosity judgments should exhibit unconscious priming during interocular suppression by CFS. However, the results by Fang and He (2005) have so far not been confirmed in other CFS studies (Fogelson et al., 2014; Hesselmann & Malach, 2011; Ludwig, Kathmann, Sterzer, & Hesselmann, 2014). It still remains unclear what differences in study design and analysis might explain the divergent results.

In a series of priming experiments, Almeida and colleagues reported faster RTs to visible tools when they were preceded by invisible tools rather than invisible animals in a categorical decision task (Almeida, Mahon, Nakayama, & Caramazza, 2008). Importantly, no such effect was demonstrated when participants had to respond to pictures of animals; furthermore, the category-specific priming effect was limited to CFS. Again, these results were interpreted as evidence for a privileged access of suppressed information to the dorsal visual stream under CFS. Follow-up studies suggested that the priming effects were in fact based on elongated shape and not on tool category [(Sakuraba, Sakai, Yamanaka, Yokosawa, & Hirayama, 2012), also see (Hebart & Hesselmann, 2012)], a finding which seems harder to reconcile with the notion of preserved action-relevant representations in parietal cortex under CFS. Note that prime durations were considerably longer (1000–3000 ms) in the numerical priming experiments by Bahrami et al. (2010) than in most previously mentioned category priming experiments (~200 ms), which might have negatively influenced the numerical priming effects under interocular suppression (Barbot & Kouider, 2012).

Another line of behavioral research investigated whether participants could learn to reach for visible stimuli which were rendered invisible by CFS. While one study indeed showed such learning effects (Roseboom & Arnold, 2011), another study showed that participants learned to better detect invisible targets but did not improve in a grasping task (Ludwig et al., 2013). More work will be needed to elucidate the exact conditions under which healthy participants can be trained to grasp invisible objects.

Together, we conclude that there is so far no compelling neuroimaging or behavioral evidence for a privileged access of visual information to dorsal stream processing under CFS.

4.5. Concluding remarks

Thus far, the rapidly accumulating evidence regarding the degree of unconscious processing that occurs during CFS has been mixed. To resolve the many discrepancies in the obtained results, we believe that future CFS studies will need to not

primarily focus on the scope and limits of unconscious processing, but aim to further investigate its boundary conditions by taking theoretical considerations as well as neural underpinnings more into account.

Acknowledgments

GH, PS & AK were supported by the German Research Foundation (Grants HE 6244/1-1, STE 1430/2-1, STE 1430/6-1, KN 959/2-1). We report all data exclusions, manipulations, and measures, and how we determined our sample sizes. We thank Kerstin Wolf, Leon Kerger, and Julia Gerb for their invaluable help with data acquisition.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.concog.2014.10.009.

References

Almeida, J., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *PNAS, 105*, 15214–15218.
Bachmann, T., Breitmeyer, B. G., & Ogmen, H. (2007). *The experimental phenomena of consciousness: A brief dictionary*. Oxford: University Press.
Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science, 21*, 224–233.
Barbot, A., & Kouider, S. (2012). Longer is not better: Nonconscious overstimulation reverses priming influences under interocular suppression. *Attention, Perception, & Psychophysics, 74*(1), 174–184.
Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
Breitmeyer, B. G. (2014). *The visual (Un)Conscious and its (Dis)Contents – A microtemporal approach*. New York: Oxford University Press.
Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology, 18*(6), 425–428.
Burr, D., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*(6), 20.
Cave, C. B., Blake, R., & McNamara, T. P. (1998). Binocular rivalry disrupts visual priming. *Psychological Science, 9*, 299–302.
Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*(1), 7–29.
Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: a neuronal model. *Journal of Cognitive Neuroscience, 5*(4), 390–407.
Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., et al (1998). Imaging unconscious semantic priming. *Nature, 395*(6702), 597–600.
Faivre, N., Berthet, V., & Kouider, S. (2012). Nonconscious influences from emotional faces: A comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Psychology, 3*, 129.
Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience, 10*, 1380–1385.
Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.
Fogelson, S. V., Kohler, P. J., Miller, K. J., Granger, R., & Tse, P. U. (2014). Unconscious neural processing differs with method used to render stimuli invisible. *Frontiers in Psychology, 5*, 1–11.
Gayet, S., Van der Stigchel, S., & Paffen, C. L. (2014). Breaking continuous flash suppression: competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology, 5*, 460.
Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods, 43*(4), 981–986.
Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95–112.
Hebart, M. N., & Hesselmann, G. (2012). What visual information is processed in the human dorsal stream? *Journal of Neuroscience, 32*(24), 8107–8109.
Hesselmann, G. (2013). Dissecting visual awareness with FMRI. *Neuroscientist, 19*(5), 495–508.
Hesselmann, G., & Knops, A. (2014). No conclusive evidence for numerical priming under interocular suppression. *Psychological Science*.
Hesselmann, G., & Malach, R. (2011). The link between fMRI-BOLD activation and perceptual awareness is "stream-invariant" in the human visual system. *Cerebral Cortex, 21*, 2829–2837.
Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences, 9*, 1–23.
Huckauf, A., Knops, A., Nuerk, H. C., & Willmes, K. (2008). Semantic processing of crowded stimuli? *Psychological Research, 72*(6), 648–656.
Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *Journal of Neuroscience, 31*(38), 13535–13545.
Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *American Journal of Psychology, 62*(4), 498–525.
Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible 'invisible'. *Trends in Cognitive Sciences, 9*, 381–388.
Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short term memory and enumeration. *Journal of Neuroscience, 34*, 9857–9866.
Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 1882–1905.
Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*, 857–875.
Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88*(2), 223–242.
Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 16*(1), 149–157.
Ludwig, K., Kathmann, N., Sterzer, P., & Hesselmann, G. (2014). Investigating category- and shape-selective neural processing in ventral and dorsal visual stream under interocular suppression. *Human Brain Mapping*. http://dx.doi.org/10.1002/hbm.22618. [Epub ahead of print].
Ludwig, K., Sterzer, P., Kathmann, N., Franz, V. H., & Hesselmann, G. (2013). Learning to detect but not to grasp suppressed visual stimuli. *Neuropsychologia, 51*(13), 2930–2938.
Miles, W. R. (1930). Ocular dominance in human adults. *Journal of General Psychology, 3*, 412–429.
Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex, 11*(10), 966–974.
Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.
Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition, 121*(1), 147–153.
Ramsoy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences, 3*, 1–23.
Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19*(6), 607–614.

Roggeman, C., Verguts, T., & Fias, W. (2007). Priming reveals differential coding of symbolic and non-symbolic quantities. *Cognition, 105*(2), 380–394.
Roseboom, W., & Arnold, D. H. (2011). Learning to reach for 'invisible' visual input. *Current Biology, 21*(13), R493–R494.
Sackur, J., Naccache, L., Pradat-Diehl, P., Azouvi, P., Mazevet, D., Katz, R., et al (2008). Semantic processing of neglected numbers. *Cortex, 44*(6), 673–682.
Sakuraba, S., Sakai, S., Yamanaka, M., Yokosawa, K., & Hirayama, K. (2012). Does the human dorsal stream really process a category for tools? *Journal of Neuroscience, 32*(11), 3949–3953.
Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences USA, 109*(48), 19614–19619.
Sterzer, P., Stein, T., Ludwig, K., Rothkirch, M., & Hesselmann, G. (2014). Neural processing of visual information under interocular suppression: A critical review. *Frontiers in Psychology*.
Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*, 1096–1101.
Tsuchiya, N., Koch, C., Gilroy, L. A., & Blake, R. (2006). Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry. *Journal of Vision, 6*, 1068–1078.
Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review, 15*(2), 419–425.
Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience, 16*(9), 1493–1504.
Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*, 475–482.
Yang, E., Brascamp, J., Kang, M. S., & Blake, R. (2014). On the use of continuous flash suppression for the study of visual processing outside of awareness. *Frontiers in Psychology, 5*, 724.
Zimba, L. D., & Blake, R. (1983). Binocular rivalry and semantic processing: Out of sight, out of mind. *Journal of Experimental Psychology: Human Perception and Performance, 9*(5), 807–815.

OPEN ACCESS Freely available online

PLOS ONE

In How Many Ways is the Approximate Number System Associated with Exact Calculation?

Pedro Pinheiro-Chagas[1,2,3]*, Guilherme Wood[4], André Knops[5], Helga Krinzinger[6], Jan Lonnemann[7,8], Isabella Starling-Alves[3], Klaus Willmes[9], Vitor Geraldi Haase[3,10,11]

1 Cognitive Neuroimaging Unit, Institut National de la Santé et de la Recherche Médicale (INSERM) U992, Gif/Yvette, France, **2** NeuroSpin Center, Institute of BioImaging Commissariat à l'Energie Atomique (CEA), Gif/Yvette, France, **3** Developmental Neuropsychology Laboratory, Department of Psychology, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil, **4** Department of Neuropsychology, Institute of Psychology, Karl-Franzens University of Graz, Graz, Austria, **5** Faculty of Life Sciences, Humboldt Universität zu Berlin, Berlin, Germany, **6** Section Child Neuropsychology, Department of Child and Adolescent Psychiatry, University Hospital of the RWTH, Aachen, Germany, **7** German Institute for International Educational Research, Frankfurt, Germany, **8** Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt, Germany, **9** Section Neuropsychology, Department of Neurology, University Hospital of the RWTH, Aachen, Germany, **10** Programa de Pós-graduação em Neurociências, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil, **11** Programa de Pós-graduação em Saúde da Criança e do Adolescente, Faculdade de Medicina, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

Abstract

The approximate number system (ANS) has been consistently found to be associated with math achievement. However, little is known about the interactions between the different instantiations of the ANS and in how many ways they are related to exact calculation. In a cross-sectional design, we investigated the relationship between three measures of ANS acuity (non-symbolic comparison, non-symbolic estimation and non-symbolic addition), their cross-sectional trajectories and specific contributions to exact calculation. Children with mathematical difficulties (MD) and typically achieving (TA) controls attending the first six years of formal schooling participated in the study. The MD group exhibited impairments in multiple instantiations of the ANS compared to their TA peers. The ANS acuity measured by all three tasks positively correlated with age in TA children, while no correlation was found between non-symbolic comparison and age in the MD group. The measures of ANS acuity significantly correlated with each other, reflecting at least in part a common numerosity code. Crucially, we found that non-symbolic estimation partially and non-symbolic addition fully mediated the effects of non-symbolic comparison in exact calculation.

Citation: Pinheiro-Chagas P, Wood G, Knops A, Krinzinger H, Lonnemann J, et al. (2014) In How Many Ways is the Approximate Number System Associated with Exact Calculation? PLoS ONE 9(11): e111155. doi:10.1371/journal.pone.0111155

Editor: Daniel Ansari, The University of Western Ontario, Canada

Received February 12, 2014; **Accepted** September 23, 2014; **Published** November 19, 2014

Funding: This research was granted by the CAPES/DAAD Probral Program, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 307006/2008-5, 401232/2009-3) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, APQ-02755-SHA, APQ-03289-10). Vitor G. Haase is supported by CNPq fellowships (308157/2011-7 and 307975/2010-0, respectively). Guilherme Wood is supported by a FWF research project (nr. P22577). Pedro Pinheiro-Chagas has a Science Without Borders Fellowship from CNPq (nr. 246750/2012-0). CNPq: http://cnpq.br/FAPEMIG: http://www.fapemig.br/FWF: http://www.fwf.ac.at/The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: ppinheirochagas@gmail.com

Introduction

An extensive literature that comprises psychophysical [1,2], electrophysiological [3], and neuroimaging data [4] has demonstrated that human infants and adults share an approximate number system (ANS), which is dedicated to representing large magnitudes in an analog fashion. Number representation within the ANS is very similar to the intuition that we have for space and time magnitudes [5] and can be well described by Weber-Fechner's law [3,4,6,7]. Because the ANS is already present in newborns [8] and interacts with culturally derived symbolic representations during development [9], it is considered to be an important *start-up* tool for the acquisition of mathematical knowledge [10].

Converging evidence from correlational [11], cross-sectional [12,13], longitudinal [14–18] and training studies [19] has provided robust support for the link between the ANS and arithmetics. Halberda et al. [11] showed that the Weber fraction calculated from a non-symbolic number comparison task in adolescents retroactively correlated with standardized math achievement scores from Kindergarten up to the sixth grade. A series of other studies replicated this finding using not only general standardized math achievement scores [18,20–22] but also simple arithmetics operations [13,23].

Importantly, a number of other studies failed to find an association between non-symbolic comparison and math achievement, but rather found significant associations between math achievement and the symbolic version of the task (see [24] for a review of the inconsistency of those findings).

However, two recent meta-analyses confirmed the existence of a robust association between non-symbolic comparison and math achievement from childhood to adulthood. Fazio, Bailey, Thompson and Siegler [25] analyzed 19 published studies and found that although non-symbolic processing is less strongly correlated with math achievement compared to symbolic processing, there is a robust and specific significant association between non-symbolic comparison and math achievement. Chen and Li [26] investigated 36 cross-sectional studies and found that the association between

non-symbolic comparison and math achievement is moderate but statistically significant (r = 0.20, 95% CI = [0.14, 0.26]), even after controlling the effect of general cognitive abilities. Importantly, non-symbolic comparison was found to prospectively predict later math performance (r = 0.24, 95% CI = [0.11, 0.37]; 6 samples) and it is also retrospectively correlated to early math achievement (r = 0.17, 95% CI = [0.07, 0.26]; 5 samples). Based on the estimated effect sizes, the authors conducted power analyses and confirmed that many previous studies failed to find a significant association between non-symbolic comparison and math achievement because of insufficient statistical power due to small sample sizes.

Moreover, other measures of ANS acuity, such as number estimation, were also found to correlate with math achievement in children and adolescents [12,16,27].

Noticeably, longitudinal and training studies have provided evidence for a foundational role of the ANS on the development of math abilities. Using the non-symbolic number comparison task, Mazzoco et al. [18] showed that the ANS acuity measured prior to formal mathematical instruction was selectively predictive of arithmetics achievement in the first grade (see also Libertus et al. [17]). Similarly, Gilmore et al. [15] found that non-symbolic calculation abilities measured in a group of Kindergarten children were a robust predictor of later math achievement. Complementary, Park and Brannon [19] showed that training adults in a non-symbolic addition and subtraction task specifically improves exact addition and subtraction. Interestingly, the ANS acuity was also found to improve with mathematical education. Piazza, Pica, Izard, Spelke, and Dehaene [28] investigated a group of Mundurukus, an indigenous population in Brazil that does not have a system for representing exact numbers [29], and found that the ANS acuity, as quantified by a non-symbolic number comparison task, was modulated by the level of formal instruction at the standard Brazilian school system. This result provides support for a bidirectional association between the most basic forms of number processing and math abilities. Importantly, an analogous bidirectionality has long been found in the reading domain, such as the fact that phonological abilities serve as the base for reading competence and are improved by literacy [30,31].

Finally, group studies demonstrated that children with developmental dyscalculia (DD), a learning disability specific to calculation, have an impaired ANS compared to their typically achieving (TA) peers. Piazza et al. [13] showed that the ANS acuity in children with DD at 10 years old, as quantified by the internal Weber fraction, was equivalent to the acuity observed in TA Kindergarten children. Similar results were obtained by Mazzoco et al. [12], who showed not only that adolescents with DD present higher internal Weber fractions than their TA peers but also that they have an impairment in estimating numerical magnitudes.

Other studies that investigated younger children with DD found only an impairment in the symbolic version of the number comparison task (see review by Noël & Rousselle [32]), which casts doubt on the assumption of a critical role of the ANS in the acquisition of exact number representations. Chu, van Marle, and Geary [33] found that the ANS acuity significantly predicted the risk for DD in children, but measures of symbolic number knowledge were more robust predictors. However, those studies used only one measure to assess the ANS acuity: the non-symbolic number comparison task. Based on Gilmore et al. [15] it could be the case that different forms of approximate manipulation of numerical information, such as calculation, could be additional important predictors of risk for DD.

Although much progress has been achieved in the establishment of an association between the ANS and arithmetics, it remains largely elusive in how many ways the ANS interacts with exact calculation and through which cognitive mechanisms this association could be grounded. The ANS allows for comparing two different magnitudes, to approximately grasp how many objects are present in a scene and to manipulate quantities using simple operations such as addition and subtraction [34]. In this sense, different tasks have been used to measure the ANS acuity, such as comparison [11–13,35], estimation [12,23,27,36] and approximate calculation [15,37,38]. Importantly, very little attention was given to the fact that these measures are tapping different instantiations of the ANS. Comparison, estimation and approximate calculation, although possibly operating at the same level of representation (the ANS), involve very different computational processes and consequently could have specific contributions to the development of exact number representations and mathematics. Indeed, Mazzocco et al. [12] found that non-symbolic comparison and estimation accounted for unique proportions of the variance when predicting math achievement.

Moreover, given the complexity of arithmetics, the link between basic number processing (e.g., magnitude comparison) and exact calculation is possibly not direct and might involve the recruitment of other cognitive processes. Indeed, the study by Lyons and Beilock [39] found that the ability to identify the order of a series of digits fully mediated the association between the ANS acuity (as measured with the non-symbolic number comparison task) and exact calculation in adults. Importantly, van Marle, Chu, Li and Geary [40] provided a conceptual replication of the study of Lyons and Beilock [39] in children, and proposed that the ANS acuity facilitated the early acquisition of symbolic number knowledge and was indirectly associated with math achievement through this knowledge. In line with these findings, it might also be the case that there is a type of hierarchical association between different instantiations of the ANS, from the most elementary abilities to more complex operations and manipulations of magnitude information. That is, non-symbolic estimation and calculation could be intermediate steps between simple number discrimination and exact calculation.

Surprisingly, to date there is only one study in adults and no study with children that directly compared different measures of the ANS. Gilmore, Attridge, and Inglis [41] measured the ANS with non-symbolic versions of the number comparison and approximate addition tasks and found null correlations between the internal Weber fractions calculated from each task, placing in doubt the assumption of a single underlying ANS. However, this result is very puzzling and deserves further examination, because both tasks used non-symbolic magnitudes and, even though different cognitive mechanisms might be recruited during performance, both tasks should at least partially activate the representation of numbers and its underlying brain circuitry. Indeed, using a conjunction analysis, Park, Park and Polk [42] recently showed that non-symbolic comparison and non-symbolic addition activated common brain circuitries in the right parietal cortex.

Therefore, a more comprehensive investigation of the association between different instantiations of the ANS (comparison, estimation and calculation), their cross-sectional trajectories in children with typical and atypical math abilities and how they interact with exact calculation is needed.

The present study

Measures that are related to the ANS acuity appear to be normally distributed in the population [11] and are systematically associated with arithmetics achievement. In this sense, the present

study first addressed the hypothesis that children with *math difficulties* (MD) who were selected according to a relatively liberal criterion (below the 25th percentile on a standardized math achievement test [43]) would present with lower ANS acuity compared to their TA peers. To this end, we calculated specific psychophysical parameters for each of three different tasks as indices of ANS acuity. The internal Weber fraction (w) [1] was calculated for the non-symbolic number comparison task, and the coefficient of variation (cv) was calculated for the non-symbolic estimation and non-symbolic addition tasks. The cv is a normalized measure of dispersion of a probability distribution and it is defined as the ratio of the standard deviation to the mean. Therefore, like the w, the higher the cv, the lower the precision. Based on the previous results obtained by Mazzocco et al. [12] and Piazza et al. [13], we expected TA children to have higher ANS acuity (lower values in the psychophysical parameters) compared to children with MD. Second, as noted by Noël and Rousselle [32], one should expect to find differences between the TA and MD groups in the cross-sectional trajectories of the ANS.

More specifically, group differences in ANS acuity, at least as measured by non-symbolic number comparison, should have a trend to increase across development. We finally tested the degree of association between the measures of ANS acuity and exact calculation. Because the psychophysical parameters extracted from the tasks that measure the ANS acuity are at least partially related to the degree of noise in the representation of numerosity, they should be positively correlated to one another. Moreover, based on previous studies [11,18], it is expected that the ANS acuity will have a specific impact on exact calculation, even after controlling for the effects of general developmental factors and other abilities that are related to mathematics, such as language. Crucially, based on the results by Lyons and Beilock [39], who showed that number ordering fully mediated the effect of non-symbolic comparison in exact calculation, we further investigated the relationship between the ANS acuity and calculation using mediation models. Six mediation models were estimated with all possible permutations between measures of ANS acuity as predictors or mediators and exact calculation as the outcome.

Materials and Methods

Participants

This study was approved by the ethics review board of the Federal University of Minas Gerais, Brazil (COEP–UFMG). Informed consent was obtained in written form from the parents and orally from the children. Children from first to sixth grade were recruited from public and private schools in Brazil and were assigned to different groups according to their performance in the Arithmetics and word Spelling subtests of the Brazilian School Achievement Test (Teste de Desempenho Escolar, TDE [44]). The typical achievement group (TA, n = 172) was composed of children who performed above the 25th percentile in both the Arithmetics and Spelling subtests of TDE. The mathematical difficulties group (MD, n = 45) performed below the 25th percentile in the Arithmetics and above that in the Spelling subtest of the TDE.

There were no statistically significant differences in age and sex between groups. All of the children had normal intelligence, as measured by Raven's Colored Progressive Matrices (IQ scores above 85).

Children were assessed using an exact calculation task comprising addition, subtraction and multiplication, a simple reaction time task and three tasks that measured the ANS acuity: non-symbolic comparison, non-symbolic estimation and non-symbolic addition (see the detailed description of the tasks below).

A subgroup of 10 children from the TA and 5 from the MD group were excluded from further analyses, because either they had a poor fit (R^2) for estimation of the w on the non-symbolic comparison task ($R^2<0.2$), and/or they showed a w that exceeded the limit of discriminability of the comparison task ($w>0.6$). The final sample was composed of 162 TA children and 40 children with MD. The subject details are presented in Table 1 (for the descriptive data of the individual assessment samples by grade, see Table S1 in the Supporting Information).

Tasks

The Brazilian School Achievement Test. The TDE [44] is the most widely used standardized test of school achievement that has norms for the Brazilian population (see also Oliveira-Ferreira et al. [45]). We used the Arithmetics and Spelling subtests, which can be applied in groups. Norms are provided for school-aged children between first and sixth grade. The Arithmetics subtest is composed of three simple verbally presented word problems (e.g., "If you had three candies and received four, how many candies do you have now?") and 35 written arithmetic calculations of increasing complexity (e.g., very easy: "4–1"; easy: "1230+150+1620"; intermediate: "823 * 96"; hard: "3/4+2/8"). The Spelling subtest constitutes a dictation of 34 words that have increasing syllabic complexity (e.g., "toca"; "balanço"; "cristalização"). The reliability coefficients (*Cronbach's* α) of the TDE subtests are 0.89 or higher. The children were instructed to work on the problems to the best of their capacity but without time limit.

Raven's Coloured Progressive Matrices. General intelligence was assessed with the Raven's Coloured Progressive Matrices, according to Brazilian norms [46].

Exact Calculation. The task was divided in two sets of items: symbolic and written verbal calculations. The symbolic calculation set was composed of additions (27 items), subtractions (27 items) and multiplications (28 items). Problems that were printed on separate sheets of paper. Children were instructed to answer as fast and as accurately as possible. Arithmetic operations were balanced at two levels of complexity and were presented to children in separate blocks: one block was composed of simple arithmetic table facts and the other block had more complex facts. Simple addition items had results below 10 (e.g., 3+5), while complex addition results ranged between 11 and 17 (e.g., 9+5). Tie problems (e.g., 4+4) were not considered for addition. Simple subtraction was composed of problems in which the operands were below 10 (e.g., 9–6), while in complex subtractions, the first operand ranged from 11 to 17 (e.g., 16–9). No negative results were included in the subtraction problems. Simple multiplication constituted operations that had results below 25 or that had the number 5 as one of the operands (e.g., 2 * 7, 5 * 6), whereas for the complex multiplication, the result of the operands ranged from 24 to 72 (e.g., 6 * 8). Tie problems were not used for multiplication. The time limit per block was set to 1 minute. The written verbal calculation set was composed of four additions and eight subtractions with single-digit operands (e.g. "Isabella has 9 cents. She gives 3 to Pedro. How many cents does Isabella have now?"). Problems were presented to children on a sheet of paper and read aloud by the examiner to avoid reading proficiency bias. The child had to solve the problems mentally and write down the answer in Arabic format as fast and as accurately as possible. The time limit per problem was 1 minute. The total score was calculated as a simple sum of all correct answers combining both symbolic and written verbal items (max score = 94). The task was highly reliable (all *Cronbach's* $\alpha>0.90$) [45,47].

Table 1. Descriptive data of the individual assessment sample.

Categorical Variables	TA (n = 162)		MD (n = 40)		χ^2	*df*	*p-value*	
Sex (% female)	59.26		52.50		0.601	1	0.274	
School type (% public)	86.42		87.50		0.032	1	0.547	
Continuous Variables	**Mean**	**SD**	**Mean**	**SD**	*t*	*df*	*p-value*	*d*
Age (months)	121.562	13.599	118.325	16.513	1.290	200	0.199	0.228
Raven (IQ score)	110.612	10.545	103.619	9.659	3.817	200	<0.001	0.674
TDE Arithmetics	108.924	11.406	85.784	5.073	12.510	200	<0.001	2.209
TDE Spelling	110.078	8.130	101.220	8.929	6.050	200	<0.001	1.068

TA: typically achieving; MD: mathematical difficulties. Both TDE Arithmetics and TDE Spelling scores are in a standardized form with mean = 100 and SD = 15; d = Cohen's d.

doi:10.1371/journal.pone.0111155.t001

Simple Reaction Time. The computerized Reaction Time (RT) task was a simple task in which a picture of a wolf (height = 9.31 cm; length = 11.59 cm) was displayed in the center of a black screen for a maximum time of 3,000 ms [47]. Upon appearance of the wolf on screen, children were instructed to press the spacebar on the keyboard at the moment they saw the wolf, as fast as possible. Trials terminated with the first key press. The task had 30 trials, with an inter-trial interval of 2,000 ms, 3,500 ms, 5,000 ms, 6,500 ms or 8,000 ms. This task was used to control for possible differences in basic processing speed that were not related to numerical tasks.

Non-symbolic Comparison. In the non-symbolic comparison task, participants were instructed to compare two sets of black dots, which were simultaneously presented in two white circles on the left and on the right side of the screen, and they were instructed to choose the larger numerosity by pressing a key congruent to its side (left or right) (see Figure 1) [45,47,48]. Black dots were presented on a white circle against a black background. On each trial, one of the two white circles contained 32 dots (reference numerosity), and the other contained 20, 23, 26, 29, 35, 38, 41, or 44 dots. Each numerosity was presented eight times, and every presentation was arranged in a different configuration. The task comprised 64 testing trials. The maximum stimulus presentation time was 4,000 ms, and the intertrial interval was 700 ms. Between trials, a fixation point appeared on the screen for 500 ms; the fixation point was a cross printed in white and that had 3 cm for each line. To prevent the use of non-numerical cues, the sets of dots which represent the non-symbolic numerosities were designed and generated using a MATLAB script [49] such that on half of the trials, dot size remained constant, and total dot area covaried positively with the numerosity; on the other half of the trials, total dot area was held constant and dot size covaried negatively with numerosity. The data were trimmed for each child to exclude responses of ± 3 SD from the individual mean RT. The w was calculated for each child as a measure of ANS acuity, based on the Log-Gaussian model of the number representation [1], with the methods described by Piazza et al. [4].

Non-symbolic Estimation. In the non-symbolic estimation task, participants were asked to estimate, with a verbal response, the quantity of dots shown on the computer screen [48] (see Figure 1). Black dots were presented on a white circle against a black background. The numerosities were 10, 16, 24, 32, 48, 56 or 64 dots. Each numerosity was presented 5 times, every time in a different configuration such that the same numerosity never appeared in consecutive trials. The task comprised 35 testing trials. To avoid counting, the maximum stimulus presentation time was set to 1000 ms. As soon as the child responded, the examiner, who was seated next to the child, pressed the spacebar on the keyboard and typed the child's answer. Between individual trials, a fixation point appeared on the screen, which was a cross printed in white, with 3 cm for each line. To prevent the use of non-numerical cues, the sets of dots were generated using MATLAB, in such a way that dot size changed but total dot area in a given set was always fixed across the stimuli. Thus, the total occupied area could not serve as a cue for distinguishing between the different numerosities. As a result of this manipulation, the average item size covaried inversely with numerosity. To avoid memorization effects due to the repetition of a specific stimulus, on each trial, the stimuli were randomly chosen from a set of 10 precomputed images with the given numerosity. The data were trimmed for each subject, to exclude the responses ± 3 SD from the mean chosen value across all of the trials. As a measure of ANS acuity, we calculated the mean cv of the responses for each child.

Non-symbolic Addition. The non-symbolic addition task was based on Knops, Viarouge, and Dehaene [50] (see Figure 1). Participants were instructed to solve approximate addition problems with operands presented in a non-symbolic notation (dots patterns). To adapt the paradigm for the use of children, the addition task was embedded in a small history of a monkey having a box of balls. Hence, a trial started with the presentation of the monkey's face, which was followed by the appearance of a brown box against a black background and the first set of dots that moved into the box. Next, another set of dots moved into the same box. Afterward, the box disappeared from the screen and was replaced by the top-view of five boxes that contained different numerosities. The boxes were arranged in a circular manner around the middle of the screen. The children were to choose which numerosity was the closest to the correct outcome by clicking with the left mouse button on the respective box. The task comprised 2 learning trials and 32 testing trials. In the training trials, the boxes were framed after each response. In a case in which the response was correct, the frame was green, which indicated that the child had chosen the box with the correct number of balls. If the response was incorrect, then the frame was red, and the children were instructed to choose another box. This procedure was repeated until the child had chosen the correct box. Before starting the testing, the children were asked if they had understood the task, and if not, the training was repeated until they confirmed that they understood the task. In the testing period, the childrens' choices were indicated by a neutral blue frame around the chosen box, regardless of whether the response was correct or not. All of the addition problems added up to four possible results (i.e., 10, 16, 26 and 40), which

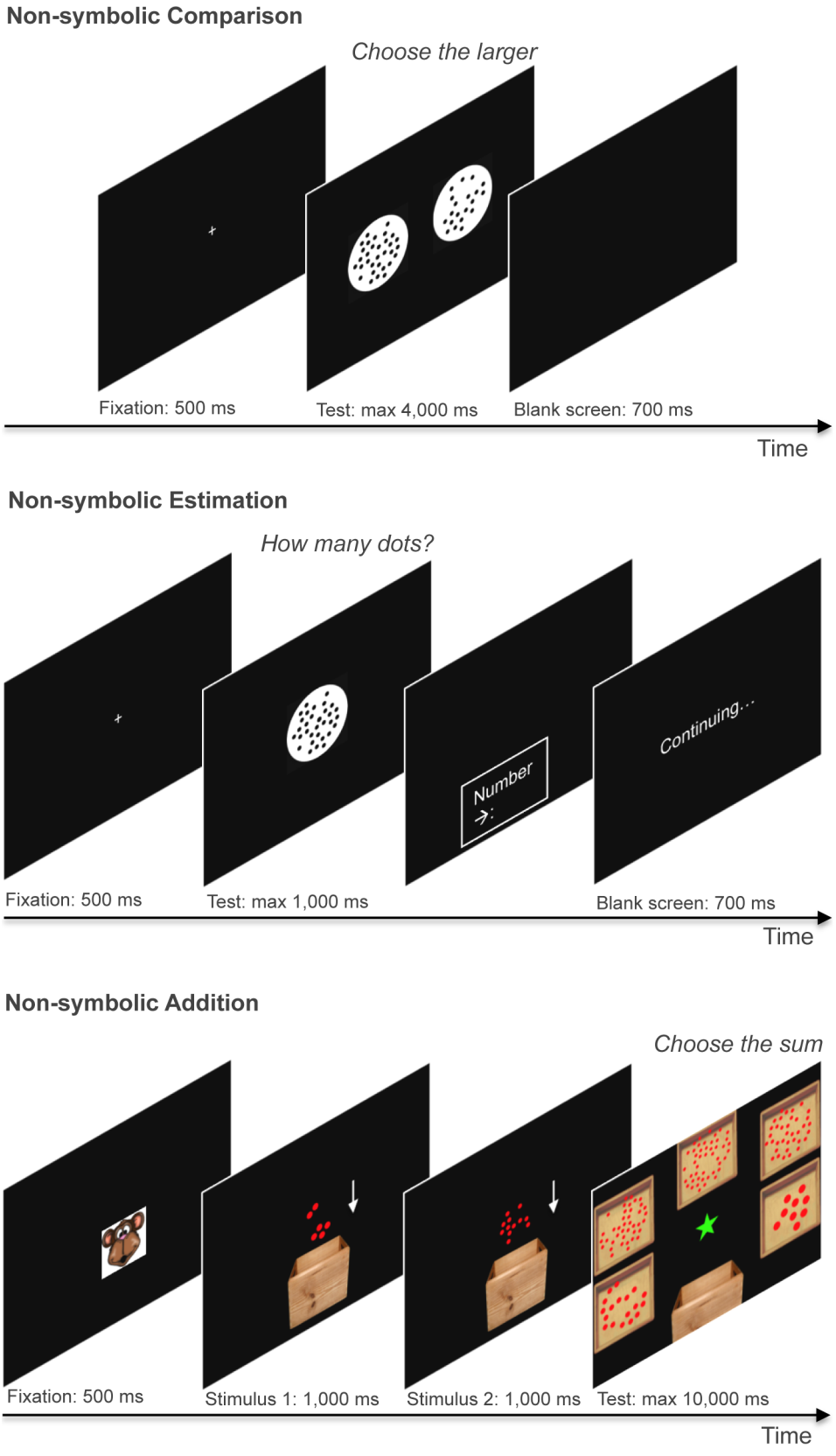

Figure 1. Psychophysical tasks used to measure ANS acuity, with non-symbolic comparison, non-symbolic estimation and non-symbolic addition. The white arrows are used in the bottom picture to illustrate the movement of the dots into the box.
doi:10.1371/journal.pone.0111155.g001

combined ten different operands (4, 5, 6, 8, 10, 13, 14, 18, 20 or 26). To prevent the subjects from memorizing the problems, the operands were randomly "jittered" by adding a random value r, with $r \in J$ and $J = \{-1,0,1\}$. For each correct response, 7 response alternatives were generated as round ($c \times 2.5^{i/3}$), where c is the correct result and i ranges from −3 to +3. To discourage the use of non-arithmetic strategies, such as "Always choose a response alternative in the middle of the presented range", only five of the seven possible results were presented in a trial, such that, in half of the trials, the presented results were the upper five (high range), and thus, the correct response was the second largest numerosity. In the other half of the trials, the lower five results were shown (low range), and the correct response was the fourth largest numerosity. To prevent the use of non-numerical cues, the sets of dots were generated using MATLAB, in such a way that dot size covaried inversely with numerosity. To avoid memorization effects due to the repetition of a specific stimulus, on each trial, the stimuli were randomly chosen from a set of 10 precomputed images with the given numerosity. The data were trimmed for each subject, to exclude the responses ±3 SD from the mean chosen value across all of the trials. As a measure of ANS acuity, we calculated the mean *cv* of the four different results.

Analyses

Initially, the TA and MD groups were compared with regard to exact calculation and the three measures of ANS acuity. Next, the cross-sectional trajectories of the ANS were investigated by calculating the slopes of the regressions between ANS acuity and age for each group separately. Finally, the association between the measures within the ANS and between the ANS acuity and the exact calculation was investigated in three steps. First, cross-correlations of the measures of ANS acuity and exact calculation were determined. Second, to estimate the specific contributions of ANS acuity measures to explain exact calculation, multiple regression models were conducted with exact calculation as the dependent variable and the three measures of ANS acuity as the predictor variables, regressing out the effects of age, schooling, general intelligence and spelling abilities. Finally, to investigate more deeply the possible mediation effects between the ANS instantiations and exact calculation, six mediation models were estimated with all of the possible permutations between the measures of ANS acuity as predictors or mediators and exact calculation as the outcome. All of the statistical analyses were performed using R statistical software [51]. Raw data is available in the Supporting Information (Data S1).

Results

First, we verified whether the children's performances in the measures of ANS acuity followed Weber's law. In the non-symbolic comparison task, we calculated the R^2 of the fitting procedure to calculate the w for each child. In both the TA and MD groups, the R^2 values were high (TA: mean = 0.883, SD = 0.082; MD: mean = 0.849, SD = 0.108), which indicates that the children's performances were well described by the *Log-Gaussian* model of number representation [1]. For the non-symbolic estimation task, we calculated the coefficients of the regression between the correct outcomes and the mean *cv* per child in each presented numerosity. In both the TA and MD groups, the β coefficients were small (TA: mean = 0.145, SD = 0.513; MD: mean = 0.072, SD = 0.449), which indicates that children's responses have scalar variability. Nevertheless, the mean slope was significantly different from 0 in TA, but not in the MD group (TA: $t(161) = 4.105$, $p<0.001$; MD: $t(39) = 0.883$, $p = 0.383$). Similar β coefficients were obtained in the non-symbolic addition task, (TA: mean = 0.094, SD = 0.531; MD: mean = 0.056, SD = 0.581). In this case, the mean slope was not significantly different from zero in both groups (TA: $t(161) = 0.122$, $p = 0.223$; MD: $t(39) = 1.119$, $p = 0.270$). Taken together, these results demonstrate that the performance in all of the tasks that measure the ANS acuity from both the TA and MD groups can be well described by Weber's law. Group differences are presented in the next section.

Differences between the TA and MD groups in ANS acuity

Although all of the children with MD had normal intelligence (IQ>85) and normal spelling achievement (above the 25th percentile), they scored significantly lower in these measures when compared to their TA peers (Table 1; see Table S1 in the Supporting Information for the descriptive data separated by grade). For this reason, intelligence and spelling were included as covariates for group comparisons in exact calculation and ANS acuity (see Table 2 for statistics). As expected, the TA group showed better performance in exact calculation when compared to the children with MD. No group difference was found in the simple reaction time task. More importantly, the TA group presented higher ANS acuity, with significant lower w the non-symbolic comparison task. Moreover, TA children had lower *cv* in the non-symbolic estimation task; however, this difference was only marginally significant. Finally, a significantly lower *cv* was found in the non-symbolic addition task for the TA compared to the MD group.

Cross-sectional trajectories of ANS acuity

Cross-sectional trajectories of the different measures of ANS acuity were investigated separately for the two groups (see Figure 2). The w was found to decrease monotonically with age in the TA children ($\beta = -0.184$, $p = 0.015$), but it remained stable in the children with MD ($\beta = -0.016$, $p = 0.897$). This result suggests that difference between the MD and TA groups in ANS acuity measured by the w increases during development. Second, the *cv* from non-symbolic estimation was found to monotonically decrease with age more or less to the same extent in both the TA and MD groups (TA: $\beta = -0.338$; $p<0.001$; MD: $\beta = -0.425$; $p = 0.006$). Finally, the *cv* from non-symbolic addition was also found to decrease with age by the same extent in both groups, but was only marginally significant in the TA group and was non-significant for the MD group (TA: $\beta = -0.154$, $p<0.050$; MD: $\beta = -0.149$, $p = 0.375$).

To confirm the results of the cross-sectional trajectory of the number comparison task, given the possible lack of statistical power in the MD group to detect significant coefficients, we ran a bootstrap analysis with the regression coefficients of the three measures of the ANS and age. First, we generated 10,000 samples with N = 40 (N of the MD group), allowing repetitive cases for each group separately. Next, we calculated for each sample one β coefficient for each of the regressions: age and non-symbolic comparison, age and non-symbolic estimation, and age and non-symbolic addition. Afterward, we calculated the percentage of positive coefficients in each group, which we use as a likelihood index for the true direction of the association in the population. In the non-symbolic comparison task, the coefficients of the TA group were found to be negative in 91.27% of the generated samples. This finding was not the case in the MD group, in which only 53.67% of the samples showed negative coefficients. For the other measures of ANS acuity, TA and MD showed similar patterns (non-symbolic estimation: TA = 97.30%, MD = 99.30%;

Table 2. ANCOVAs comparing the TA and MD groups in mathematics and in the measures of ANS acuity controlling for intelligence and spelling.

	TA (n = 162)		MD (n = 40)					
Measures	**Mean**	**SD**	**Mean**	**SD**	***F***	***df***	***p-value***	η^2
Exact Calculation (acc)	68.624	19.342	42.048	21.380	33.894	1;198	<0.001	0.146
Simple RT (ms)	416.876	84.201	444.911	102.362	2.363	1;198	0.126	0.012
Nsymb Comparison (w)	0.245	0.085	0.298	0.079	7.560	1;198	<0.01	0.037
Nsymb Estimation (cv)	0.278	0.086	0.312	0.111	3.273	1;198	0.072	0.016
Nsymb Addition (cv)	0.297	0.076	0.351	0.096	7.855	1;198	<0.01	0.038

TA = typically achieving; MD = mathematical difficulties; w: internal Weber fraction; cv: coefficient of variation; acc = accuracy in % correct.

doi:10.1371/journal.pone.0111155.t002

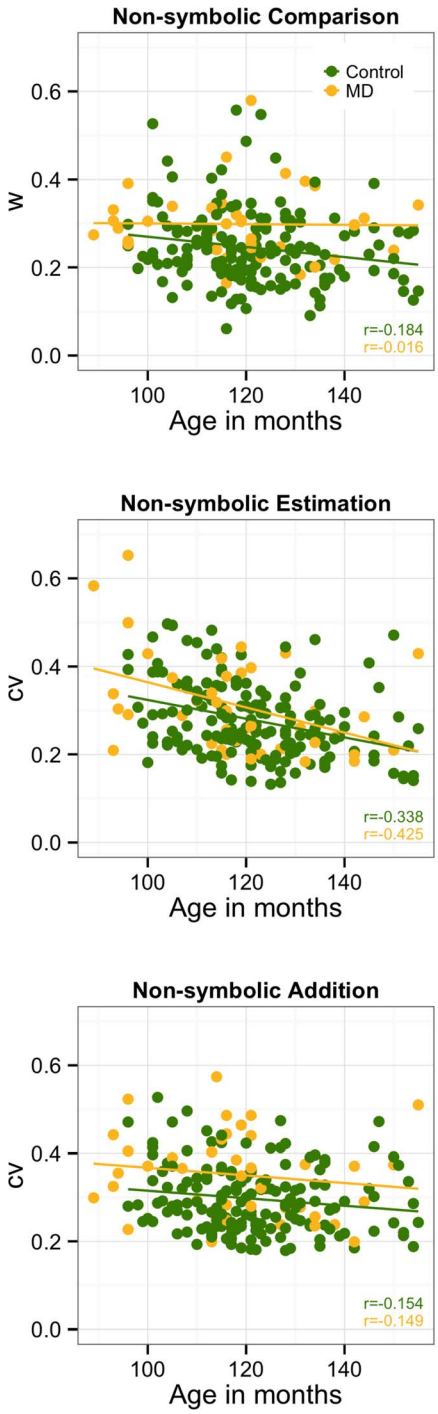

Figure 2. Cross-sectional trajectories of the measures of ANS acuity for the TA and MD groups.
doi:10.1371/journal.pone.0111155.g002

non-symbolic addition: TA: 80.98%, MD: 82:61%). Considering a confidence interval of 90%, both groups showed developmental changes in non-symbolic estimation, but the results were less robust in non-symbolic addition. Crucially, TA showed significant improvement in non-symbolic comparison in contrast to their MD peers, who definitively did not show any sign of improvement in this task during development.

Relationship between the measures of ANS acuity and exact calculation

As expected, all of the three measures of ANS acuity showed significant positive correlations among themselves, even after controlling for the effects of age, spelling and intelligence, which indicates that they share a common construct. Importantly, all three measures of ANS acuity also correlated with exact calculation (Table 3).

Following the suggestion of one the reviews based on the inhibitory control account of the relationship between the non-symbolic comparison task and exact calculation [52,53], we ran separate correlations between the *w* calculated from two sets of stimuli (*wSize*: size control; *wArea*: area control; see Methods) used in the non-symbolic comparison task and exact calculation. Partial correlations controlling for the effects of age, intelligence and spelling revealed that both *wSize* and *wArea* significantly correlated with exact calculation ($r = -0.13$, $p = 0.033$ and $r = -0.175$, $p = 0.007$, respectively). Importantly, Fisher's r-to-z transformation revealed that there was no significant difference between the two correlation coefficients ($z = -0.459$, $p = 0.359$). Therefore, we used the *w* calculated from all trials in further analyses.

Next, the specific contributions of the different instantiations of the ANS in exact calculation were determined by calculating three multiple regression models. In all three models, two blocks of variables were defined. In the first block, the intervening variables age, schooling, general intelligence and spelling abilities were added using the method "*enter*". In the second block, the three measures of ANS acuity were included as single predictors in three separate models. The first block of variables explained 57.8% of the variance in exact calculation (see the coefficients in Table 4). Importantly, all three measures of ANS acuity remained significant predictors of exact calculation after removing the effects of the intervening variables (non-symbolic comparison: $\beta = -0.135$, $p = 0.005$; non-symbolic estimation: $\beta = -0.162$, $p < 0.001$; non-symbolic addition: $\beta = -0.167$; $p < 0.001$). This finding indicates that all three instantiations of the ANS contribute to explaining exact calculation independently of age, schooling, general intelligence and spelling abilities.

Because all measures of ANS acuity are intercorrelated, the extent to which different instantiations of the ANS present unique contributions to exact calculation was determined. A multiple regression model was calculated with the same structure as before. The first block of variables included the same variables as before, and the second block of variables considered the three measures of ANS acuity simultaneously by using the "*stepwise*" method. The regression model kept non-symbolic estimation and non-symbolic addition but excluded non-symbolic comparison. Non-symbolic addition raised the variance explained to 60.1% and non-symbolic estimation to 61.4% (Table 4). Therefore, the multiple regression analysis showed that both non-symbolic estimation and non-symbolic addition have unique contributions to exact calculation. Moreover, the contribution of non-symbolic comparison to explain exact calculation is shared with other instantiations of the ANS.

Together, these results reveal that all of the instantiations of the ANS contribute to explaining exact calculation, but their contributions are not always unique. More specifically, the effects of non-symbolic comparison on exact calculation appear to be fully shared by non-symbolic estimation and non-symbolic addition. In contrast, a portion of the impact of these two variables on exact calculation appears to be unique. That is, the effect of non-symbolic comparison on exact calculation is indirect, because it is common to non-symbolic estimation and non-symbolic addition. Are these results due to mediation processes that act inside the ANS? To specifically test this hypothesis, one must test whether the effect of one instantiation of the ANS (X) on exact calculation (Y) is significantly absorbed by another instantiation of the ANS (M) [54,55]. Moreover, to increase the confidence in the direction of the mediation effect, it is necessary to determine whether the effect of M in exact calculation is also reduced to the same extent by the inclusion of X as a mediator variable.

To investigate these possible mediation effects, we conducted Causal Mediation Analysis [54,55], as implemented in the R package *mediation* (version 4.2.2)" [56]. Six models that analyzed all of the possible combinations of different measures of ANS acuity as both predictors (X) and mediators (M) and exact calculation as the outcome (Y) were calculated. In each model, the total effect of each instantiation of the ANS on exact calculation was decomposed into a mediation and a direct effect. The regression coefficients as well as their confidence intervals and statistical significance are depicted in Table 5. To determine the statistical significance of the coefficient estimates, a nonparametric bootstrap method was employed. To obtain reliable estimates, a total of 10,000 samples for bootstrapping were drawn.

As can be seen in Table 5, only Models 1.1 (X = non-symbolic comparison, M = non-symbolic estimation) and 2.1 (X = non-symbolic comparison; M = non-symbolic addition) presented directional mediation effects ($p = 0.038$ and $p = 0.026$, respectively). These results revealed that both non-symbolic estimation and non-symbolic addition mediate the total effect of non-symbolic comparison on exact calculation. While non-symbolic estimation has a partial mediation effect, because the direct effect between non-symbolic comparison to exact calculation remained significant ($p = 0.012$), non-symbolic addition has a complete mediation effect, because the direct effect from non-symbolic comparison to exact calculation failed to reach statistical significance ($p = 0.064$).

Table 3. Partial correlations between measures of ANS acuity and exact calculation, controlling for age, intelligence and spelling.

	TA and MD group (n = 202)		
Measures	***Exact Calculation***	***Nsymb Estimation (cv)***	***Nsymb Addition (cv)***
Nsymb Comparison (w)	−0.212*	0.253*	0.346**
Nsymb Estimation (cv)	−0.237**	-	0.249**
Nsymb Addition (cv)	−0.233**	-	-

TA = typically achieving; MD = mathematical difficulties; w: internal Weber fraction; cv: coefficient of variation. *Correlation is significant at the 0.05 level; **Correlation is significant at the 0.01 level.

doi:10.1371/journal.pone.0111155.t003

Table 4. Stepwise regression with exact calculation as the dependent variable and non-symbolic comparison, non-symbolic estimation and non-symbolic addition as predictors, regressing out the effects of age, schooling, general intelligence and spelling abilities.

Model	Predictors	*B*	*SE*	*Beta*	*t*	*p-value*
Block 1 $R^2 = 0.578$	Age	0.349	0.156	0.236	2.231	0.027
	Grade	8.736	1.862	0.485	4.692	<0.001
	Raven	8.882	1.530	0.302	5.805	<0.001
	TDE Spelling	8.330	1.719	0.237	4.845	<0.001
Block 2 $R^2 = 0.037$	Nsymb Addition (cv)	−34.688	12.689	−0.136	−2.734	0.007
	Nsymb Estimation (cv)	−29.214	11.568	−0.128	−2.525	0.012

TA = typically achieving; MD = mathematical difficulties; w: internal Weber fraction; cv: coefficient of variation.
doi:10.1371/journal.pone.0111155.t004

The direction of the possible causal direction between the measures of ANS acuity and exact calculation was further corroborated by the fact that the alternative models 1.2 and 2.2 with non-symbolic comparison as the mediator variable did not show any significant mediation effect ($p = 0.069$ and $p = 0.094$, respectively). Finally, the results revealed no mediation directional effects between non-symbolic estimation and non-symbolic addition to exact calculation ($p = 0.055$ and $p = 0.058$, respectively).

Discussion

The present study investigated the relationship between three measures of ANS acuity (non-symbolic comparison, estimation and addition), their cross-sectional trajectories in children with typical and atypical arithmetic abilities, and their specific contributions to exact calculation. The children with MD were found to have impairments in multiple instantiations of the ANS, more specifically in non-symbolic comparison and non-symbolic addition. Moreover, the TA children were more accurate in mapping between non-symbolic magnitudes and number words compared to the children with MD, although this difference was only marginally significant.

Interestingly, the acuity of the non-symbolic comparison was found to develop normally in TA, but not in the MD group. The children with MD did not show any improvement with age in this task. A bootstrapping analysis confirmed that this difference was not due to a lack of statistical power given the smaller sample size of the MD group. Regarding the acuity of non-symbolic addition, both groups improved with age, but the children with MD were less accurate compared to their TA peers. Finally, both groups also improved to the same extent in the non-symbolic estimation task.

The three measures of ANS acuity significantly correlated with each other, which possibly reflects at least in part a common numerosity code, as proposed by Dehaene [34]. Importantly, all three measures of ANS acuity significantly correlated with exact calculation. However, a multiple regression analysis revealed that only non-symbolic estimation and addition contributed with unique proportions of variance in explaining exact calculation. Mediation analysis showed that the effect of non-symbolic comparison on exact calculation was mediated to different degrees by non-symbolic estimation and non-symbolic addition.

Differences between the TA and MD groups in ANS acuity

In line with previous studies that have investigated the cognitive mechanisms that underlie MLD, the ANS acuity as measured by non-symbolic comparison was found to be impaired in children

Table 5. Mediation models with measures of ANS acuity as either predictors (X) or mediators (M) and exact calculation as the outcome (Y).

Models	Variables	Effects	Estimate 95%	CI Lower 95%	CI Upper	*p-value*
1.1	X = Nsymb Comparison	Direct	−25.113	−43.782	−6.943	0.012
	M = Nsymb Estimation	Mediation	−7.665	−14.791	−1.864	0.038
1.2	X = Nsymb Estimation	Direct	−30.810	−50.139	−9.162	0.002
	M = Nsymb Comparison	Mediation	−6.239	−13.116	−1.337	0.069
2.1	X = Nsymb Comparison	Direct	−21.982	−45.660	−1.744	0.064
	M = Nsymb Addition	Mediation	−10.795	−19.664	−2.047	0.026
2.2	X = Nsymb Addition	Direct	−34.319	−60.102	−6.306	0.008
	M = Nsymb Comparison	Mediation	−8.308	−18.448	−0.791	0.094
3.1	X = Nsymb Estimation	Direct	−29.214	−48.779	−8.317	0.003
	M = Nsymb Addition	Mediation	−7.835	−15.859	−1.565	0.055
3.2	X = Nsymb Addition	Direct	−34.688	−59.325	−9.521	0.006
	M = Nsymb Estimation	Mediation	−7.939	−16.354	−1.630	0.058

X = predictor variable, M = mediator variable.
doi:10.1371/journal.pone.0111155.t005

with MD compared to their TA peers [12,13,57]. Importantly, we were able to detect deficits in the ANS even in a group of MD children selected with a more liberal criterion, which probably includes children with high cognitive heterogeneity [43]. This finding lends support to the view that the different forms of MD are better described as a continuous spectrum rather than qualitatively different categories.

The children with MD were also found to be impaired in the acuity of non-symbolic addition. To our knowledge, this study was the first that demonstrated that non-symbolic addition is impaired in children with low achievement in math. De Smedt and Gilmore [38] found no impairment in a non-symbolic addition task in children with DD during the first year of formal schooling. However, the authors used a two-alternative forced choice task and analyzed only the mean accuracy of the responses. In the present study, the task used allowed children to compare their internally generated sum with five different options that were presented. Accordingly, the acuity of the internal representation of numbers could be determined by calculating the *cv*. Therefore, our measure is more sensitive to capturing differences.

Finally, in line with previous studies [12,27], TA children were more accurate in mapping between non-symbolic stimuli and number words compared to children with MD, although in our study this difference was only marginally significant.

Both the MD and TA groups had normal intelligence (IQ>85) and normal spelling achievement (above the 25th percentile), but group differences in those domains were still observed. For this reason, intelligence and spelling abilities were included as covariates in the group comparison analyses. These results are similar to other studies that also found medium to high effect sizes when comparing language-related abilities [13] and intelligence [58] between TA and DD groups. Lower language-related and intelligence performance could reflect the more widespread impairment that is frequently observed in children who have specific developmental disorders [59].

Cross-sectional trajectories of ANS acuity

More detailed analyses of the cross-sectional trajectories of ANS acuity in typically and atypically developing children revealed several findings that merit discussion. First, while the *w* for the TA children decreased with age, this relationship was not the case for the children with MD. While longitudinal studies [16,18] have found that ANS acuity measured by the non-symbolic comparison task prior to formal schooling is a specific predictor of later mathematics achievement in TA children, group studies with younger children (6 to 9 years old) failed to find differences in this task between TA and children with DD (for a review, see Noël & Rousselle [32]). However, studies with younger children did not use the *w* as an index of ANS acuity, which might be a more sensitive parameter to capture group differences compared to the commonly used distance effect (e.g. Oliveira-Ferreira et al. [45]). Furthermore, those studies tended to use a more liberal criterion to classify the children with DD, for example, below the 15th percentile [35,38]. In contrast, studies with older children that used a more stringent criterion to characterize the DD group (below the 10th/5th percentile) reported an elevated *w* for these children compared to their TA peers [12,13]. Because no correlation was found between non-symbolic comparison and age in children with MD as opposed to their TA peers, our results suggest that it might be easier to detect group differences in this task in older children, because differences seem to increase over the course of development.

Importantly, we were able to detect group differences in the non-symbolic comparison task in a group of 10-year-olds even when using a very liberal criterion to select children with math difficulty. The acuity of the non-symbolic addition was found to increase more or less to the same degree in both the TA and MD groups; however, TA children were systematically more accurate than the children with MD. This finding suggests that even from the initial years of formal schooling, MD children might already present a detectable impairment in more complex manipulations of numeric information. In line with the present results, non-symbolic addition measured before formal math instruction was found to be a specific predictor of later math achievement [15]. A similar pattern compared to non-symbolic addition was observed in the non-symbolic estimation task; however, a group comparison revealed that there was only a marginally significant result.

Taken together, the results suggest that the hypothesis put forward by Noël and Rousselle [32], which states that the ANS is not impaired in children with DD, based solely on the results of a single measure of ANS acuity (non-symbolic comparison) might be too simplistic. The present results indicate that compared to TA children, younger children with low achievement in math selected even with a liberal criterion already present a lower acuity in non-symbolic addition, which is a task that probably calls for more manipulations within the ANS than non-symbolic comparison. The characterization of cross-sectional trajectories can be considered as an important step towards the understanding of the evolution of developmental disorders, but should be confirmed in future longitudinal studies [60].

Relationship between the measures of ANS acuity and exact calculation

As expected, significantly positive correlations were found between the three measures of ANS acuity. This finding is consistent with the data from Mazzocco et al. [12], who reported an association between non-symbolic comparison (*w*) and non-symbolic estimation (*cv*) in 14-year-old adolescents. However, the results are inconsistent with the only study that directly investigated the association between more than one measure of ANS acuity [41]. These authors found null correlations between the *w* calculated from non-symbolic versions of the number comparison and approximate addition tasks, suggesting that these tasks are measuring completely different constructs. However, this study has an important limitation, which is that for the non-symbolic comparison task, only three ratios were used to fit the psychometric function. As is known from the psychophysical literature, it is very difficult to have good fits from using only three points in the psychometric curve [61]. Thus, the lack of correlation reported by the authors could simply reflect a poor estimation of the coefficients. In contrast, in the present study, a much larger range of data points (eight) was used to calculate the *w* in the non-symbolic comparison task. Therefore, the moderate but significant correlations between non-symbolic comparison, non-symbolic estimation and non-symbolic addition found in the present study are compatible with the existence of different instantiations of the ANS which at least partially activate a common underlying numerosity code. Nevertheless we agree with both Gilmore et al. [41] and Park and Brannon [62] that it is very misleading to select only one task involving non-symbolic numerical stimuli and present it as a valid index of a supposedly unique ANS. This is a very frequent practice in the numerical cognition literature that should be avoided in further studies.

The indices *w* and *cv* correspond to the degree of noise in the internal representation of numerosity and are mathematically equivalent, in the sense that they are on the same scale (for a comprehensive review on the mathematical basis of the internal Weber fraction, see Dehaene [1]; for an intuitive explanation

about the relationship between the Weber fraction and the coefficient of variation, see Halberda [63]). Indeed, the values of the parameters in the three tasks have been revealed to be similar (see Table 2). However, one should not expect them to be equal, because the tasks that were used to extract them involve very different cognitive processes: simple discrimination, mapping from non-symbolic magnitudes to number words and more complex manipulations of magnitudes in the context of an arithmetic operation, for non-symbolic comparison, estimation and addition. Interestingly, the coefficient values increased from non-symbolic comparison to non-symbolic addition; thus, it is tempting to speculate that this result possibly reflects the summation of noise during the process of accumulation of evidence in more complex forms of magnitude manipulation. For example, the internally generated sum of two given numerosities is possibly noisier than the representation of the numerosities themselves, because during addition there is another source of noise, which arises from the operation itself. In fact, the two existent mathematical models of approximate calculation [37,64] account for this additional source of variation, by including a scaling factor that corresponds to the amount of noise due to the calculation. The same logic could be applied to non-symbolic estimation. In this task, one needs to not only discriminate the magnitudes but also transform the represented value in a symbolic label, which certainly corresponds to another source of noise [1]. The precise mechanisms that are involved in different instantiations of the ANS and how their interaction occurs is a very exciting topic, but scarcely addressed in the literature. Therefore, future studies should investigate empirically the different sources of noise during magnitude manipulations and their spatial-temporal neural dynamics.

Next, we investigated more deeply the association between measures of ANS acuity and exact calculation. Significant correlations between all three measures of ANS acuity and exact calculation were found. Importantly, all three measures of ANS acuity have specific contributions to explain variance in exact calculation, which remain significant after partialling out the more general effects of age, schooling, general intelligence and spelling abilities. Accordingly, the link between the different instantiations of the ANS and exact calculation does not seem to be generated by general cognitive processes, but rather by magnitude processing abilities underlying the tasks. However, our results are not definitive. It is still possible that other general cognitive abilities that we didn't include in this study, such as executive functions and especially inhibitory control [52,53] can account for this link.

Previous studies also found that non-symbolic estimation [12,27,36] and non-symbolic addition [15] are significantly correlated with exact calculation. Regarding the non-symbolic comparison, the literature is more inconsistent and contains both positive and negative results (see De Smedt et al. [24] for a comprehensive review). As noted by De Smedt et al. [24], this inconsistency can probably be attributed to methodological differences in the non-symbolic comparison tasks, in the index calculated from behavior (e.g., RT, accuracy, distance effect, Weber fraction) and also in the math tests used. This inconsistency has led some authors to radically argue that the ANS does not make any contribution to explaining exact calculation [32]. However, this conclusion may have been premature, since two recent meta-analyses reported an association, although moderate, between non-symbolic comparison and math achievement from childhood to adulthood [25,26]. Moreover, the present study also showed that partial correlations between ANS acuity and exact calculation are in the same order of magnitude as that reported by Chen and Li [26], and Fazio et al. [25], which cannot be reduced to more general cognitive processes. After inspecting all these results, one can be confident to assume the existence of a specific link between very basic ANS related processes and exact calculation.

An alternative hypothesis is that the non-symbolic comparison task is not a measure of the precision of numerical representations, but rather a measure of inhibitory control, since it is necessary to inhibit the processing of continuous visual parameters to be able to accurately discriminate between two numerosities [52,53]. To ensure that participants are not using non-numerical variables to judge which collection of dots is the larger, researchers normally use different sets of stimuli varying continuous visual properties, such as dot size and dot total area. For example, Fuchs and McNeil [52] used three different sets of stimuli: dot total area was constant and dot size positively covaried with numerosity; mean dot size was constant and dot total area positively covaried with numerosity; dot total area and mean dot size were both inversely covaried with numerosity ('inverse ANS acuity' trials). Interestingly, results demonstrated that only the accuracy in the inverse ANS acuity set was significantly correlated with math achievement. Furthermore, accuracy in this set showed the highest correlation coefficient with a measure of inhibitory control. Similarly, Gilmore et al. [53] used two sets of stimuli: dot total area and dot size positively covaried with numerosity ('congruent' trials); dot total area and dot size negativelly covaried with numerosity ('incongruent' trials). Consistently with Fuchs and McNeil [52], results showed that incongruent trials were significantly correlated with math achievement but congruent trials were not.

As described in the Methods, in the present study we used two different sets of stimuli: dot size was constant and dot total area positively covaried with numerosity (size control); dot total area was constant and dot size negatively covaried with numerosity (area control). Therefore, we didn't have the 'inverse ANS acuity' [52] or the 'incongruent' [53] sets of trials. Nevertheless, we calculated the w separately for the size control and area control items. Partial correlations controlling for the effects of age, intelligence and spelling revealed that both *wSize* and *wArea* significantly correlated with exact calculation and no significant difference were found between the two coefficients. However, we cannot rule out the possibility that inhibitory skills could account at least in part for the relationship between the non-symbolic comparison task and exact calculation in the present study. It is important to note that a serious limitation of separating the items in two different categories, which the above-mentioned studies didn't take into account, is that the number of observations in each category of stimuli dramatically decreases and consequently compromises the stability of the measure. Therefore, we decided to use the w calculated from all items in all the analyses.

Importantly for our results, higher correlations between non-symbolic comparison and math abilities were reported in studies that measured math ability with standardized achievement batteries, which normally include items closely associated with the representation and manipulation of numerical quantity without invoking knowledge of arithmetic (e.g. TEMA-3 [11,18]). In fact, more recently, Libertus, Feigenson and Halberta [65] analyzed the association between the non-symbolic comparison task and the items present in the widely used TEMA-3 separated in two distinct categories: items associated with informal (e.g. enumeration and number comparison) and formal (e.g. transcoding and exact calculation) mathematical abilities. Results demonstrated that the performance in the non-symbolic comparison task was only significantly correlated with informal mathematical abilities. Similarly, Piazza et al. [13] found the performance in the non-symbolic comparison task significantly

correlated only with a subtest of a standardized math achievement battery that required children to compute the proximity relations between different numbers, but not with the other subtests, which measured transcoding and exact calculation abilities. Therefore, it is most likely that non-symbolic comparison, a very basic form of number manipulation within the ANS, has an indirect and consequently moderate effect in exact calculations.

Accordingly, Lyons and Beilock [39] found that the ability to order a series of digits fully mediated the correlation between non-symbolic comparison and basic symbolic arithmetical operations in adults. In the same line, van Marle et al. [40] found that the association between non-symbolic comparison and math achievement in children was also fully mediated by a series of symbolic numerical tasks, mainly the knowledge of cardinal value. Following the same logic, it is also possible that other instantiations of the ANS that require different forms of numerical manipulation also account for the effect of non-symbolic comparison in exact calculation. Indeed, results of our multiple regression model revealed that non-symbolic comparison did not uniquely contribute to explain the variance of exact calculation, when all three measures of ANS acuity were considered simultaneously. To our knowledge, no previous study has systematically investigated the effects of non-symbolic comparison, estimation and addition on exact calculation. In order to do that, we calculated six mediation models with all combinations of measures of ANS acuity as either predictors or mediators and exact calculation as the outcome. The results revealed first that non-symbolic estimation partially mediates the relation between non-symbolic comparison and exact calculation. This finding is in line with Mazzocco et al. [12], who demonstrated that both non-symbolic comparison and non-symbolic estimation accounted for unique proportions of variance in a math achievement task. Second, a full mediation effect of non-symbolic addition was found to be present in the relation between non-symbolic comparison and exact calculation. These results are fully compatible with the ones recently reported by in Park and Brannon [61], who demonstrated that the training on non-symbolic addition but not in non-symbolic comparison has a significant transfer effect to exact calculation. Therefore, the authors suggested that the active process of manipulating numerical information is the critical mechanism underlying the association between the basic number processing and exact calculation.

Although significantly correlated, the three ANS related tasks investigated in the present study involve different cognitive processes. The non-symbolic comparison task involves a very basic operation of magnitude discrimination, which is found to be already present in infancy. Differently, the non-symbolic estimation task involves a transcoding process from approximate to exact symbolic representations of numbers. Finally, the non-symbolic addition involves a more complex process of arithmetical transformations. Results of the mediation analyses show the existence of multiple associations between the different measures of ANS acuity and exact calculation and suggest the existence of a hierarchy of complexity between different instantiations of the ANS. These different instantiations seems to be organized from the more basic and less cognitively demanding forms of number processing to more elaborate operations that involve more active manipulation of magnitudes. The crucial evidence supporting this hypothesis is that alternative models with non-symbolic comparison as the mediator variable for the association of non-symbolic estimation and non-symbolic addition with exact calculation showed no significant mediation effects. At the neural level, this hierarchical organization of different processes underlying number representation and manipulation might reflect the increasing functional connectivity between and within the left and right parietal cortices, as observed during the performance of number-related tasks with increasing demands on the processing of numerical information [42]. Finally, the hierarchical structure of the different instantiations of the ANS can account for the finding that exact calculation is more strongly associated with non-symbolic estimation and non-symbolic addition, compared to non-symbolic comparison.

Conclusions

Benefiting from high statistical power, we showed that children with MD, even when selected with a more liberal criterion, present lower acuity in multiple instantiations of the ANS (non-symbolic comparison and addition), even after controlling for the effects of intelligence and spelling abilities. This finding lends support to the view that the different forms of MD are better described as a continuous spectrum rather than qualitatively different categories. Second, the analyses of the cross-sectional trajectories showed that the ANS acuity measured by all three tasks positively correlated with age in TA children, while no correlation was found between non-symbolic comparison and age in the MD group. A plausible explanation for this result is that number discrimination, as the most basic form of numerical manipulation, is less prone to compensatory strategies that MD children could have developed to solve the other number-related tasks. Third, for the first time, we demonstrated that the three instantiations of the ANS investigated were significantly correlated among each other, reflecting at least in part a common numerosity code. Finally, mediation models revealed that non-symbolic estimation partially and non-symbolic addition fully mediated the effects of non-symbolic comparison in exact calculation. Therefore, the present study represents an important step towards a deeper understanding of the cognitive mechanisms underlying the relationship between basic number processing and mathematics. Given the highly hierarchical nature of mathematics, further studies should focus on precisely investigating the association between each instantiation of the ANS and different forms of mathematical reasoning. This will certainly help to a better understanding of the typical normal development of mathematical abilities as well as the nature of developmental dyscalculia.

Supporting Information

Table S1 Descriptive data of the individual assessment sample by grade.
(DOCX)

Data S1 Raw data. Data_S1.zip. (sheet 1: data; sheet 2: variables specifications).
(ZIP)

Author Contributions

Conceived and designed the experiments: JL HK AK PPC VGH KW. Performed the experiments: PPC IS. Analyzed the data: PPC GW IS. Contributed reagents/materials/analysis tools: PPC GW. Wrote the paper: PPC GW VGH.

References

1. Dehaene S (2007) Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In: Haggard P, Rossetti Y, Kawato W, editors. Sensorimotor foundations of higher cognition:

Attention and performance. Cambridge (UK): Oxford University Press. pp. 527–574.
2. Xu F, Spelke ES (2000) Large number discrimination in 6-month-old infants. Cognition 74: B1–B11.
3. Nieder A, Miller EK (2003) Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. Neuron 37: 149–157.
4. Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S (2004) Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron 44: 547–555.
5. Dehaene S, Brannon EM (2010) Space, time, and number: a Kantian research program. Trends Cogn Sci 14: 517–519.
6. Izard V, Dehaene S (2008) Calibrating the mental number line. Cognition 106: 1221–1247.
7. Moyer RS, Landauer TK (1967) Time required for judgements of numerical inequality. Nature 215: 1519–1520.
8. Izard V, Sann C, Spelke ES, Streri A (2009) Newborn infants perceive abstract numbers. Proc Natl Acad Sci USA 106: 10382–10385.
9. Dehaene S (1992) Varieties of numerical abilities. Cognition 44: 1–42.
10. Piazza M (2010) Neurocognitive start-up tools for symbolic number representations. Trends Cogn Sci 14: 542–551.
11. Halberda J, Mazzocco MM, Feigenson L (2008) Individual differences in non-verbal number acuity correlate with maths achievement. Nature 455: 665–668.
12. Mazzocco MM, Feigenson L, Halberda J (2011) Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). Child Dev 82: 1224–1237.
13. Piazza M, Facoetti A, Trussardi AN, Berteletti I, Conte S, et al. (2010) Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. Cognition 116: 33–41.
14. Bonny JW, Lourenco SF (2013) The approximate number system and its relation to early math achievement: evidence from the preschool years. J Exp Child Psychol 114: 375–388.
15. Gilmore CK, McCarthy SE, Spelke ES (2010) Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. Cognition 115: 394–406.
16. Libertus ME, Feigenson L, Halberda J (2011) Preschool acuity of the approximate number system correlates with school math ability. Dev Sci 14: 1292–1300.
17. Libertus ME, Feigenson L, Halberda J (2013) Is approximate number precision a stable predictor of math ability? Learn Individ Differ 25: 126–133.
18. Mazzocco MM, Feigenson L, Halberda J (2011b) Preschoolers' precision of the approximate number system predicts later school mathematics performance. PLoS One 6: e23749.
19. Park J, Brannon EM (2013) Training the approximate number system improves math proficiency. Psychol Sci 24(10): 2013–2019.
20. Desoete A, Ceulemans A, De Weerdt F, Pieters S (2012) Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. Br J Educ Psychol 82: 64–81.
21. Inglis M, Attridge N, Batchelor S, Gilmore C (2011) Non-verbal number acuity correlates with symbolic mathematics achievement: but only in children. Psychon Bull Rev 18: 1222–1229.
22. Landerl K, Kölle C (2009) Typical and atypical development of basic numerical skills in elementary school. J Exp Child Psychol 103: 546–565.
23. Mundy E, Gilmore CK (2009) Children's mapping between symbolic and nonsymbolic representations of number. J Exp Child Psychol 103: 490–502.
24. De Smedt B, Noël M-P, Gilmore C, Ansari D (2013) How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. Trends Neurosci Edu 2: 48–55.
25. Fazio LK, Bailey DH, Thompson CA, Siegler RS (2014) Relations of different types of numerical magnitude representations to each other and to mathematics achievement. J Exp Child Psychol 123: 53–72.
26. Chen Q, Li J (2014) Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. Acta Psychol 148: 163–172.
27. Mejias S, Mussolin C, Rousselle L, Gregoire J, Noël MP (2011) Numerical and nonnumerical estimation in children with and without mathematical learning disabilities. Child Neuropsychol 18(6): 550–75.
28. Piazza M, Pica P, Izard V, Spelke ES, Dehaene S (2013) Education enhances the acuity of the nonverbal approximate number system. Psychol Sci 24: 1037–1043.
29. Pica P, Lemer C, Izard V, Dehaene S (2004) Exact and approximate arithmetic in an Amazonian indigene group. Science 306: 499–503.
30. Bradley L, Bryant PE (1983) Categorizing sounds and learning to read - a causal connection. Nature 301: 419–421.
31. Castles A, Coltheart M (2004) Is there a causal link from phonological awareness to success in learning to read? Cognition 91: 77–111.
32. Noël MP, Rousselle L (2011) Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. Front Hum Neurosci 5: 165.
33. Chu FW, van Marle K, Geary DC (2013) Quantitative deficits of preschool children at risk for mathematical learning disability. Front Psychol doi: 10.3389/fpsyg.2013.00195
34. Dehaene S (2009) Origins of mathematical intuitions: the case of arithmetic. Ann N Y Acad Sci 1156: 232–259.
35. Rousselle L, Noël MP (2007) Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs. non-symbolic number magnitude processing. Cognition 102: 361–395.
36. Booth JL, Siegler RS (2006) Developmental and individual differences in pure numerical estimation. Dev Psychol 42: 189–201.
37. Barth H, La Mont K, Lipton J, Dehaene S, Kanwisher N, et al. (2006) Non-symbolic arithmetic in adults and young children. Cognition 98: 199–222.
38. De Smedt B, Gilmore CK (2011) Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. J Exp Child Psychol 108: 278–292.
39. Lyons IM, Beilock SL (2011) Numerical ordering ability mediates the relation between number-sense and arithmetic competence. Cognition 121(2): 256–61.
40. van Marle K, Chu FW, Li Y, Geary DC (2014) Acuity of the approximate number system and preschoolers' quantitative development. Dev Sci doi: 10.1111/desc.12143.
41. Gilmore C, Attridge N, Inglis M (2011) Measuring the approximate number system. Q J Exp Psychol (Hove) 64: 2099–2109.
42. Park J, Park DC, Polk TA (2013) Parietal functional connectivity in numerical cognition. Cereb Cortex 23: 2127–2135.
43. Mazzocco MM (2007) Defining and differentiating mathematical learning disabilities and difficulties. In Berch DB, Mazzocco MM, editors. Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities. Baltimore: Brookes. pp. 29–47.
44. Stein LM (1994) TDE – Teste do Desempenho Escolar. Manual para aplicação e interpretação. São Paulo: Casa do Psicólogo. 204 p.
45. Oliveira-Ferreira F, Costa DS, Micheli LR, Pinheiro-Chagas P, Haase VG (2012) School Achievement Test: normative data for a representative sample of elementary school children. Psychology & Neuroscience 5(2): 157–164
46. Angelini AL, Alves ICB, Custódio EM, Duarte WF, Duarte JLM (1999) Matrizes progressivas coloridas de Raven – escala especial. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia 435 p.
47. Costa AJ, Silva JB, Chagas PP, Krinzinger H, Lonneman J, et al. (2011) A hand full of numbers: a role for offloading in arithmetics learning? Front Psychol 2: 368.
48. Júlio-Costa A, Antunes AM, Lopes-Silva JB, Moreira BC, Vianna GS, et al. (2013) Count on dopamine: influences of COMT polymorphisms on numerical cognition. Front Psychol 4: 531.
49. Dehaene S, Izard V, Piazza M (2005) Control Over Non-Numerical Parameters In Numerosity Experiments. Available: www.unicog.org/docs/DocumentationDotsGeneration.doc. Accessed 12 February 2014
50. Knops A, Viarouge A, Dehaene S (2009) Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. Atten Percept Psychophys 71: 803–821.
51. The R Development Core Team (2008) R: a language and environment for statistical computing. Available: http://www.lsw.uni-heidelberg.de/users/christlieb/teaching/UKStaSS10/R-refman.pdf. Accessed 12 February 2014. Accessed 2014 Feb 12.
52. Fuchs MW, McNeil NM (2013) ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. Dev Sci 16: 136–148.
53. Gilmore C, Attridge N, Clayton S, Cragg L, Johnson S, et al. (2013) Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. PLoS One 8(6): e67374.
54. Imai K, Keele L, Tingley D (2010) A general approach to causal mediation analysis. Psychol Methods 15: 309–334.
55. Imai K, Tingley D, Yamamoto T (2013) Experimental designs for identifying causal mechanisms. J Royal Stat Soci 176(1): 5–51.
56. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K (2014) mediation: R package for causal mediation analysis. J Stat Softw 59(5); 1–38.
57. Mussolin C, Mejias S, Noël MP (2010) Symbolic and nonsymbolic number comparison in children with and without dyscalculia. Cognition 115: 10–25.
58. Ashkenazi S, Rosenberg-Lee M, Tenison C, Menon V (2012) Weak task-related modulation and stimulus representations during arithmetic problem solving in children with developmental dyscalculia. Dev Cogn Neurosci 2 Suppl 1: S152–166.
59. Johnson MH (2012) Executive function and developmental disorders: the flip side of the coin. Trends Cogn Sci 16: 454–457.
60. Thomas MS, Anaaz D, Ansari D, Scerif G, Jarrold C, et al. (2009) Using developmental trajectories to understand developmental disorders. J Speech Lang Hear Res 52(2):336–358.
61. Kingdom FAA, Prins N. (2010) Psychophysics: a practical introduction. London: Academic Press. 304 p.
62. Park J, Brannon EM (2014) Improving arithmetic performance with number sense training: An investigation of underlying mechanism. Cognition 133: 188–200.
63. Halberda J (2011) What is a Weber fraction? Available: http://pbs.jhu.edu/research/halberda/publications/pdf/HalberdaWeberChp110124.pdf. Accessed 2014 Feb 12.
64. Cordes S, Gallistel CR, Gelman R, Latham P (2007) Nonverbal arithmetic in humans: light from noise. Percept Psychophys 69: 1185–1203.
65. Libertus ME, Feigenson L, Halberda J (2013) Numerical approximation abilities correlate with and predict informal but not formal mathematics abilities. J Exp Child Psychol 116: 829–838.

Commentary

Psychological Science
2014, Vol. 25(11) 2116–2119

DOI: 10.1177/0956797614548876
pss.sagepub.com

No Conclusive Evidence for Numerical Priming Under Interocular Suppression

Guido Hesselmann[1] and André Knops[2]
[1]Visual Perception Laboratory, Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany, and [2]Department of Psychology, Humboldt-Universität zu Berlin

Received 4/1/14; Revision accepted 8/1/14

The scope and limits of unconscious priming effects are a central topic of research in cognitive science and neuroscience (Dehaene & Changeux, 2011). The findings of one recent study published in *Psychological Science* (Bahrami et al., 2010) suggest that numerical processing of small quantities (1–3) can escape interocular suppression and lead to robust unconscious priming in an enumeration task, contradicting earlier work that showed a lack of semantic priming effects with words and a lack of repetition priming effects with pictorial images when prime stimuli were suppressed under binocular rivalry (Cave, Blake, & McNamara, 1998; Zimba & Blake, 1983). Bahrami et al. used a variant of interocular suppression called continuous flash suppression (CFS), which disrupts conscious perception of visual stimuli for up to seconds or even minutes (Tsuchiya & Koch, 2005). Thus far, the evidence regarding the degree of unconscious processing that occurs during CFS has been mixed (Sterzer, Stein, Ludwig, Rothkirch, & Hesselmann, 2014).

In the experiments of Bahrami et al., participants had to enumerate a target set on each trial, and the reported response time (RT) data seemed to show that unconsciously perceived nonsymbolic and symbolic primes (Gabor patches and Arabic digits, respectively) induced a distance-dependent priming effect for nonsymbolic targets. Here, we argue that it is premature to conclude that the experiments demonstrated unconscious high-level priming specific to the quantity relationship between target and prime. Our argument rests on the design specifics of the paradigm used. Both targets and primes could be 1, 2, or 3, so the possible target-prime distances were −2, −1, 0, +1, and +2. Bahrami et al. collapsed across different target-prime combinations to test the effect of target-prime distance. For example, target-prime distances of −2 and +2 were instantiated by only one target-prime combination each (target: 1, prime: 3, and target: 3, prime: 1, respectively), whereas the target-prime distance of 0 was instantiated by three target-prime combinations involving all targets (see Fig. 1a).

Irrespective of target numerosity, Bahrami et al. calculated a baseline from RTs in prime-absent trials and subtracted this baseline from RTs in prime-present trials. We conducted a replication experiment[1] and found a large priming effect using this common no-prime baseline, $F(4, 48) = 18.27$, $p < .001$, Greenhouse-Geisser's $\varepsilon = .46$. This effect was linearly modulated by target-prime distance. Thus, we replicated the results of Bahrami et al.: Whereas the effect signaled interference in the case of negative distances (i.e., slower RTs relative to baseline) and facilitation in the case of positive distances, it was virtually absent when the distance was 0 (Fig. 1a).

However, use of the common no-prime baseline rests on the assumption of equal target-specific RTs. If this assumption is violated, the effect of target-prime distance is unequivocally confounded with an effect of target numerosity. In our experiment, participants were approximately 100 ms faster when responding to three Gabor patches than when responding to one[2] (Fig. 1b). This RT difference is reflected in the difference between the priming effects at target-prime distances of −2 and +2 (Fig. 1a).

In their Supplemental Material, Bahrami et al. acknowledged this confound but argued that in their data, target-specific RTs in prime-absent trials did not differ significantly from one another, $F(2, 32) = 1.78$, $p > .1$, ignoring the fact that this result does not imply equality of target-specific RTs in prime-present trials. In an attempt to rule out the possibility that the confound had an effect

Corresponding Author:
Guido Hesselmann, Charité Berlin, Department of Psychiatry and Psychotherapy, Charitéplatz 1, Berlin 10117, Germany
E-mail: guido.hesselmann@charite.de

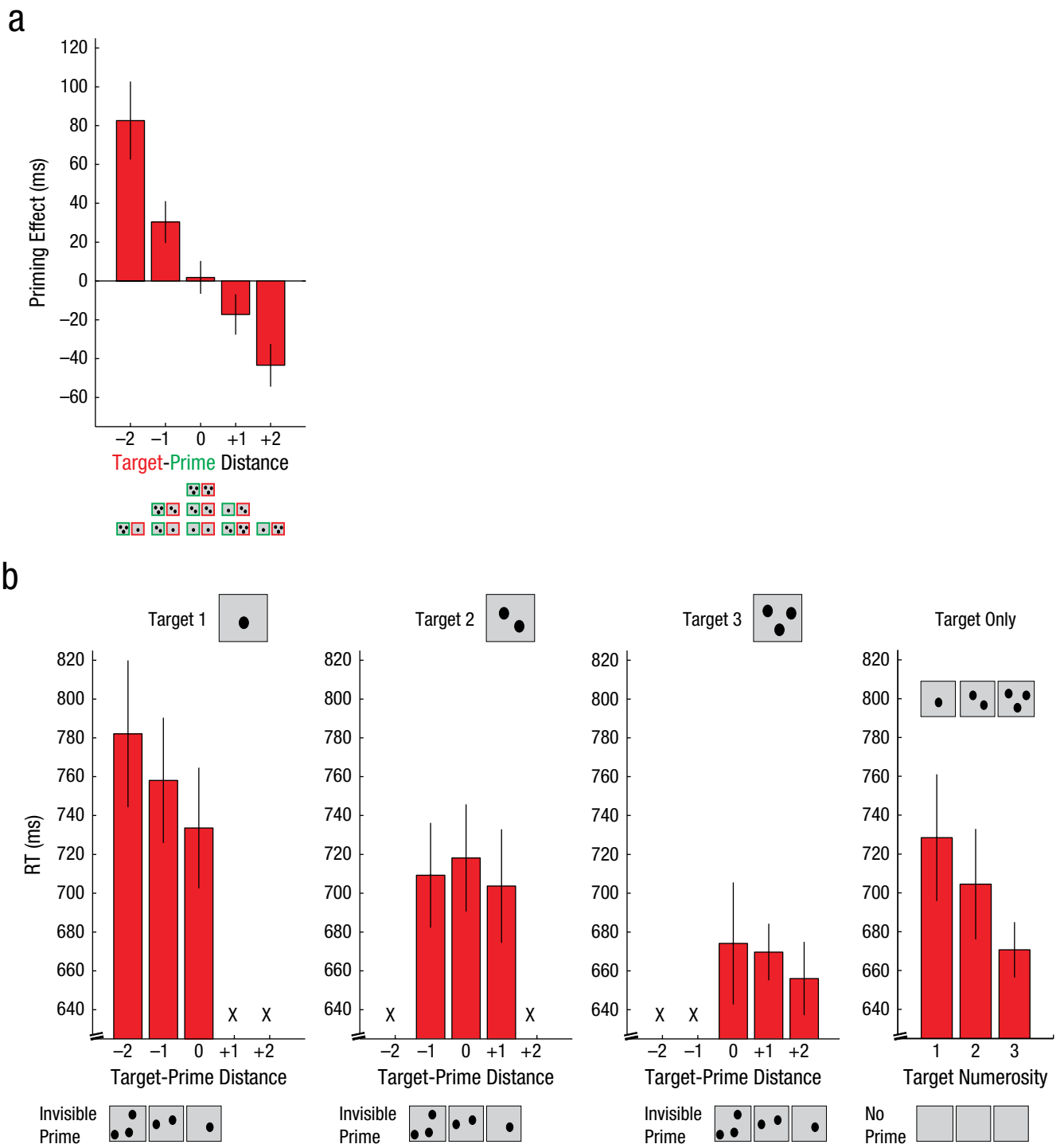

Fig. 1. Experimental design and results of the replication experiment ($N = 13$). The graph in (a) shows the mean priming effect on response times (RTs) as a function of target-prime distance, collapsed across target numerosities (e.g., targets of 1, 2, and 3 were all used across trials with a target-prime distance of 0). The priming effect was calculated by subtracting RT in the common no-prime baseline (i.e., prime-absent trials) from RT in the prime-present trials. Below the graph, the green (primes) and red (targets) "dice" indicate the target-prime combinations for each target-prime distance. The graphs in (b) show raw RTs separately for all target-prime combinations in prime-present trials and for all target numerosities in prime-absent (target-only) trials. For the prime-present trials, each graph shows RT as a function of target-prime distance, and the *X*s denote the target-prime distances that were missing from the experimental design. Error bars represent standard errors of the mean.

on their results, they also calculated priming effects using as their baseline target-specific RTs from prime-absent trials. Using this specific no-prime baseline left the overall pattern of results intact, which we confirmed in our replication experiment, $F(4, 48) = 7.04$, $p = .003$, $\varepsilon = .52$. This approach, however, rests on the assumption that differences across target-specific RTs are the same in prime-absent trials as in prime-present trials (i.e., that there is no Target Numerosity × Prime Presence interaction). If this assumption is violated—which it was in our experiment, which revealed a significant interaction, $F(2, 24) = 3.63$, $p = .049$, $\varepsilon = .88$—subtracting the baseline RTs will distort the shape of the priming function and introduce a spurious effect of target-prime distance.

Alternatively, multiple regression analysis (Lorch & Myers, 1990) of raw prime-present RTs can be used to directly test for different shapes of the priming function (Fig. 1b). V-shaped and steplike priming functions have previously been reported for symbolic and nonsymbolic numerical priming, respectively (Roggeman, Verguts, & Fias, 2007; Van Opstal, Gevers, De Moor, & Verguts, 2008). In our experiment, regression coefficients of the linear-function predictor did not deviate significantly from zero, mean $\beta = 12.32$, 95% confidence interval (CI) = [−4.11, 28.76], nor did we observe V-shaped effects, mean $\beta = 6.91$, 95% CI = [−7.55, 21.36], or steplike effects, mean $\beta = 0.44$, 95% CI = [−16.11, 17.00]; the coefficients of the target-numerosity predictor in this analysis did deviate from zero, mean $\beta = -36.01$, 95% CI = [−58.90, −13.12]. By contrast, regression analyses of the priming functions obtained following subtraction of the common no-prime and specific no-prime baselines, which do not fully account for effects of target numerosity, clearly indicated linearity, mean $\beta = 33.26$, 95% CI = [17.08, 49.44], and mean $\beta = 15.94$, 95% CI = [7.28, 24.60], respectively.

It turns out that the design and analysis used by Bahrami et al. does not allow a conclusive answer to the question of whether there is distance-dependent numerical priming during interocular suppression. One could overcome the pervasive confound of target numerosity, however, by increasing the range of prime numerosities so that all targets appear with the same target-prime distances (e.g., if target 3 appears in combination with primes 1, 2, 3, 4, and 5, target 4 would appear in combination with primes 2, 3, 4, 5, and 6).

To sum up, the study by Bahrami et al. opened up an interesting new avenue of research by investigating the facilitatory and interference effects of unconsciously perceived numerosity primes under CFS. It is premature, however, to conclude that their RT data demonstrate high-level priming, or specifically distance-dependent numerical priming, during interocular suppression.

Author Contributions

G. Hesselmann and A. Knops developed the study concept. G. Hesselmann performed the data analysis. G. Hesselmann and A. Knops wrote the manuscript.

Acknowledgments

We thank Natasha Darcy and Kerstin Wolf for help with data acquisition.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

G. Hesselmann and A. Knops were supported by the German Research Foundation (Grants HE 6244/1-1 and KN 959/2-1).

Notes

1. Nineteen observers participated in this experiment, which closely followed the procedure used by Bahrami et al. and used very similar stimuli. Prime and target stimuli were nonsymbolic (arrays of Gabor patches). Six participants were excluded because their discrimination of invisible primes in a control experiment was above chance ($n = 3$) or because they provided an insufficient number of trials in one of the conditions ($n = 3$).
2. RT profiles for numerosities within subitizing range are usually flat (Piazza, Fumarola, Chinello, & Melcher, 2011), but by superimposing Gabor targets on dynamic CFS masks (to minimize prime afterimages), Bahrami et al. might have added a visual search component to the enumeration task.

References

Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science*, *21*, 224–233.

Cave, C. B., Blake, R., & McNamara, T. P. (1998). Binocular rivalry disrupts visual priming. *Psychological Science*, *9*, 299–302.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*, 200–227.

Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognition research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 149–157.

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*, 147–153.

Roggeman, C., Verguts, T., & Fias, W. (2007). Priming reveals differential coding of symbolic and non-symbolic quantities. *Cognition*, *105*, 380–394.

Sterzer, P., Stein, T., Ludwig, K., Rothkirch, M., & Hesselmann, G. (2014). Neural processing of visual information under interocular suppression: A critical review. *Frontiers in Psychology*, *5*, Article 453. Retrieved from http://journal.frontiersin.org/Journal/10.3389/fpsyg.2014.00453/full

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*, 1096–1101.

Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*, 419–425.

Zimba, L. D., & Blake, R. (1983). Binocular rivalry and semantic processing: Out of sight, out of mind. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 807–815.

OPEN ACCESS Freely available online

PLOS ONE

Operational Momentum in Multiplication and Division?

Curren Katz*, André Knops

Faculty of Life Sciences, Humboldt Universität zu Berlin, Berlin, Germany

Abstract

Biases are commonly seen in numerical cognition. The operational momentum (OM) effect shows that responses to addition and subtraction problems are biased in the whole-number direction of the operation. It is not known if this bias exists for other arithmetic operations. To determine whether OM exists in scalar operations, we measured response bias in adults performing symbolic (Arabic digits) and non-symbolic (dots) multiplication and division problems. After seeing two operands, with either a multiplication (×) or division (÷) sign, participants chose among five response choices. Both non-random performance profiles and the significant contribution of both operands in a multiple regression analysis predicting the chosen values, suggest that adults were able to use numerical information to approximate the outcomes in both notations, though they were more accurate on symbolic problems. Performance on non-symbolic problems was influenced by the size of the correct choice relative to alternatives. Reminiscent of the bias in addition and subtraction, we found a significant response bias for non-symbolic problems. Non-symbolic multiplication problems were overestimated and division problems were underestimated. These results indicate that operational momentum is present in non-symbolic multiplication and division. Given the influence of the size of the correct choice relative to alternatives, an interaction between heuristic bias and approximate calculation is possible.

Citation: Katz C, Knops A (2014) Operational Momentum in Multiplication and Division? PLoS ONE 9(8): e104777. doi:10.1371/journal.pone.0104777

Editor: Michael A. Motes, Center for BrainHealth, University of Texas at Dallas, United States of America

Received September 27, 2013; **Accepted** July 18, 2014; **Published** August 14, 2014

Funding: This work was supported by a grant (KN 959/2-1) from Deutsche Forschungsgemeinschaft/German Research Council (DFG (www.dfg.de) to André Knops in support of Curren Katz. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: curren@post.harvard.edu

Introduction

Humans, as well as other animals, have an innate approximate number sense that they use to interact with the environment [1–3]. Disease, injury [4], and environmental variables (*e.g.* ineffective education) can impact this system at a high cost to individuals and society. Even in healthy populations, lower numeracy predicts poor decision making and susceptibility to bias [5]. Given the importance of numerical abilities, research has focused on understanding the underlying cognitive processes.

According to the *triple-code model*, numbers can be represented in three codes [6]. In the Arabic code, associated with bilateral occipito-temporal regions, numbers are represented as Arabic numerals and can be used to perform symbolic arithmetic. In the verbal code, associated with left perisylvian language areas, numbers are represented as words and memorized arithmetic facts. In the magnitude code, associated with bilateral parietal areas, numbers are represented as abstract magnitudes and perhaps points on a spatially oriented mental number line (MNL). Consistent with the idea of a mental number line, parietal neural populations tuned to small quantities exhibit a topographic organization [7]. This innate approximate number system (ANS) supports quantity knowledge (*e.g.* 3 is smaller than 7), as well as estimation and calculation on non-symbolic quantities [8].

Perhaps due to the spatial features of quantity representation, spatial and directional biases are frequently seen in numerical tasks. The *Spatial-Numerical Association of Response Codes* (SNARC) effect shows that smaller numbers are left-side associated and larger numbers are right-side associated [9,10]. Further evidence for the spatial nature of number representation comes from magnitude-dependent covert shifts of attention during number viewing [11,12]. Directional bias is seen in addition and subtraction, when participants overestimate for addition and underestimate for subtraction [13]. This operational momentum (OM) effect occurs in adults performing non-symbolic and, to a lesser extent, approximate symbolic arithmetic [14]. Infants exhibit OM as well, demonstrated by looking longer at arithmetic animations violating the momentum of the operation [15]. Interestingly, school-age children may overestimate non-symbolic subtraction, although this could be due to individual differences in attention [16]. OM occurs in exact symbolic arithmetic, as long as an approximate response method is used [17,18].

Although OM research has focused on whole numbers, adults and children answering symbolic arithmetic questions also show a tendency to believe addition/multiplication always makes more than the initial quantity and subtraction/division always makes less, even though this is not necessarily true with operations including non-whole rational numbers (*e.g.* $8 \times .5 = 4$) or zero (*e.g.* $8 \times 0 = 0$) [19,20]. The origin of this whole-number bias is still a matter of debate [21]. In all four arithmetic operations, the 'addition/multiplication makes bigger, subtraction/division makes smaller' intuition [22] could lead to the correct choice, over/under-estimation in the direction of the operation (an OM effect), or even over/under-estimation counter to the direction of the operation (a reverse OM effect), as long as the estimation was larger than the initial quantity. OM research demonstrating systematic over and under estimation on approximate symbolic addition and subtraction problems shows that, at least for these operations, whole numbers themselves are subject to directional biases. It is not yet known if whole-number multiplication and division are subject to directional biases.

Different explanations for OM have been proposed based on response bias in addition and subtraction. Addition and subtraction have been described as spatial movements on a mental number line [23], and the OM effect attributed to movements or shifts of attention too far along this line [14,24]. Alternatively, if the mental number line is logarithmically compressed [8,25], OM may result from flawed decompression [13], though not all OM research has supported this [16,17]. A simple rule of accepting more than the original operand for addition and less for subtraction may also explain the observed bias [15,19–21]. Whether these explanations for OM can be reconciled remains unclear. Since the term operational momentum could imply a spatial origin of the observed bias, it is important to separate proposed cognitive underpinnings (*e.g.* spatial, attentional shifts, etc.) from the observed bias. When we use the term OM, we refer to the observed empirical response pattern without any assumptions about the underlying mechanism [26,27].

Although a fair amount of research has focused on OM in addition and subtraction, scalar operations such as multiplication and division have never been tested. In this article, "scalar operations" refer to problems where a quantity element is modified by a scalar element [45]. Studies of other operational biases have only used symbolic formats [19,20]. This may be due to the small number of studies addressing non-symbolic scalar operations, most of which have focused on children prior to instruction [28,29]. Children in kindergarten and 1st grade can double and halve discrete (dot arrays) and continuous (lines) stimuli [28]. Children in this age group are also able to quadruple and even multiply by a fraction (*e.g.* 2.5) [29]. The limited existing research supports non-symbolic multiplication and division ability and therefore the possibility of studying OM in these operations. Demonstration of OM in scalar operations would add to our understanding of OM in particular and numerical decision making in general.

In this context, we designed a study to test whether OM exists in whole-number multiplication and division by presenting symbolic and non-symbolic problems and measuring response bias. Our first goal was to see whether OM exists in whole-number multiplication and division. Finding OM in multiplication and division could suggest that the ANS influences scalar operations. Our second goal was to see whether participants could use the ANS to solve non-symbolic multiplication and division problems using larger quantities (operands and results) than previous studies. We found that participants based their responses on a combination of both operands, implying reliance on numerical information rather than mere guessing or plausibility checks. Most importantly, they demonstrated OM in non-symbolic problems.

Materials and Methods

Ethics Statement

The study was approved by the Humboldt University Department of Psychology Ethics committee (Nr.: 2010–12) on October, 8, 2010. Written informed consent was obtained. Participants were reimbursed 8€/hour for participation in the study.

Participants

Sixteen native German-speaking right-handed participants (12 female; 20–65 years old, mean = 33.88, SD = 13.12) were recruited in Berlin, Germany, using a Humboldt University department database. Participants who reported a history of psychiatric illness were excluded.

Stimuli

Twenty-four multiplication and 24 division problems were created (table 1). To control for correct value size, the same response choices, including dot arrays for non-symbolic problems, were used for multiplication and division. The task design was based on a previously reported adult OM assessment method [14,16]. The correct result (*C*) and 6 incorrect results were created in a geometric series (symbolic: C x $1.5^{i/3}$ & non-symbolic: $C \times 2^{i/3}$; *i* from −3 to 3). Previous research has shown that subjects tend to avoid extreme results in symbolic calculation [14]. To increase the likelihood of finding an OM effect in symbolic problems, 1.5 rather than 2 was used. To control for parity, symbolic response alternatives were rounded to the closest value with the same parity as the correct result. To avoid the strategy of choosing the middle value, only 5 of the 7 possible results were presented. In 50% of trials the low range was presented and the 4th result was correct. In the other 50% of trials the high range was presented and the 2nd result was correct (Fig. 1).

Non-symbolic stimuli were created using MATLAB (The MathWorks, Inc., 2012) and the Psychophysics Toolbox extension [30,31], using the method described by Gebuis and Reynvoet [32]. Previous research has varied intensive (*e.g.* dot size) and extensive (*e.g.* envelope size, area, density) parameters separately. In this case, although participants cannot rely on one feature for all trials they could, for example, use area in half the trials to accurately predict quantity and dot size in the other trials, choosing the best strategy for each trial. We overcame this by first generating 2 dot arrays for each of the five response choice values (comp_dots_-version180112.m, http://titiagebuis.eu). We then selected an optimal combination of 5 dot arrays by testing the correlation of visual parameters and quantity for all possible combinations. We chose combinations with individual correlations less than.4 to create groups of uncorrelated dot arrays. The mean correlations between quantity and extensive and intensive visual parameters were 0.05 and 0.02, respectively (area subtended: mean r = 0.05, SD = 0.21; mean dot size: mean r = 0.02, SD = 0.19).

Procedure

The task was created and presented using OpenSesame [33]. A total of 384 trials were presented in 16 blocks with 24 calculation trials (12 high range & 12 low range) per block. Breaks were given between blocks. Operands were presented simultaneously to reduce working memory confounds. The problem was shown horizontally for 3s with either a multiplication (×) or division (÷) sign between the operands, followed by a screen with the 5 response choices arranged in a circle (Fig. 1). Responses were made using a mouse. The task advanced after a response was made or after a maximum of 4 seconds. The participants were told to answer quickly, even if they were not certain of the exact answer, and not to count the dots in non-symbolic problems.

Analysis

Data were visualized and analyzed using SPSS 20. To confirm that response choices were not random and check for a significant response bias, repeated measures ANOVA was used. Since interpretation of main effects in the presence of a significant interaction is not recommended [34], simple effects analysis was used when a significant interaction was present. The Bonferroni method was used to correct for multiple comparisons. When Mauchley's test of sphericity indicated that the assumption of sphericity had been violated, the Greenhouse-Geisser correction was used. Consistent with the notion of a logarithmically compressed mental magnitude representation and previous

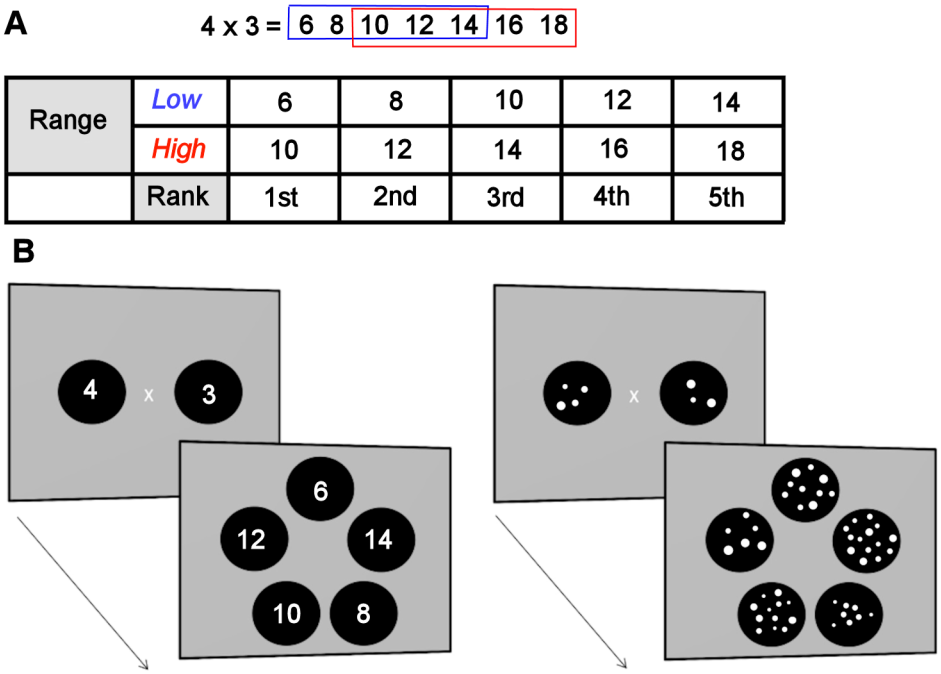

Figure 1. Task design. The correct result (C) and 6 incorrect results were created in a geometric series (symbolic: C x $1.5^{i/3}$ & non-symbolic: C x $2^{i/3}$; *i* from −3 to 3). To avoid the correct result corresponding to the 'middle one' among presented answer choices, only 5 of the 7 results were presented in a given trial. (A) A low range (blue) with the 4th choice correct and a high range (red) with the 2nd choice correct were created. (B) Illustration of the procedure using 4×3 low (4th choice correct), in symbolic (left) and non-symbolic (right) format. The problem was presented horizontally for 3s, followed by the answer choices for a maximum of 4s. Responses were made with a mouse.
doi:10.1371/journal.pone.0104777.g001

research [14], correct and response values were log-transformed prior to ANOVA [35].

Results

Nonrandom distribution of responses

To investigate the effects of notation, operation, rank (1–5) and range (4th or 2nd choice correct) on response percentage, a series of repeated measures ANOVAs were used. The correct choice can only be inferred using both range and rank variables. Thus, since response percentage results are only meaningful when both variables are considered, *interactions* including only one of these variables were not included. There was a significant interaction between notation, operation, rank and range ($F(4, 60) = 7.802$, $p < .001$, partial $\eta^2 = .342$), qualifying other main and lower-order interaction effects (notation × range × rank: $F(4, 60) = 259.672$, $p < .001$, partial $\eta^2 = .945$; operation × range × rank: $F(4, 60) = 5.488$, $p = .001$, partial $\eta^2 = .268$; range × rank: $F(4, 60) = 251.132$, $p < .001$, partial $\eta^2 = .944$; notation: $F(1, 15) = 6.808$, $p = .020$, partial $\eta^2 = .312$; operation: $F(1, 15) = 1.676$, $p = .215$, partial $\eta^2 = .100$; range: $F(1, 15) = 1.436$, $p = .249$, partial $\eta^2 = .087$; rank: $F(4, 60) = 84.525$, $p < .001$, partial $\eta^2 = .849$). Following up on the significant four-way interaction, we submitted response percentages to 2×5×2 repeated measures ANOVAs, separately for symbolic and non-symbolic notations, with the factors operation (multiplication vs. division), rank (1–5), and range (4th or 2nd choice correct). In symbolic problems, there was a significant interaction between operation, range and rank ($F(4, 60) = 17.054$, $p < .001$, partial $\eta^2 = .532$), qualifying other main and lower-order interaction effects (range × rank: $F(4, 60) = 333.555$, $p < .001$, partial $\eta^2 = .957$; operation: $F(1, 15) = 4.776$, $p = .045$, partial $\eta^2 = .242$; range: $F(1, 15) = 2.020$, $p = .176$, partial $\eta^2 = .119$; rank: $F(4, 60) = 218.421$, $p < .001$, partial $\eta^2 = .936$). In non-symbolic problems, there was no significant interaction between operation, range, and rank ($F(4, 60) = .978$, $p = .427$, partial $\eta^2 = .061$). The interaction between range and rank for all non-symbolic problems was significant ($F(4, 60) = 31.626$, $p < .001$, partial $\eta^2 = .678$), qualifying the main effects (operation: $F(1, 15) = 4.655$, $p < .048$, partial $\eta^2 = .237$; range: $F(1, 15) = .075$, $p < .787$, partial $\eta^2 = .005$; rank: $F(4, 60) = .334$, $p = .854$, partial $\eta^2 = .022$). Therefore, we analyzed the impact of rank and range using 2-way repeated measures ANOVAs, separately for each combination of notation and operation.

Influence of rank and range on response percentage

We first checked whether responses were non-randomly distributed, to confirm that participants were not guessing. Since five possible answer choices were used, with the rank of the correct choice depending on the range presented, random responding would be a flat line for both low and high ranges, with 20% of responses in each of the five choices. Based on visual inspection, responses appeared non-random in all conditions (Fig. 2). In symbolic problems, participants chose the correct answer 87% of the time for multiplication (4th choice correct: 1st = 1%, 2nd = 3%, 3rd = 3%, 4th = 87%, 5th = 5%; 2nd choice correct: 1st = 7%, 2nd = 87%, 3rd = 3%, 4th = 2%, 5th = 1%) and around 74% for division (4th choice correct: 1st = 4%, 2nd = 3%, 3rd = 7%, 4th = 75%, 5th = 11%; 2nd choice correct: 1st = 10%, 2nd = 73%, 3rd = 7%, 4th = 4%, 5th = 7%). In non-symbolic multiplication, participants chose close to the correct answer, with a trend towards overestimation, when the 4th choice was correct (1st = 9%, 2nd = 12%, 3rd = 22%, 4th = 26%, 5th = 31%), but chose randomly when the 2nd choice was correct (1st = 21%, 2nd = 18%, 3rd = 22%, 4th = 21%, 5th = 18%). In non-symbolic division, the opposite

Table 1. Multiplication and division problems and response choice values.

×	÷	Response Choice Values													
		Symbolic							*Non-symbolic*						
		Correct x $1.5^{i/3}$							*Correct x* $2^{i/3}$						
i		−3	−2	−1	0	1	2	3	−3	−2	−1	0	1	2	3
4×3	48/4	6	8	10	**12**	14	16	18	6	8	10	**12**	16	20	24
6×3	36/2	12	14	16	**18**	22	24	26	9	12	14	**18**	22	28	36
7×3	63/3	13	17	19	**21**	23	27	31	11	13	17	**21**	27	33	41
8×3	78/2	16	18	22	**24**	28	32	36	12	16	20	**24**	30	38	48
9×3	54/2	19	21	23	**27**	31	37	41	13	17	21	**27**	35	43	53
6×4	96/4	16	18	22	**24**	28	32	36	12	16	20	**24**	30	38	48
7×4	112/4	18	22	24	**28**	32	36	42	14	18	22	**28**	36	44	56
8×4	128/4	18	24	28	**32**	36	42	48	16	20	26	**32**	40	50	64
9×4	144/4	24	28	32	**36**	42	48	54	18	22	28	**36**	46	58	72
7×6	126/3	28	32	36	**42**	48	56	62	22	26	34	**42**	52	66	84
8×6	192/4	32	36	42	**48**	54	62	72	24	30	38	**48**	60	76	96
9×6	162/3	36	42	48	**54**	62	72	82	26	34	42	**54**	68	86	108
12×3	108/3	24	28	32	**36**	42	48	54	18	22	28	**36**	46	58	72
14×3	168/4	28	32	36	**42**	48	56	62	22	26	34	**42**	52	66	84
16×3	144/3	32	36	42	**48**	54	62	72	24	30	38	**48**	60	76	96
17×3	153/3	31	39	47	**51**	59	67	79	25	33	41	**51**	65	81	101
19×3	171/3	37	43	49	**57**	63	73	87	29	35	45	**57**	71	91	113
12×4	96/2	32	36	42	**48**	54	62	72	24	30	38	**48**	60	76	96
13×4	156/3	34	38	46	**52**	58	68	78	26	32	42	**52**	66	82	104
16×4	128/2	42	48	56	**64**	74	84	96	32	40	50	**64**	80	102	128
17×4	136/2	46	52	54	**68**	78	92	102	34	42	54	**68**	86	108	136
19×4	152/2	52	58	68	**76**	86	98	114	38	48	60	**76**	96	120	152
13×6	156/2	52	58	68	**78**	92	102	118	38	50	62	**78**	98	124	156
14×6	168/2	56	64	74	**84**	96	112	126	42	52	66	**84**	106	134	168

Response choice values were rounded to match parity of the correct choice value.

doi:10.1371/journal.pone.0104777.t001

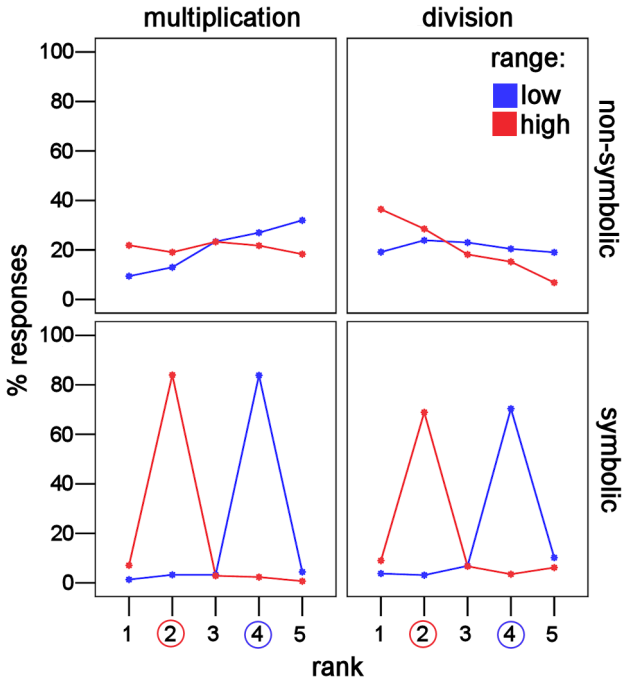

Figure 2. Non-random distribution of responses. Symbolic response percentages were non-random and peaked at the correct result. Rank (1–5) had a significant effect on response percentage for both low (blue, 4nd correct) and high (red, 2nd correct) ranges. Non-symbolic responses were non-random, depending on the response range presented. Rank had a significant effect on response percentage for multiplication when the low (blue) range was presented and for division when high (red) range was presented. This indicates that subjects were not guessing, but rather using a calculation based strategy.
doi:10.1371/journal.pone.0104777.g002

trend was seen. Participants chose randomly when the 4th choice was correct (1st = 18%, 2nd = 23%, 3rd = 22%, 4th = 19%, 5th = 18%), but chose close to the correct answer, with a tendency towards underestimation, when the 2nd choice was correct (1st = 35%, 2nd = 27%, 3rd = 17%, 4th = 14%, 5th = 6%).

This was confirmed using two-way repeated measures ANOVAs, separately for each condition, with response percentage as the dependent variable and rank of the response choice (1–5) and range (high: 2nd or low: 4th choice correct) as factors. When a significant interaction was present, simple effects analysis was performed to see whether rank had an effect on response percentage, separately for low and high ranges.

Symbolic multiplication. Responses were non-random regardless of range of response choices presented. The assumption of sphericity was violated according to Mauchley's Test of sphericity, $\chi^2(9) = 65.378$, $p<.001$; therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. The interaction between rank and range on response percentage was significant ($F(4, 60) = 421.783$, $p<.001$, partial $\eta^2 = .966$, $\varepsilon = .312$), qualifying significant main effects (range: $F(1, 15) = 2.246$, $p = .155$, partial $\eta^2 = .130$; rank: $F(4, 60) = 332.318$, $p<.001$, partial $\eta^2 = .957$). Therefore, a simple effects analysis was performed. There was a statistically significant difference in response percentage between the five response choices for both 2nd (partial $\eta^2 = .963$) and 4th (partial $\eta^2 = .968$) choice correct trials (table 2).

Non-symbolic multiplication. Responses were not random but, unlike symbolic calculations, this depended on the range of response choices presented. Responses where non-random when the 4th choice was correct but random when the 2nd choice was correct. The interaction between rank and range on response percentage was significant ($F(4,60) = 13.667$, $p<.001$, partial $\eta^2 = .477$), qualifying the significant main effects (range: $F(1, 15) = .024$, $p = .879$, partial $\eta^2 = .002$; rank: $F(4, 60) = 4.648$, $p = .002$, partial $\eta^2 = .237$). Therefore, a simple effects analysis was performed. There was a statistically significant difference in response percentage between the five response choices when the 4th (partial $\eta^2 = .433$) rather than the 2nd choice was correct (partial $\eta^2 = .049$) (table 2).

Symbolic division. Like symbolic multiplication, responses were not random regardless of the range of response choices presented. The assumption of sphericity was violated according to Mauchley's Test of sphericity, $\chi^2(9) = 66.359$, $p<.001$; therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. The interaction between rank and range on response percentage was significant ($f(4,60) = 188.257$, $p<.001$, partial $\eta^2 = .926$, $\varepsilon = .317$), qualifying significant main effects (range: $F(1, 15) = .135$, $p = .718$, partial $\eta^2 = .009$; rank: $F(4, 60) = 101.163$, $p<.001$, partial $\eta^2 = .871$). Therefore, a simple effects analysis was performed. There was a statistically significant difference in response percentage between the five response choices for both 2nd (partial $\eta^2 = .911$) and 4th (partial $\eta^2 = .905$) choice correct trials (table 2).

Non-symbolic division. Responses were not random but, similar to non-symbolic multiplication, this depended on the response range presented. In contrast to non-symbolic multiplication, responses were non-random when the 2nd rather than the 4th choice was correct. The interaction between rank and range on response percentage was significant ($F(4,60) = 18.765$, $p<.001$, partial $\eta^2 = .556$), qualifying significant main effects (range: $F(1, 15) = .135$, $p = .718$, partial $\eta^2 = .009$; rank: $F(4, 60) = 6.827$, $p< .001$, partial $\eta^2 = .313$). Therefore, a simple effects analysis was performed. There was a statistically significant difference in response percentage between the five response choices for 2nd choice correct trials ($\eta^2 = .531$), but not 4th choice correct trials (partial $\eta^2 = .053$) (table 2).

Linear increase of response value with correct value

Before analyzing response bias, we wanted to determine whether the logarithm of response and correct values should be used as in past research [14,16], and in-line with statistical recommendations [35]. Weber's law predicts that the variability of response values will increase with numerical magnitude. On the linear scale, mean response value and variability increased as a function of the correct value, whereas on the log scale, variability was constant (Fig. 3 B, C). To confirm this, we plotted the original linear and log-transformed response value as a function of the linear and log-transformed correct value (Fig. 3A) and tested the slope against a null value of zero using multi-level modeling, with participants as a random effect (table 3). There was significant linear dependence of the mean chosen value on the correct value for all conditions, for both the linear and log-transformed data. For all conditions, the rate of change of the conditional mean of the response value with respect to the correct value was greater than zero, for both the linear (symbolic multiplication: B = 0.9763, 95% C.I. [0.9656, 0.9871]; non-symbolic multiplication: B = 1.2427, 95% C.I. [1.1975, 1.2880]; symbolic division: B = 0.9868, 95% C.I. [0.9724, 1.0013]; non-symbolic division: B = 0.9306, 95% C.I. [0.8937, 0.9675]) and log-transformed data (symbolic multiplication: B = 0.9890, 95% C.I. [0.9810, 0.9970]; non-symbolic multiplication: B = 1.0731, 95% C.I. [1.0421, 1.0423]; symbolic division: B = 0.9809, 95% C.I. [0.9677–

Table 2. Effect of response choice rank on response percentage for low and high range.

					Sphericity	
Range	**df**	**F**	***P***	**Partial η^2**	χ^2	ε
Symbolic multiplication						
Low, 4th correct*	4,60	452.392	<.001	.968	69.344*	.333
High, 2nd correct*	4,60	391.147	<.001	.963	92.150*	.292
Non-symbolic multiplication						
Low, 4th correct*	4,60	11.460	<.001	.433	27.291*	.453
High, 2nd correct	4,60	0.777	.460	.049	29.160*	.463
Symbolic division						
Low, 4th correct*	4,60	142.784	<.001	.905	75.556*	.305
High, 2nd correct*	4,60	153.190	<.001	.911	69.293*	.313
Non-symbolic division						
Low, 4th correct	4,60	0.839	.447	.053	28.050*	.527
High, 2nd correct*	4,60	16.979	<.001	.531	35.074*	.475

Bonferroni corrected for multiple comparisons and Greenhouse-Geisser corrected for violations of sphericity as measured by Mauchley's Test of Sphericity.
*p<.001.
doi:10.1371/journal.pone.0104777.t002

0.9941]; non-symbolic division: B = 0.9542, 95% C.I. [0.9238, 0.9845]). The log-transformed response and correct values appeared to better prepare the data for ANOVA since constant variation is assumed. Therefore, the logarithm of the response and correct value was used in all OM analyses.

Contribution of both operands to response value

To determine whether participants considered both operands when choosing a response value, we performed multi-level multiple regression, separately for symbolic multiplication, non-symbolic multiplication, symbolic division, and non-symbolic division, using log-transformed values and participants as a random effect. For all conditions, there was a significant ($p<$ .001) contribution of both operands to the mean response value and a significant rate of change of the conditional mean of the response value with respect to the first (op1) and second (op2) operands (symbolic multiplication: op1 $t = 208.252$, B = 0.9951, 95% C.I. [0.9857, 1.0045]; op2 $t = 127.287$, B = 0.9731, 95% C.I. [0.9581, 0.9881]; non-symbolic multiplication: op1 $t = 51.852$, B = 1.0990, 95% C.I. [1.0574, 1.1406]; op2 $t = 29.389$, B = 0.9962, 95% C.I. [0.9297, 1.0627]; symbolic division: op1 $t = 131.116$, B = 1.0105, 95% C.I. [0.9953, 1.0256]; op2 $t = -81.827$, B = −0.9733, 95% C.I. [−0.9966, −0.9500]; non-symbolic division: op1 $t = 48.655$, B = 0.9873, 95% C.I. [0.9475, 1.0271]; op2 $t = -29.691$, B = −0.9266, 95% C.I. [−0.9878, −0.8653]). The positive slopes for op2 in multiplication problems and negative slopes in division problems are consistent with the operations since a larger 2nd operand in division would result in a smaller result value. Based on the conservative test of non-overlapping confidence intervals [36,37], magnitudes of the slopes (absolute value of ***B***) where not significantly different between op1 and op2. These findings suggest that participants based their response on a combination of both operands and provide evidence against pure guessing.

Taken together, these results imply that participants did not consistently use a random guessing strategy. Rather, they relied on both operands, although perhaps not to an equal degree, to formulate a response. This supports the use of approximate calculation versus consideration of one operand. For symbolic problems, choices clearly peaked at the correct response. For non-symbolic problems, the pattern of results was more complex. However, the interaction between rank, range, and operation in non-symbolic problems implies that participants' choices depended on the range of presented response alternatives in a given trial. Since the two ranges were presented in random order and participants were unaware of the low/high range design, the results are unlikely to be due to a completely non-numeric strategy. The increase of the mean chosen value as a function of the correct value, in all conditions, further supports this interpretation.

Operational momentum effect

To investigate operational momentum, we looked at the response bias, defined as the difference between the log chosen and the log correct values. To test the influence of operation and notation on response bias, a 2-way repeated measures ANOVA was used. The interaction of operation and notation had a significant effect on response bias ($F(1,15) = 16.023$, $p = .001$, partial $\eta^2 = .516$), qualifying significant main effects (operation: $F(1,15) = 14.077$, $p = .002$, partial $\eta^2 = .484$; notation: $F(1,15) = .297$, $p = .594$, partial $\eta^2 = .019$). Therefore, simple effects analysis was performed to see whether operation had an effect on mean response bias, separately for non-symbolic and symbolic notations.

Non-symbolic notation. For non-symbolic problems, operation had a significant effect on response bias at the Bonferroni corrected $p<.025$ level ($F(1,15) = 15.315$, $p = .001$, partial $\eta^2 = .505$). There was a significant difference in the mean log response bias between non-symbolic multiplication and division (M = 0.069, Bonferroni 95% C.I. [0.031, 0.106, p = .001]). To see if participants overestimated multiplication (mean log response bias>0) and underestimated division (mean log response bias<0), we performed one-sample t-tests against a null value of zero. We found that participants significantly overestimated multiplication problems ($t(15) = 2.449$, M = 0.02987, 95% C.I. [0.0039, 0.0559], $p = .027$) and underestimated division problems ($t(15) = -3.136$, M = −0.03879, 95% C.I. [−0.0652, −0.0124], $p = .007$). These results indicate that non-symbolic response bias is significantly

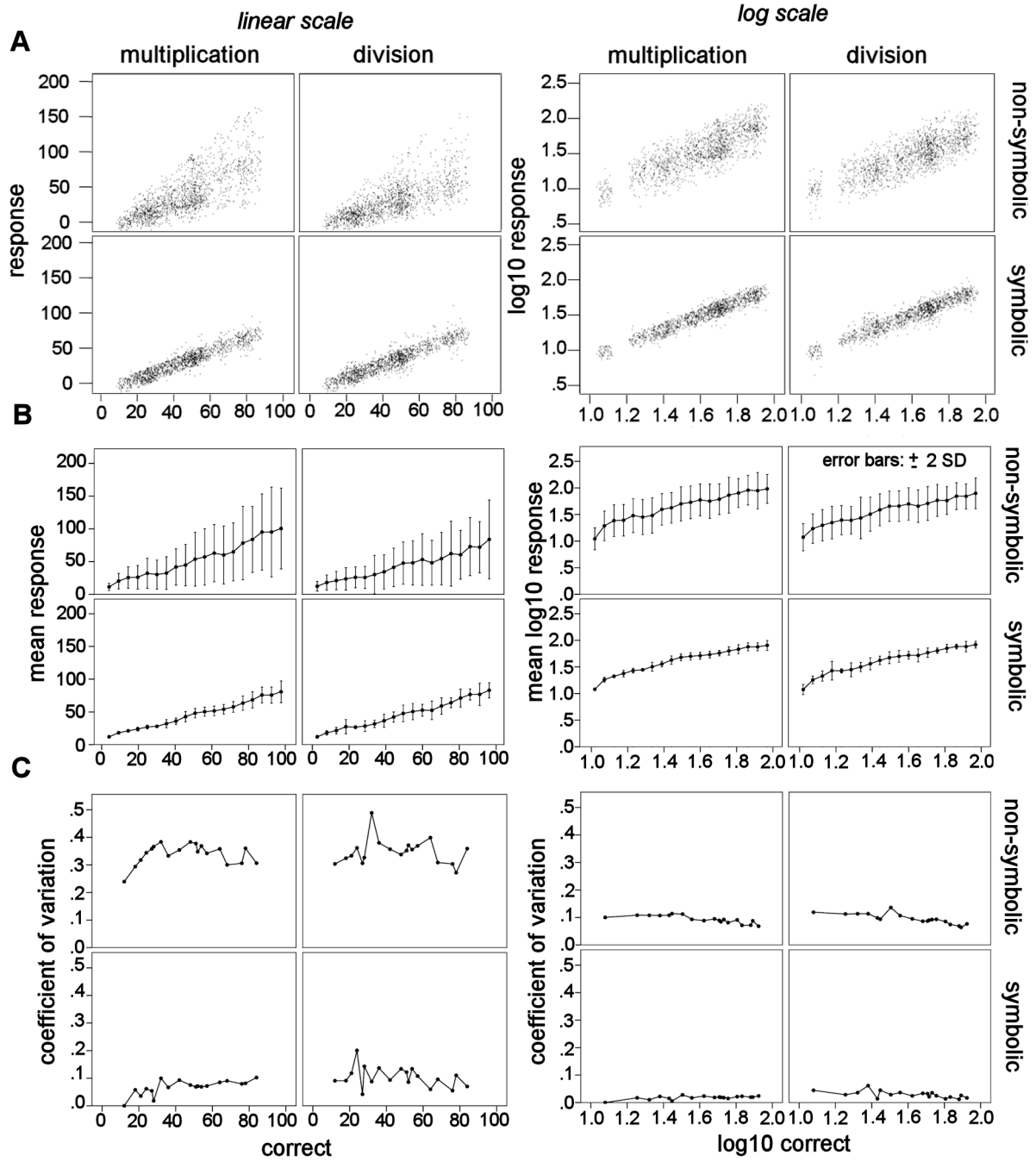

Figure 3. Response value as a function of correct value on linear and log-scale data. (A) Non-aggregated response value as a function of correct value, on the linear (left) and log (right) scale. Number of cases is shown by increased density (*i.e.* darker color). (B) Linear response value and SD increased as a function of correct value, consistent with Weber's law. Log-transformed response value, but not SD, increased as a function of log-transformed correct value (*i.e.* linear on the log scale). (C) Dispersion of response choices, measured by the coefficient of variation, was constant across correct values on the log scale, but not on the linear scale. Dispersion was constant when log-transformed values were used.
doi:10.1371/journal.pone.0104777.g003

influenced by operation. Consistent with our hypothesis, non-symbolic multiplication problems were overestimated while division problems were underestimated (Fig. 4).

Symbolic notation. For symbolic problems, operation did not have a significant effect on response bias ($F(1,15) = 4.049$, $p = .063$, partial $\eta^2 = .213$). There was not a significant difference in the mean log response bias between symbolic multiplication and division (M = −.005, Bonferroni 95% C.I. [−0.01, 0.00], p = .063). The mean log response bias was not significantly different from zero in multiplication (t(15) = −1.365, M = −0.002, 95% C.I. [−0.005, 0.001], $p = .192$) or division (t(15) = 1.538, M = 0.003, 95% C.I. [−0.001, 0.006], $p = .145$) (Fig. 4).

Discussion

Our primary goal in the present study was to determine if there was an OM effect for multiplication and division, like for addition

Table 3. Linear increase of response value with correct value.

	Linear Scale			*Log scale*		
	t	**Slope**	**95% CI**	***t***	**Slope**	**95% CI**
Multiplication						
Symbolic	178.3*	.9763	.9656–.9871	242.7*	.9890	.9810–.9970
Non-symbolic	53.9*	1.2427	1.1975–1.2880	67.7*	1.0731	1.0421–1.0423
Division						
Symbolic	133.8*	.9868	.9724–1.0013	145.4*	.9809	.9677–.9941
Non-symbolic	49.4*	.9306	.8937–.9675	61.6*	.9542	.9238–.9845

Bonferroni corrected for multiple comparisons.
*$p<.013$.
doi:10.1371/journal.pone.0104777.t003

and subtraction. Our second goal was to see if participants could use the ANS to perform non-symbolic multiplication and division, on larger quantities than previously studied. We hypothesized that participants could perform non-symbolic multiplication and division and would overestimate for multiplication, as they do for addition, and underestimate for division, as they do for subtraction. We found that participants could perform non-symbolic multiplication and division and their response patterns were consistent with use of the ANS. Participants significantly overestimated non-symbolic multiplication problems and underestimated non-symbolic division problems. Unlike symbolic addition and subtraction, for symbolic problems we observed no significant modulation of responses by operation. These findings expand the mathematical operations subject to response bias to include non-symbolic whole-number multiplication and division.

Non-symbolic multiplication and division ability

To our knowledge, this is the first study to look at non-symbolic multiplication and division in adults. Previous research has focused on children and used smaller numbers (*e.g.* halving/doubling) [28,29]. Based on the observed response pattern, participants likely used an approximate calculation based strategy for non-symbolic multiplication and division. If participants had ignored all numeric information and responded randomly, we would have seen a flat distribution across response choices (20% for each answer choice) both when the 2nd and 4th choices were correct (Fig. 2). This was obviously not the case, as seen in figure 2 and the significant interaction between rank and range in both notations. Alternatively, if participants had used a heuristic of choosing a relatively large number of dots for multiplication and small for division, we would have seen a distribution peaked at the high or low end of response choices, regardless of the response range presented. Again, this hypothesis was not supported by the data. Instead, participants overestimated multiplication when the 4th of five answer choices was correct, but seemed to guess when the 2nd lowest choice was correct. The reverse was found for division. Participants underestimated when the 2nd choice was correct, but seemed to guess when the 4th choice was correct. Since participants were not aware of the experimental design, let alone when they were answering in the low or high range, the consistent differences between ranges were most likely driven an approximate evaluation of the operands. Additionally, the mean chosen value increased with correct value (Fig. 3) and both the first and second operands independently contributed to the response value. Taken together, these findings indicate that participants were using calculation strategies that were influenced by operation. The influence of operation is consistent with the presence of OM in non-symbolic problems. Similar to addition and subtraction, the ANS might be used to solve whole-number non-symbolic multiplication and division. However, it should be noted that OM may be driven by non-calculation based strategies as was seen in infants [15]. Thus, the likely approximate calculation we have demonstrated is not a precondition for OM.

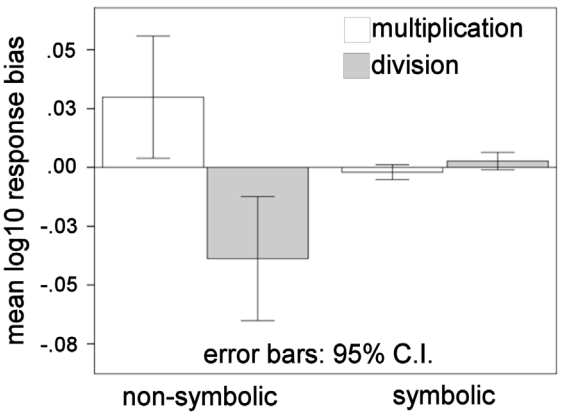

Figure 4. Operational momentum (OM) in non-symbolic, but not symbolic, notation. A significant response bias occurred for non-symbolic problems, indicating an OM effect. Symbolic response bias was not significant. Non-symbolic multiplication (light grey) problems were overestimated and division (dark gray) problems were underestimated. Response bias was calculated as the $\log_{10}$ response value – $\log_{10}$ correct value. Positive response bias indicates overestimation and negative indicates underestimation. Error bars represent 95% confidence interval (C.I.).
doi:10.1371/journal.pone.0104777.g004

Operational momentum effect in non-symbolic multiplication and division

This is the first study to look at OM in whole-number multiplication and division. Consistent with past research [13–15], we found an OM effect in non-symbolic calculations. Specifically, participants overestimated for multiplication and underestimated for division. Finding an OM effect is reminiscent of the whole-number bias [19,20], as well as an extension of the 'multiplication makes bigger, division makes smaller' (than the original quantity) (MMBDMS) belief [22]. All of the response alternatives fit this belief, yet there was a bias towards over or underestimating. That is, over and above the predictions of the

MMBDMS belief, we observed a modulation of mean chosen value by operation.

Three hypotheses for OM have been proposed: First, the compression hypothesis states that flawed decompression from the log scale results in response bias [13,26]; however our data do not directly speak to this issue. Second, the attentional shifts hypothesis states that OM occurs as a result of left/right shifts of attention along a mental number line and a preference for outcomes in the whole-number biased [19,20,22] direction of the calculation [14]. Our findings support a preference for outcomes in the whole-number biased direction of the operation, although we did not test the role of attention. Finally, the heuristic (MMBDMS, in our study) hypothesis explains OM as using a rule of accepting more than the original operand for addition and less for subtraction [15,19,20,22]. In principle, this could apply to multiplication and division. However, in our study the response choices were always numerically larger than both operands for multiplication and numerically smaller than the first operand for division. Therefore, any response would fit this rule. We found a bias within the presented choices even though they were equally likely to be chosen based on this heuristic. However, the influence of the range of presented response choices suggests that a similar heuristic could partially explain our findings. A combination of heuristic bias, similar to MMBDMS, and approximate calculation might best explain OM in whole-number multiplication and division. Based on the current results, we cannot rule out the possibility that approximate calculation was influenced by attentional shifts. Further research is needed to clarify the role of attention during approximate mental arithmetic.

The difference between low (4^{th} choice correct) and high (2^{nd} choice correct) range response choices suggests a more complex strategy than hypothesized for addition and subtraction. If participants had overestimated in multiplication (*i.e.* 2^{nd} correct: chosen 3^{rd}, 4^{th}, 5^{th}; 4^{th} correct: chosen 5^{th}) and underestimated in subtraction (*i.e.* 2^{nd} correct: chosen 1^{st}; 4^{th} correct: chosen 1^{st}, 2^{nd}, 3^{rd}), then a directionally biased approximate calculation hypothesis would explain our findings. Since participants were naïve to the study aims and manipulated factors (*e.g.* range), it remains to be seen what determines strategy choice in a given trial. If participants had chosen the largest response choice for multiplication (*i.e.* chosen 5^{th} in both 2^{nd} and 4^{th} correct) and smallest for division (*i.e.* chosen 1^{st} in both 2^{nd} and 4^{th} correct), then a modified MMBDMS heuristic hypothesis would explain our findings. Since they overestimated in multiplication only when the 4^{th} choice was correct and underestimated in division only when the 2^{nd} choice was correct, a combination of the two hypotheses might best explain our findings. Another possibility is that the ratio between the largest (multiplication) or smallest (division) response alternative and the correct outcome was too small for participants to exclude extreme results. For the high response range in multiplication and low in division, even extreme response alternatives were not considered too large (multiplication, high range) or small (division, low range), leading to a lack of tapering. However, this lack of tapering could be due to the operational momentum effect. That is, the operational momentum effect might be the reason why extreme values (too large for multiplication or small for division) did not seem extreme enough to exclude. More empirical data is needed to disentangle these possibilities. Thus, although our data cannot be explained by the traditional MMBDMS bias, a related heuristic strategy incorporating approximate calculation seems likely. This might be described as 'multiplication makes relatively large, division makes relatively small' (MRLDRS).

Although past OM research assumes that participants use a single strategy, they might use multiple strategies, especially for more difficult tasks. Symbolic arithmetic is thought to use global processes to evaluate solutions alongside fact retrieval [38,39]. These biased global processes may originate from, or be exacerbated by, early educational methods [22]. Similarly, both heuristic evaluation (multiplication = relatively large answer; division = relatively small answer) (MRLDRS) and approximate calculation may be used for non-symbolic calculation and their interaction might explain OM. When a plausible response choice (dot array) based on the mentally represented approximate calculation is small relative to alternatives for multiplication (high range, 2^{nd} correct) or large for division (low range, 4^{th} correct), the approximate calculation and the MRLDRS heuristic evaluation lead to different response choices. When this conflicting information is present, accuracy is likely to decrease [40]. This is also inline with the role of inhibitory control in numerical cognition [41], including OM [16].

In contrast to some research [14,17,18], we did not find OM in symbolic problems. The inclusion of the correct answer might have made performance too accurate to detect response bias. Studies finding OM in symbolic arithmetic have used approximate response methods, such as pointing to a line marked only with endpoint numbers [17], manual dot array generation [18], or jittering the correct result [14]. Though children may use the ANS to support symbolic arithmetic [1,42,43], reliance on rote verbal memory may limit ANS influence in adults [44]. Alternatively, regrouping performed in multiplication and division problems may prevent, and even reverse, OM. This explanation has been proposed for reverse OM in symbolic addition and subtraction [18]. A final possibility is that adult exposure to multiplication and division with rational numbers attenuates bias in whole-number symbolic calculation [22]. However, the demonstration of whole-number bias in adults suggests that directional bias is not fully corrected [19,20]. Future research could use an approximate response method and a regrouping variable to understand whole-number symbolic OM.

Limitations

This study has some limitations. Only five answer choices were presented, which might have put a ceiling or floor effect on response bias. Despite this, we were able to demonstrate OM. We also chose to keep the correct value identical across operations and to roughly match the 2^{nd} operands. Therefore, the size of the first operands was not matched between multiplication and division problems. However, if participants relied on the first operand we would expect, if anything, a reverse OM effect since division had larger first operands than multiplication.

Conclusions

We have demonstrated that OM occurs in whole-number multiplication and division. This is the first time OM has been found in scalar operations. Additionally, we have shown that adults do not randomly guess or use a purely heuristic strategy, but rather use approximation, based on the operands, to perform non-symbolic multiplication and division. Non-symbolic multiplication problems are overestimated and non-symbolic division problems are underestimated. Interestingly, response patterns depend on the magnitude of the correct choice relative to the alternatives. These findings suggest that a combination of approximate calculation and an operationally dependent bias towards large or small quantities might explain OM. When multiple choices are given, response may depend on an interaction between approximate calculation and a heuristic evaluation. This interaction could

reconcile these two previously proposed explanations. Future research should consider the use of multiple strategies, depending on difficulty and task design. However, regardless of the response strategy, the demonstration of OM in multiplication and division advances understanding of this phenomenon and shows that OM can be found in all whole-number arithmetic operations.

Acknowledgments

We thank Carolin Utecht and Hannes Hösterey for their help with recruitment and data collection and Seda Cavdaroglu for her helpful comments on an earlier draft of this paper. We would also like to thank academic editor Michael Motes and the two anonymous reviewers for their helpful suggestions as well as Brian Greer for providing his book chapter.

Author Contributions

Conceived and designed the experiments: CK AK. Performed the experiments: CK. Analyzed the data: CK. Wrote the paper: CK AK.

References

1. Gilmore CK, McCarthy SE, Spelke ES (2007) Symbolic arithmetic knowledge without instruction. Nature 447: 589–591. doi:10.1038/nature05850.
2. Xu F, Spelke ES, Goddard S (2005) Number sense in human infants. Dev Sci 8: 88–101. doi:10.1111/j.1467-7687.2005.00395.x.
3. Livingstone MS, Srihasam K, Morocz IA (2010) The benefit of symbols: monkeys show linear, human-like, accuracy when using symbols to represent scalar value. Anim Cogn 13: 711–719. doi:10.1007/s10071-010-0321-1.
4. McCloskey M (1992) Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. Cognition 44: 107–157. doi:10.1016/0010-0277(92)90052-J.
5. Peters E (2012) Beyond Comprehension: The Role of Numeracy in Judgments and Decisions. Curr Dir Psychol Sci 21: 31–35. doi:10.1177/0963721411429960.
6. Dehaene S, Cohen L (1995) Towards an anatomical and functional model of number processing. In: Butterworth B, editor. Mathematical Cognition 1, No. East Sussex, UK: Psychology Press, Vol. 1. 83–120.
7. Harvey BM, Klein BP, Petridou N, Dumoulin SO (2013) Topographic representation of numerosity in the human parietal cortex. Science 341: 1123–1126. doi:10.1126/science.1239052.
8. Dehaene S (2003) The neural basis of the Weber-Fechner law: a logarithmic mental number line. Trends Cogn Sci 7: 145–147. doi:10.1016/S1364-6613(03)00055-X.
9. Fias W (1996) The Importance of Magnitude Information in Numerical Processing: Evidence from the SNARC Effect. Math Cogn 2: 95–110. doi:10.1080/135467996387552.
10. Fischer MH (2001) Number processing induces spatial performance biases. Neurology 57: 822–826. doi:10.1212/WNL.57.5.822.
11. Nicholls MER, Loftus a M, Gevers W (2008) Look, no hands: A perceptual task shows that number magnitude induces shifts of attention. Psychon Bull Rev 15: 413–418. doi:10.3758/PBR.15.2.413.
12. Fischer MH, Castel AD, Dodd MD, Pratt J (2003) Perceiving numbers causes spatial shifts of attention. Nat Neurosci 6: 555–556. doi:10.1038/nn1066.
13. McCrink K, Dehaene S, Dehaene-Lambertz G (2007) Moving along the number line: Operational momentum in nonsymbolic arithmetic. Percept Psychophys 69: 1324–1333. doi:10.3758/BF03192949.
14. Knops A, Viarouge A, Dehaene S (2009) Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. Atten Percept Psychophys 71: 803–821. doi:10.3758/APP.71.4.803.
15. McCrink K, Wynn K (2009) Operational momentum in large-number addition and subtraction by 9-month-olds. J Exp Child Psychol 103: 400–408. doi:10.1016/j.jecp.2009.01.013.
16. Knops A, Zitzmann S, McCrink K (2013) Examining the presence and determinants of operational momentum in childhood. Front Psychol 4: 325. doi:10.3389/fpsyg.2013.00325.
17. Pinhas M, Fischer MH (2008) Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. Cognition 109: 408–415. doi:10.1016/j.cognition.2008.09.003.
18. Lindemann O, Tira MD (2011) Operational Momentum in Numerosity Production Judgements of Multi-digit Number Problems. J Psychol 219: 50–57.
19. Vamvakoussi X, Van Dooren W, Verschaffel L (2012) Naturally biased? In search for reaction time evidence for a natural number bias in adults. J Math Behav 31: 344–355. doi:10.1016/j.jmathb.2012.02.001.
20. Vamvakoussi X, Dooren W, Verschaffel L (2013) Educated adults are still affected by intuitions about the effect of arithmetical operations: evidence from a reaction-time study. Educ Stud Math 82: 323–330. doi:10.1007/s10649-012-9432-8.
21. Ni Y, Zhou Y-D (2005) Teaching and Learning Fraction and Rational Numbers: The Origins and Implications of Whole Number Bias. Educ Psychol 40: 27–52. doi:10.1207/s15326985ep4001_3.
22. Greer B (1994) Extending the meaning of multiplication and division. In: Harel G, Confrey J, editors. The development of multiplicative reasoning. Albany, NY: State University of New York Press. 61–85.
23. Hubbard EM, Piazza M, Pinel P, Dehaene S (2005) Interactions between number and space in parietal cortex. Nat Rev Neurosci 6: 435–448. doi:10.1038/nrn1684.
24. Knops A, Thirion B, Hubbard EM, Michel V, Dehaene S (2009) Recruitment of an area involved in eye movements during mental arithmetic. Science 324: 1583–1585. doi:10.1126/science.1171599.
25. Izard V, Dehaene S (2008) Calibrating the mental number line. Cognition 106: 1221–1247. doi:10.1016/j.cognition.2007.06.004.
26. Chen Q, Verguts T (2012) Spatial intuition in elementary arithmetic: a neurocomputational account. PLoS One 7: e31180. doi:10.1371/journal.pone.0031180.
27. Klein E, Huber S, Nuerk H-C, Moeller K (2014) Operational momentum affects eye fixation behaviour. Q J Exp Psychol Advance online publication. doi:10.1080/17470218.2014.902976.
28. Barth H, Baron A, Spelke E, Carey S (2009) Children's multiplicative transformations of discrete and continuous quantities. J Exp Child Psychol 103: 441–454. doi:10.1016/j.jecp.2009.01.014.
29. McCrink K, Spelke ES (2010) Core multiplication in childhood. Cognition 116: 204–216. doi:10.1016/j.cognition.2010.05.003.
30. Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10: 433–436. doi:10.1163/156856897X00357.
31. Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10: 437–442. doi:10.1163/156856897X00366.
32. Gebuis T, Reynvoet B (2011) Generating nonsymbolic number stimuli. Behav Res Methods 43: 981–986. doi:10.3758/s13428-011-0097-5.
33. Mathôt S, Schreij D, Theeuwes J (2012) OpenSesame: an open-source, graphical experiment builder for the social sciences. Behav Res Methods 44: 314–324. doi:10.3758/s13428-011-0168-7.
34. Keppel G, Wickens TD (2004) Design and Analysis: A Researcher's Handbook. 4th ed. Upper Saddle River, NJ, USA: Prentice Hall.
35. Berry DA (1987) Logarithmic Transformations in ANOVA. Biometrics 43: 439. doi:10.2307/2531826.
36. Payton ME, Greenstone MH, Schenker N (2003) Overlapping confidence intervals or standard error intervals: What do they mean in terms of statistical significance? J Insect Sci 3: 34.
37. Katz C, Yaseen ZS, Mojtabai R, Cohen LJ, Galynker II (2011) Panic as an independent risk factor for suicide attempt in depressive illness: findings from the National Epidemiological Survey on Alcohol and Related Conditions (NESARC). J Clin Psychiatry 72: 1628–1635. doi:10.4088/JCP.10m06186blu.
38. Ashcraft MH, Stazyk EH (1981) Mental addition: A test of three verification models. Mem Cognit 9: 185–196. doi:10.3758/BF03202334.
39. LeFevre J-A, Bisanz J, Daley KE, Buffone L, Greenham SL, et al. (1996) Multiple routes to solution of single-digit multiplication problems. J Exp Psychol Gen 125: 284–306. doi:10.1037/0096-3445.125.3.284.
40. Friedman D, Nessler D, Cycowicz YM, Horton C (2009) Development of and change in cognitive control: a comparison of children, young adults, and older adults. Cogn Affect Behav Neurosci 9: 91–102. doi:10.3758/CABN.9.1.91.
41. Gilmore C, Attridge N, Clayton S, Cragg L, Johnson S, et al. (2013) Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. PLoS One 8: e67374. doi:10.1371/journal.pone.0067374.
42. Gilmore CK, McCarthy SE, Spelke ES (2010) Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. Cognition 115: 394–406. doi:10.1016/j.cognition.2010.02.002.
43. Inglis M, Attridge N, Batchelor S, Gilmore C (2011) Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. Psychon Bull Rev 18: 1222–1229. doi:10.3758/s13423-011-0154-1.
44. Dehaene S, Piazza M, Pinel P, Cohen L (2003) Three parietal circuits for number processing. Cogn Neuropsychol 20: 487–506. doi:10.1080/02643290244000239.
45. Otto A, Hancock SW, Caldwell J (2011) Developing Essential Understanding of Multiplication and Division for Teaching Mathematics in Grades 3–5. First edit. Zbiek RM, editor Reston: National Council of Teachers of Mathematics.

Behavioral/Cognitive

A Shared, Flexible Neural Map Architecture Reflects Capacity Limits in Both Visual Short-Term Memory and Enumeration

André Knops,[1,2] Manuela Piazza,[1,3,4]* Rakesh Sengupta,[5] Evelyn Eger,[3,4] and David Melcher[1]*

[1]Center for Mind/Brain Sciences, University of Trento, I-38068 Rovereto, Italy, [2]Department of Psychology, Humboldt University of Berlin, D-12489 Berlin, Germany, [3]National Institute of Health and Medical Research, Unit 992, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France, [4]Atomic Energy and Alternative Energies Commission, Departement of Life Sciences (DSV)/Biomedical Imaging Institute (I2BM), NeuroSpin Center, F-91191 Gif/Yvette, France, and [5]Center for Neural and Cognitive Sciences, University of Hyderabad, 500046 Hyderabad, India

Human cognition is characterized by severe capacity limits: we can accurately track, enumerate, or hold in mind only a small number of items at a time. It remains debated whether capacity limitations across tasks are determined by a common system. Here we measure brain activation of adult subjects performing either a visual short-term memory (vSTM) task consisting of holding in mind precise information about the orientation and position of a variable number of items, or an enumeration task consisting of assessing the number of items in those sets. We show that task-specific capacity limits (three to four items in enumeration and two to three in vSTM) are neurally reflected in the activity of the posterior parietal cortex (PPC): an identical set of voxels in this region, commonly activated during the two tasks, changed its overall response profile reflecting task-specific capacity limitations. These results, replicated in a second experiment, were further supported by multivariate pattern analysis in which we could decode the number of items presented over a larger range during enumeration than during vSTM. Finally, we simulated our results with a computational model of PPC using a saliency map architecture in which the level of mutual inhibition between nodes gives rise to capacity limitations and reflects the task-dependent precision with which objects need to be encoded (high precision for vSTM, lower precision for enumeration). Together, our work supports the existence of a common, flexible system underlying capacity limits across tasks in PPC that may take the form of a saliency map.

Introduction

Visual cognition is characterized by high flexibility but also capacity limits. Although the visual system can adapt its representational accuracy, the number of items concurrently processed is limited: in tasks as different as rapid object enumeration or visual short-term memory (vSTM), subjects can only process three or four items at a time. These capacity limits may reflect a general mechanism of object individuation (Piazza et al., 2011; Wutz and Melcher, 2013), commonly accessed in many different attentional tasks and that we suggested may take the form of a saliency (or priority) map (Bisley and Goldberg, 2003). Saliency maps topographically represent the conspicuity (or "saliency") of items at every location. Map-like architectures for spatial attention have been observed previously in the monkey lateral intraparietal (LIP) area (Bisley and Goldberg, 2003) and the putative human homolog posterior parietal cortex (PPC; Connolly et al., 2002). Critically, PPC has been implicated in studies of capacity limits in both enumeration (Piazza et al., 2002) and vSTM (Todd and Marois, 2004), as well as in visuospatial attention tasks in general, suggesting a shared neural substrate for capacity limits across tasks (Colby and Goldberg, 1999). Proof for the hypothesis of shared neural systems across tasks remains scarce because of a lack of studies investigating more than one task at a time (but see Silk et al., 2010). Here we directly test the hypothesis that a map architecture in human PPC (Gottlieb, 2007; Bays et al., 2010; Melcher and Piazza, 2011; Franconeri et al., 2013) represents individual items with a flexible degree of precision (e.g., modulable by context and task requirements) and reflects capacity limits across different tasks. Recent empirical and computational evidence link lateral inhibition strength between items to the precision of represented items within a map (Roggeman et al., 2010; Dempere-Marco et al., 2012; Sengupta and Melcher, 2014). High inhibition reduces the noise within a map, allowing for precise representations of items, but restricts capacity to few items. Conversely, low inhibition allows for a larger number of items to be represented yet less precisely. The representational precision of a given item varies with the observer's current goals. Whereas in a vSTM task participants encode multiple features, such as location and orientation of items, in enumeration tasks, no precise encoding of object features is necessary. The mere individuation of items is sufficient to encode them as units (Melcher and Piazza, 2011; Wutz and Melcher, 2013).

Here, we manipulated the required representational precision of objects by engaging participants in two tasks: (1) a vSTM task,

Received June 28, 2013; revised May 19, 2014; accepted May 22, 2014.

Author contributions: A.K., M.P., and D.M. designed research; A.K. and R.S. performed research; A.K., R.S., and E.E. analyzed data; A.K., M.P., E.E., and D.M. wrote the paper.

This work was supported by the Italian Ministry of Universities and Research (Research Programs of Relevant National Interest 2009), the Caritro Foundation (Fondazione Cassa di Risparmio di Trento e Rovereto), German Research Foundation Grant KN 959/2 (A.K.), India-Trento Programme for Advanced Research (R.S.), and European Research Council Starting Grant 313658 (D.M.).

*M.P. and D.M. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. André Knops, Department of Psychology, Humboldt University of Berlin, Rudower Chaussee 18, D-12489 Berlin, Germany. E-mail: knops.andre@gmail.com.

DOI:10.1523/JNEUROSCI.2758-13.2014

 0270-6474/14/349857-10$15.00/0

requiring high encoding precision; and (2) an enumeration task, requiring low encoding precision. This allowed us to test whether there were common PPC maps for the two tasks with changing neural response profile dependent on task demands. We predict a nonlinear increase of PPC activation with increasing number of to-be-encoded objects in a task-dependent manner: in a task requiring low precision, activation should increase only when the set exceeds three or four items, whereas in a task requiring high precision, activation should increase already beyond one item.

Materials and Methods

The current study comprises two experiments, both described below: (1) a main experiment; and (2) a control experiment.

Main experiment

Participants. A total of 19 healthy adults with normal or corrected-to-normal vision and no history of neurological or psychiatric illness participated in the study, which was approved by the Ethics Committee of the University of Trento. Two participants were excluded from subsequent analysis because of extensive head motion during scanning. All subsequent analyses are based on data from 17 participants (seven females; mean ± SD age, 25.78 ± 10.3 years).

Stimuli. Stimuli consisted of a variable number of Gaussian modulated sinusoidal grating (Gabor) patches. A given Gabor in a set was individually tilted from vertical to the left or right with a random angle between 15° and 45°. In the main experiment, subjects performed two tasks in different fMRI runs of a single session: (1) enumeration; and (2) vSTM (see below). For the enumeration task, the number varied between one and eight, whereas for the vSTM, it varied between one and six. Each numerosity was presented equally often and at least twice in each of four blocks. Numerosity was fully crossed with saliency that had two levels. For low-saliency displays, all the Gabors had the same contrast of 35%. In high-salient displays, one Gabor was flickering at 20 Hz between 100 and 33% contrast. Because this saliency manipulation did not have a significant effect on behavioral or functional imaging data, we collapsed across saliency levels for all analyses. Two sets of Gabors were created to control for non-numerical factors of the stimuli (http://www.unicog.org/pm/pmwiki.php/Main/Arithmetics). In one set, the overall surface of the Gabor patches was kept constant across numerosities, thus individual item size (varying between 2.65° and 0.93° visual angle for numerosities 1 and 8, respectively) and density (defined as total surface covered by the Gabors divided by convex hull of Gabors) were inversely related to numerosity. In the second set, the individual item size was kept constant (1.26° visual angle), thus total area covered by the Gabors and density increased with increasing numerosity. The average overall surface covered by the Gabor patches across different numerosities was identical for the two sets.

Procedure of main tasks. Two seconds before each trial, a red fixation dot appeared and remained on the screen for 1000 ms to indicate the upcoming trial. Each trial began with the presentation of a gray fixation dot. After a delay of 800 ms, the fixation dot disappeared for 200 ms, signaling the subsequent onset of a stimulus. A number of Gabor patches (sample) appeared on screen for 200 ms (500 ms for the vSTM task), followed by a white fixation cross presented for 500 ms (600 ms for vSTM trials). In the enumeration task, subjects were to name as quickly as possible the number of Gabors. In the vSTM task, a second display (test) was presented that included only one of the previously shown Gabor patches from the sample set (Fig. 1). Subjects were instructed to judge whether the orientation of the test stimulus was changed with respect to the item in the sample set that had been presented in that location. The orientation of the test Gabor was either identical to the sample item or was a mirrored version along the vertical axis. Subjects were required to give their response within a range of 1.7 s. The next trial started after a variable delay (±0–500 ms) with a mean duration of 7400 ms (7000 ms for enumeration) within which the red fixation appeared. The average trial length was 11.2 s for the vSTM task and 10.4 s for the enumeration task. Each experiment was divided into four fMRI runs. Each run lasted ~7.1 min for the vSTM task and 6.9 min for the enumeration task. Figure 1 schematically depicts a trial in the main task.

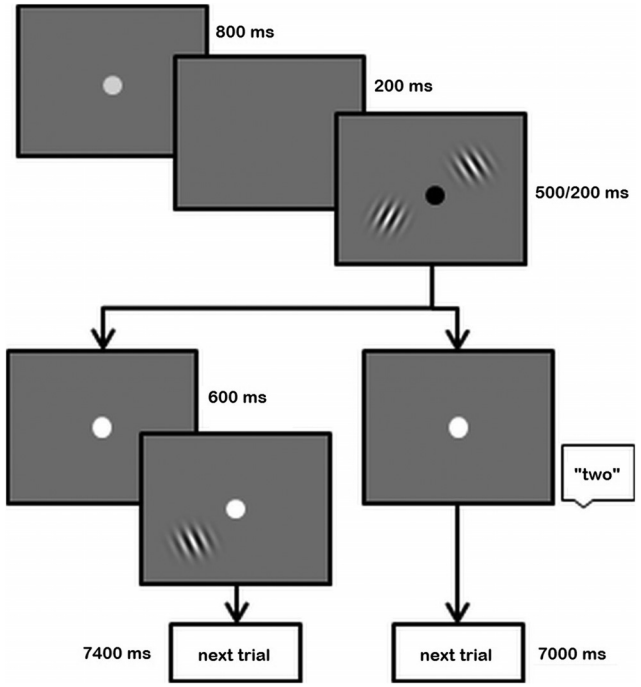

Figure 1. Schematic depiction of a trial in the main experiment. After the initial presentation of a gray fixation cross for 800 ms, followed by a brief blank period, a variable number of Gabors appeared on screen (for 500 and 200 ms in the vSTM and enumeration tasks, respectively). In the vSTM task (left part), a delay period of 600 ms was followed by the presentation of an arbitrarily chosen Gabor that had to be evaluated via button press with respect to a change in orientation (here: orientation changed). In the enumeration task (right part), participants were asked to immediately utter an estimate of the number of Gabors on screen that was recorded and transcoded offline.

Procedure of saccades localizer. In one additional fMRI run, subjects performed 10 blocks of eye movements, each followed by a baseline period in which identical visual stimulation was presented but participants did not move their eyes. The change between the eye movement and the fixation task was signaled via a change in the color of a central fixation cross. Each block of saccades was composed of 14 sequential presentations of a target cross (width and height, 0.38° visual angle) that appeared ~5° (up to ±0.42° jitter in x and y) to the left or the right of fixation or near fixation (with the same jitter) for, on average, 1000 ms (±200 ms jitter; five trials of 800 and 1200 ms, four of 1000 ms). Each block used a different order, and block order was randomized across participants. The total duration of the localizer was 4 min.

Imaging parameters. Functional data in the main experiment were acquired at the Laboratory for Functional Neuroimaging at the Center for Mind/Brain Sciences in Mattarello, Italy on a 4 T MR system (Bruker MedSpec Biospin MR) as T2*-weighted echo-planar image (EPI) volumes using an eight-channel birdcage head coil. Thirty-seven axial slices covering the whole brain were obtained with a TR of 2.2 s (TE, 33 ms; flip angle, 75°; 3 × 3 × 3 mm voxels; no gap). For the saccades localizer task, the TR was 2.4 s. Before each block, we performed an additional scan to measure the point-spread function (PSF) of the acquired sequence, which served for distortion correction. The first three images (6.6 s) in each series served to guarantee stable magnetization and were not recorded. For each participant, an anatomical scan was obtained using a MPRAGE sequence with 176 slices covering the entire brain (TR, 2.7 s; TE, 4.18 ms; flip angle, 7°; voxel size, 1 × 1 × 1 mm; no gap).

Behavioral data analysis. Vocal responses in the enumeration task were recorded and manually labeled offline. Vocal onset times (VOTs) were determined for each trial using an in-house MATLAB algorithm that detected intensity changes above a participant-specific threshold. We then determined the subitizing range (capacity) per participant by fitting a bilinear function to the VOTs and accuracy rates over numerosities. The function identified the best combination of ranges, one with a 0 slope for

small numbers, followed by one with a variable positive slope for larger ones. Subitizing range was operationalized as the intersection of the two lines. As described by Cowan, 2001, vSTM capacity was determined by calculating Cowan's K using the formula $K = (\text{hit rate} + \text{correct rejection rate} - 1)n$, with n describing the number of items in a given set.

Imaging data analysis. After correcting the data for field distortions using the acquired PSF, the functional imaging data were preprocessed using SPM8 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm8). Images were corrected for motion and slice-timing differences, realigned to the first image in the series of the respective experiment, and coregistered to the individual anatomies. For the reported random effects and classifier analyses, the functional images were smoothed with a 6 mm^2 FWHM Gaussian kernel after normalization to the standard template of the Montreal Neurological Institute. The fMRI data from the two main tasks were modeled in a common design to allow for direct comparisons between tasks. The enumeration task was modeled with 16 predictors: 1 predictor for each numerosity (8 levels) × saliency (2 levels). The vSTM was modeled using 12 predictors (6 numerosities × 2 saliencies). All predictors were convolved with a canonical hemodynamic response function and its temporal derivative in SPM8. Session-specific motion parameters were included as effects of no interest to account for remaining artifacts attributable to head motion. We extracted and normalized the β weights for each value of numerosity for each participant in both tasks. Voxels were selected as follows. Based on previous fMRI data (Piazza et al., 2003; Todd and Marois, 2004) and on a computational saliency map model of individuation (Sengupta and Melcher, 2014), we hypothesized that the brain activation profile over numerosities in the enumeration task would differ from the profile in the vSTM task in the PPC regions. In the enumeration task brain, it was expected that activity would follow a "subitizing profile," characterized by a constant level of activation for small numerosities within the subitizing range (1–3) and a linear increases for higher numbers (a "flat and then increase" profile). Conversely, the "VSTM profile" would be characterized by an initial linear increase of brain activity with increasing number of items in the 1–3 range display and a plateau for displays with a larger number of items ("increase and then flat" profile). Any subsequent reference to brain regions that were active in either of the two tasks refers to the activations as defined by these two response profiles. To demonstrate that the same voxels in the PPC flexibly adapt their activation profile to the required representational precision of the task at hand, we analyzed voxels that fulfilled three criteria. First, we isolated superior parietal cortex voxels by an anatomically defined mask using the Wake Forest University (WFU) PickAtlas toolbox in SPM (Maldjian et al., 2003). Second, we used the saccades localizer random-effects contrast (saccade vs fixation) to further restrict our initial PPC anatomical voxel selection. Finally, we chose voxels on the basis of the results from one task (e.g., enumeration) and analyzed their response profile in the other task (e.g., vSTM task).

Additionally, we applied multivariate pattern analysis (MVPA) algorithms to test how far individual numerosities elicit distinguishable patterns of brain activation and how this may differ across tasks. Pattern recognition analysis applied a linear multiclass classification based on support vector machines in the implementation of LIBSVM (Library for Supports Vector Machines; http://www.csie.ntu.edu.tw/~cjlin/libsvm/), with the regularization parameter C fixed to 1. The data entering classification were 240 β images in total (20 images for each of the six numerosity conditions corresponding to individual trials in each of the two tasks). For each of the regions of interest (ROIs) mentioned below, each individual pattern was mean corrected across voxels, and, to reduce potential session- or time-related confounds, voxelwise activations were normalized by subtracting from each voxel the mean across the six numerosity conditions. Because for each session there were 10 trials per condition, this was done 10 times, starting with the first trial of each condition and repeating the same procedure up to the 10th trial for each condition in a given session. Separate classifiers were trained and tested for the enumeration and the vSTM conditions (with the theoretical chance level corresponding to ⅙ for six conditions). The classification cycle used a leave-one-out with 20-fold cross-validation, such that of the 20 patterns of each condition, the n-th pattern ($1 \leq n \leq 20$) from each condition was held out at each given cycle of the cross-validation loop while the classifier was trained on the remaining 19 patterns for each condition. Classification analysis was first applied to an ROI in PPC. To identify voxels in PPC for each participant, we contrasted all numbers versus baseline and masked the resulting first-level SPM with the active voxels from the random-effects contrast saccades versus baseline, with the additional restriction that voxels were located in the parietal cortex as defined in the WFU PickAtlas (Maldjian et al., 2003). To test whether classifier performance was specific to PPC, we identified a second group of voxels in the primary visual cortex (PVC) along the calcarine sulcus using the same functional restrictions. The number of voxels was fixed to the 250 most activated voxels for each participant, task (enumeration and vSTM), and ROI (PPC and PVC).

Control experiment

Participants. Six healthy participants (all females; mean ± SD age, 24.4 ± 0.61 years) were tested in the control experiment that was approved by the Ethics Committee of the Humboldt University of Berlin.

Stimuli. Stimuli consisted of two sets of tilted dark gray bars, displayed against a light gray background. Bars were used instead of Gabor patches to facilitate feature encoding (orientation) and increase performance in the vSTM task (Alvarez and Cavanagh, 2008; Melcher and Piazza, 2011). Two sets were created using a variant of the above described MATLAB routines to control for non-numerical stimulus features in the same way as described above. Individual bar size varied between 0.92° × 0.32° (set size 8) and 2.3° × 0,86° (set size 1) visual angle in set 1 and was fixed to 0.92° × 0.32° in set 2. As in the main experiment, participants were presented with a variable number of items in a vSTM task (one to six items) and an enumeration task (one to eight items).

Procedure and behavioral data analysis. The control study was designed to test whether minor procedural differences between the tasks in the main experiment (e.g., slightly different duration of stimulus presentation) might explain the observed behavioral performance differences and related changes in brain activation profiles (see below). Each trial started with the presentation of a black fixation dot in the center of the screen (500 ms), followed by the simultaneous presentation of a variable number of tilted bars on the screen for 150 ms (sample). In the vSTM task, the offset of the sample stimulus was followed by the presentation of a white fixation dot for 1000 ms (delay period), which was replaced by the presentation of the test stimulus, containing identical number of tilted bars in identical positions. The participants' task was to remember the orientations of the sample bars and to decide whether or not one of the bars in the test display had changed orientation by 90°, which was the case in half of the trials. The test display was replaced by black fixation dots on button press or after 1.7 s. The next trial started on average 3.8 s (minimum, 3.3 s; maximum, 4.4 s) after the response period, yielding a mean trial length of 7.15 s. In enumeration trials, a 1.7 s response period started immediately with the presentation of sample display. Subsequent trials started after an average interval of 7.8 s (minimum, 7.4 s; maximum, 8.3 s). Intertrial fixation was ensured by the presentation of a gray fixation dot in the center of the screen in both tasks.

Vocal responses in the enumeration task were recorded and transcribed offline. Because of technical limitations, no VOT determination was possible and no manual responses were recorded during the vSTM task. Six additional participants (three females; mean ± SD age, 28.3 ± 7.2 years) were tested with the identical paradigm outside the scanner.

Imaging parameters and analysis. Functional data were acquired at the Berlin Center for Advanced Neuroimaging on a 3 T TIM Trio scanner (Siemens) as T2*-weighted EPI volumes using a standard 12-channel head coil. Forty-two axial slices covering the whole brain were obtained with a TR of 2.5 s (TE, 25 ms; flip angle, 82°; 2.5 × 2.5 × 2.5 mm^2 voxels; 20% gap). The first two images (5 s) in each series served to guarantee stable magnetization and were not recorded. For each participant, an anatomical scan was obtained using an MPRAGE sequence with 192 slices covering the entire brain (TR, 1.9 s; TE, 2.52 ms; flip angle, 9°; voxel size, 1 × 1 × 1 mm^2; no gap; generalized autocalibrating partially parallel acquisitions factor, 2).

Functional imaging data were analyzed using the same software (SPM8) and routines as the main experiment. To ensure sampling brain

activation from identical voxels, we used the (resliced) second-level masks from the main experiment.

Computational saliency map model

The saliency map model we used to make quantitative predictions on the activation level of the PPC in the two tasks extended the work by Roggeman et al. (2010), who considered set sizes from 1 to 64 to the set sizes (one to eight) used in this experiment. Following Roggeman et al. (2010) and Sengupta and Melcher (2014), we constructed a recurrent on-center, off-surround network with a single layer of 70 completely interconnected nodes (Fig. 2*A*). Each node can be considered, theoretically, as a group of neurons in the parietal cortex encoding an object or location of an object in an attentional priority or saliency map. The three main parameters that define the type of network are as follows: (1) strength of self-excitation for each node (α); (2) strength of lateral inhibition between nodes (β); and (3) decay constant for the passive decay term (λ). The differential equation governing the time evolution of the network of nodes is given by the following:

$$\frac{dx_i}{dt} = -\lambda x_i + \alpha F(x_i) - \beta \sum_{j=1,\, j\neq i}^{N} F(x_j) + I_i + \text{noise}$$

where $x_i(t)$ is the activation of node at time t, and I_i represents the intensity of external input ($\forall i, 0 \leq I_i \leq 1$). In our simulation is a unit step function, i.e., it has the value 1 for a certain number of time steps for the particular node i and 0 for rest of the time steps. Input is only presented for a finite amount of time, typically much less than total time of simulation. $F(x)$ is the activation function given by the following formula:

$$F(x) = \begin{cases} 0, \text{ for } x \leq 0 \\ \dfrac{x}{1+x}, \text{ for } x > 0 \end{cases}$$

The decay parameter was set to $\lambda = 1$. We modeled the dynamics according to the discrete form of the differential equation governing the time evolution. The activation of the nodes are updated at each step according to the following equation:

$$x_i(t) = \alpha F\left(x_i(t-1)\right) - \beta \sum_{j=1,\, j\neq i}^{N} F\left(x_j(t-1)\right) + I_i + \text{noise}$$

As reported previously by Roggeman et al. (2010), the inhibition parameter determines the degree to which the network behavior can track a small or larger number of items. At high inhibition, the network activation increases with the number of items up to an upper limit that should reflect encoding capacity in the vSTM task. At medium inhibition levels, necessary to individuate but not to track fine local features of the individuated objects, the network activation should show no detectable increase within the very small number range, followed by a steeper increase for larger numerosities, which should reflect higher capacity in enumeration tasks (also referred to as subitizing range). Very low inhibition allows a larger number of nodes in the network to respond in the case of a larger number of inputs, and this might reflect activation related to numerosity estimation, a task that we did not use in our experiments. The behavior of the network can be understood intuitively in terms of competition: when there is strong competition between nodes, strong nodes inhibit other nodes, leading to a winner-take-all system. When the inhibition parameter is weak, the activity of one node does not inhibit its neighbors, allowing many different nodes to be active at the same time. Thus, there is a tradeoff between inhibition and precision. As in the studies by Roggeman et al. (2010) and Sengupta and Melcher, (2014)), the input was presented to the model for five time steps, and the simulations ran for 50 time steps. Then we plotted the mean of 100 simulations of the presentation of one to eight items under a high inhibition parameter ($\beta = 0.28$), simulating the high object feature coding precision requirements of the vSTM task and medium inhibition parameters ($\beta = 0.12$) simulating the lower object feature coding precision requirements of the enumeration task.

Results

Computational model

One important argument in favor of a common, flexible system for both enumeration and vSTM would be the ability to model such flexibility in the same computational model. To test this, we applied a model of visuospatial saliency maps in parietal cortex initially developed by Roggeman et al. (2010). The model is based on a recurrent on-center/off-surround network of connected nodes, with each node representing a spatial location. The nodes are interconnected, with a self-excitation parameter α and a lateral inhibition parameter β (Roggeman et al., 2010; Sengupta and Melcher, 2014). As shown previously, the behavior of this saliency map model depends critically on the inhibition between nodes. In a high inhibition regimen, we found that activation increased as set size went from one to three items and then reached a plateau (Todd and Marois, 2004; Kawasaki et al., 2008; Fig. 3*I*). However, for a medium level of inhibition, we found flat activation up to approximately three items (Fig. 3*G*) as would be expected for subitizing and in line with previous results (Piazza et al., 2003). Thus, changing only the inhibition parameter leads to changing response profiles as a function of the representational precision required by the task at hand. Based on these results, we predict that the activation profiles in the two tasks should reflect the varying representational precision in brain areas organized in a map-like architecture, thus resembling the activation profiles of the computational model.

Behavioral results

Performance in the enumeration task matched the expected response profiles. For numerosities 1–3, the verbal estimates were highly accurate and did not vary in speed (Fig. 2*B*,*D* for VOTs and accuracy, respectively). Beyond numerosity 3, latency increased and accuracy decreased for larger numerosity values as expected. To determine the subitizing range, we fitted the data (VOTs and error rates) using a bilinear fit algorithm. Using pairwise *t* tests, we found VOTs attaining a plateau at seven items (six vs seven items, $t_{(16)} = -3.39$, $p = 0.004$; seven vs eight items, $t_{(16)} = -0.054$, $p = 0.958$). To avoid artificially reduced estimates of subitizing range, we included only numerosities 1–7 in the VOT analysis. This analysis revealed a subitizing range of approximately three items for VOTs (mean ± SD, 2.71 ± 0.75) and of approximately four items for error rates (mean ± SD, 4.0 ± 0.98).

To confirm the presence of a subitizing range in the enumeration task, we used a one-way repeated-measures ANOVA with numerosity as the only factor to test for (1) the absence of an impact of numerosity for small numerosities and (2) an impact of numerosity for larger numerosities. We observed a marginal impact of numerosity for enumeration performance in the one to three range for VOTs ($F_{(2,32)} = 2.95$, $p = 0.07$), which was absent for error rates ($F_{(2,32)} = 0.93$, $p = 0.39$). For numerosities 1–4, we observed an impact of numerosity on VOTs [$F_{(3,48)} = 18.943$, $p < 0.001$, Greenhouse-Geisser $\varepsilon = 0.61$ (Greenhouse and Geisser, 1959)] that was marginal only in accuracy ($F_{(3,48)} = 3.3$, $p = 0.053$, $\varepsilon = 0.63$). In contrast, a marked effect of numerosity was present for larger numerosities between 4 and 8, for both VOTs ($F_{(3,48)} = 16.340$, $p < 0.001$, $\varepsilon = 0.6$) and error rates ($F_{(3,48)} = 20.203$, $p < 0.001$, $\varepsilon = 0.9$). Together, these results suggest an average subitizing range of three to four items.

Accuracy in the vSTM task significantly decreased with increasing numerosity over the whole range (Fig. 2*E*), with an average ± SD Cowan's *K* of $K = 1.2 \pm 0.53$ (maximum $K = 1.38$; Fig. 2*C*). The capacity estimate of ~1.5 items for orientation memory is consistent with previous studies using an orientation

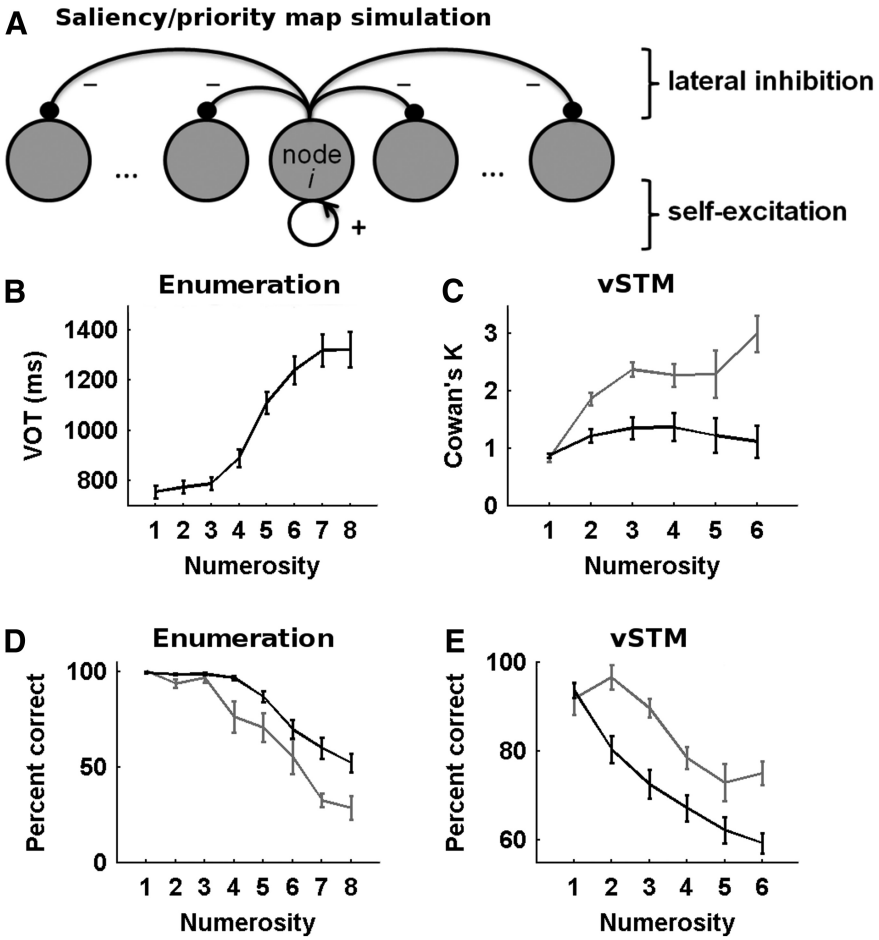

Figure 2. ***A***, Graphical illustration of the computational saliency/priority model. Self-excitation and inhibitory connection between nodes are indicated by the different endpoints (arrows and filled circles, respectively). ***B–E***, Behavioral results from the main (black) and control (gray) experiments. Average voice onset times (***B***) and percentage correct (***D***) across numerosities in the enumeration task and Cowan's *K* (***C***) and percentage correct (***E***) in the vSTM task.

task on Gabors (Alvarez and Cavanagh, 2008; Melcher and Piazza, 2011).

Behavioral performance in the control experiment was comparable with the main experiment. Enumeration accuracy was characterized by a near-perfect performance for numerosities one through three and monotonically decreasing accuracy with additional increases of set size (Fig. 2*D*,*E*). To visualize vSTM performance, we calculated Cowan's *K* (Fig. 2*C*). *K* (mean $K = 2.11$; maximum $K = 3.0$) increased with increasing set sizes until reaching a plateau.

Brain imaging results

GLM results

As a starting point, we traced brain regions that responded according to the different expected response profiles across numerosities in the enumeration and vSTM tasks. For the enumeration task, we were looking for voxels with a response profile that would parallel the behavioral results, that is, voxels that did not exhibit an increase of activation for low numerosities ($n \leq 3$) but a parametric increase in activation for higher numerosities ($n > 3$), equivalent to an exponential function. For the vSTM task, we traced voxels that showed a complementary response profile with an increase of activation for lower numerosities ($n \leq 3$), reaching a plateau for higher numerosities ($n \geq 3$), equivalent to the inverse of an exponential function. Figure 3 shows the resulting activated networks projected onto an inflated brain template using the Human PALS (population-average landmark- and surface-based)-B12 Atlas (Van Essen, 2005; Van Essen and Dierker, 2007) implemented in Caret software (Van Essen et al., 2001). vSTM (Fig. 3*A*) activated bilateral precentral regions (frontal eye fields), superior parietal cortex and occipital cortex. Figure 3*B* shows the activations elicited by the saccades localizer task, consisting mainly of superior parietal and occipital regions. Enumeration (Fig. 3*C*) activated a large network of frontal, precentral, and parietal regions extending into the occipital cortex (for a detailed list of activated sites, see Table 1). Virtually identical brain regions were obtained when using regressor profiles with lower inflection points of 3 and 2, better matching the empirically observed profiles for enumeration and vSTM, respectively.

Given our interest in the activity of saliency/priority maps, we sought to sample from the human homolog of monkey area LIP by adopting an inclusive masking approach that only included voxels that were (1) anatomically located in the parietal cortex and (2) were active in the saccades localizer task. From these voxels, we selected only those voxels that were active in either the enumeration task or the vSTM task with their differential response profiles as described above. Overlapping voxels between vSTM and saccades, and enumeration and saccades are shown in Figure 3, *D* and *E*, respectively. For voxels that exhibited an STM profile in the vSTM task, we then plotted the activation profile in the enumeration task (Fig. 3*F*). Conversely, for voxels that showed a subitizing profile, we plotted the activation profile in the vSTM task (Fig. 3*H*). The response profile changed completely as a function of the specific task at hand. Voxels that paralleled the behavioral profile in the enumeration task changed their profile in the context of the vSTM task and vice versa. In both cases, the β values varied significantly with numerosity as indicated by one-way repeated-measures ANOVA (vSTM, $F_{(5,80)} = 11.1$, $p < 0.001$, $\varepsilon = 0.70$; enumeration, $F_{(7,112)} = 31.73$, $p < 0.001$, $\varepsilon = 0.72$).

To statistically validate that voxels in PPC changed their response profile with task requirements, we analyzed the data points that were common to both tasks (i.e., numerosities 1–6) by fitting a log-linear function (Anobile et al., 2012) to the β weights in the context of both tasks. The function of the form $Y = a\left((1 - \lambda)N + \lambda \frac{N_{\max}}{\ln N_{\max}} \ln N\right)$ comprises a linear and a logarithmic component (λ) that will be positive if the response profile is (logarithmically) compressed (as should be the case in the vSTM task) and negative if the response profile takes an exponential form (as should be the case in the enumeration task for numerosities 1–6). We fitted this function to the individual mean β weights of the voxels that were identified with the above de-

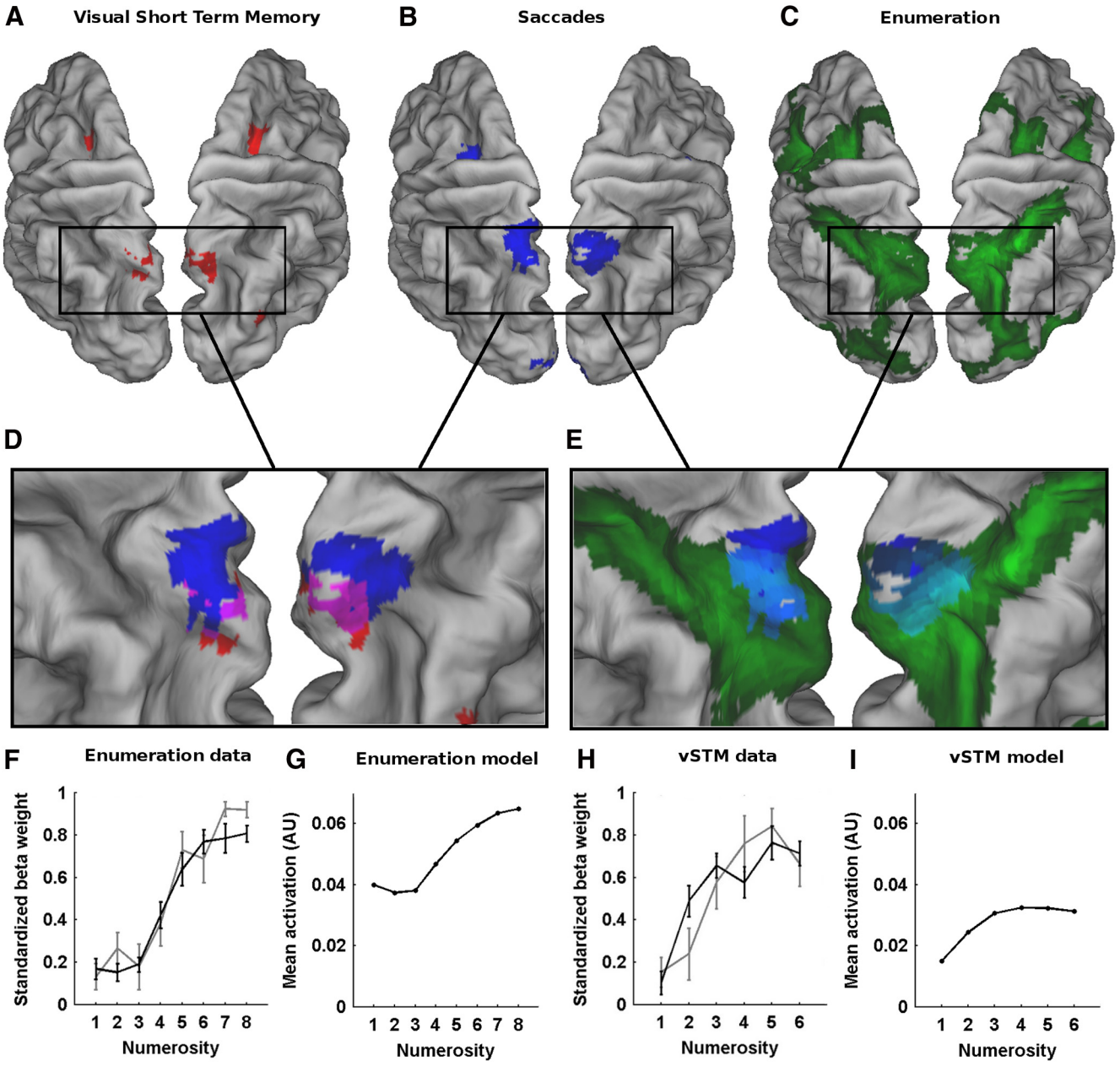

Figure 3. Brain activation results. ***A***, Brain regions exhibiting a vSTM profile. ***B***, Results of the saccades localizer. ***C***, Brain regions with a subitizing profile. All random-effects contrasts projected onto the top view of left and right hemispheres of an inflated brain template, thresholded at $p < 0.05$ (FDR-corrected) except saccades localizer ($p = 0.005$, uncorrected). ***D***, ***E***, Enlarged view of the overlapping activation in PPC. ***D***, Overlap (purple) between saccades localizer (blue) and vSTM (red) activation. ***E***, Overlap (turquoise) between saccades localizer (blue) and enumeration (green) activation. ***F–I***, Empirically observed and computational model activation profiles in the enumeration task (***F***, ***G***, respectively) and the vSTM task (***H***, ***I***, respectively) expressed as standardized β weights (data) and arbitrary units (model). Empirical activation profile for enumeration is based on voxels that have been identified by the overlap between vSTM and saccades. The profile for vSTM is based on voxels that have been identified by the overlap between enumeration and saccades. Results from the main and control experiments are shown in black and gray, respectively. Error bars depict SEM.

scribed approach in both tasks ($n = 35$ voxels from the vSTM task and $n = 51$ voxels from the enumeration task) and compared the λ parameters obtained for both tasks. Model fit was good for both response profiles ($R^2 = 0.73$ and $R^2 = 0.93$ for the vSTM and enumeration tasks, respectively). Analysis of the λ parameter revealed that parameters were significantly different from 0 ($\lambda = 1.05$, $t_{(16)} = 3.50$, $p = 0.003$; $\lambda = -0.88$, $t_{(16)} = -3.24$, $p = 0.005$ for vSTM and enumeration tasks, respectively) and were significantly different from each other ($t_{(16)} = 5.29$, $p < 0.0001$).

Very similar results, including a reversal of response profiles as a function of task requirements, were obtained from the following: (1) voxels in the parietal cortex without masking the task-related activation maps with the saccades map; and (2) voxels within the saccades map without applying the masks from the respective activation maps. This suggests that the observed pattern of results is not confined to the voxels from the mask conjunction but that other regions in PPC also exhibit a similar response behavior.

To test whether minor procedural differences between the enumeration and vSTM tasks rather than the task itself could account for the observed differences in the neural response profiles, we conducted a control experiment. In particular, in

Table 1. Brain areas from the enumeration task, the vSTM task, and saccades localizer

Task	Peak coordinates (MNI)			Peak Z-score	Cluster size	Label
	x	y	z			
Enumeration	0	20	46	6.85	289	Superior frontal lobe
	3	29	40	6.41		R superior medial frontal lobe
	0	5	67	5.16		Supplementary motor area
	−30	23	4	5.92	153	R inferior frontal cortex/insula
	−24	−58	43	5.89	229	L superior parietal lobe
	−12	−64	55	5.47		L precuneus
	−36	−49	40	5.44		L inferior parietal lobe
	33	23	4	5.89	126	R insula
	−45	8	34	5.85	99	L precentral gyrus
	−48	−4	43	5.17		L precentral gyrus
	−42	20	28	4.83		L inferior frontal lobe
	36	−49	49	5.71	294	R inferior parietal lobe
	27	−58	49	5.50		R superior parietal lobe
	15	−67	55	5.41		R superior parietal lobe
	45	8	31	5.54	51	R precentral gyrus
	−30	2	55	5.48	68	L middle frontal lobe
	−15	5	55	4.85		L superior frontal lobe
	−27	2	43	4.78		L precentral gyrus
vSTM	12	−58	58	4.40	157	R precuneus
	21	5	58	4.35	78	R superior parietal lobe
	−15	−64	55	4.18	95	L superior parietal lobe
	−3	23	40	4.05	161	L superior medial frontal lobe
	−12	8	58	3.89		L superior frontal lobe
	−21	−1	52	3.77		L superior frontal lobe
	45	−76	22	3.85	19	R middle occipital lobe
	−33	−73	22	3.56	26	L middle occipital lobe
Saccades localizer	−15	−64	52	4.02	159	R Superior parietal lobe
	15	−91	−2	3.76	593	R calcarine sulcus
	−12	−94	4	3.68		L superior occipital sulcus
	21	−79	10	3.13		R calcarine sulcus
	24	−70	61	3.69	147	R superior parietal lobe
	15	−67	64	3.47		R superior parietal lobe
	−36	−58	−29	3.27	24	L cerebellum
	−27	−10	49	3.26	99	L middle frontal lobe
	−30	−1	64	2.82		L middle frontal lobe
	−6	−76	−17	3.24	35	L cerebellum
	9	−73	−20	2.75		R cerebellum
	36	−55	−29	3.12	25	R cerebellum
	48	2	55	2.88	12	R middle frontal lobe
	42	−1	61	2.8		R middle frontal lobe
	−24	5	7	2.71	12	L putamen

For a detailed contrast description, see Materials and Methods. Because activation clusters in the enumeration task were linked to each other with one cluster containing 15,384 voxels, we report the clusters at $p = 0.05$ (FWE-corrected) for the enumeration task. L, Left; R, right.

the main experiment, there were slightly longer stimulus presentation times during the vSTM task, which in theory might have influenced the results such as by allowing for more saccadic eye movements during the vSTM task compared with enumeration.

As shown in Figure 3, *F* and *H* (depicted in gray), despite now using identical stimulus displays and identical stimulus presentation parameters in the two tasks, we again found that brain activation profiles in PPC showed remarkable differences between tasks. In line with the results from the main experiment, we found a stable activation level for numerosities 1–4 in the enumeration task and an increase of activation with increasing set size (more than three). In contrast, in the vSTM task, activation increased from one to approximately four items before reaching a plateau and decreasing again for six items. It is important to note that the pattern of results from the main experiment was replicated despite independent samples, different MR systems, imaging parameters, and stimuli, and this underlines the idea of flexible coding accuracy as the driving factor in the current findings. Thus, it is unlikely that marginal procedural task differences (e.g., stimulus duration) in the main experiment played an important role in the observed differences in response profiles between the two tasks.

Decoding results

From single-cell recordings in monkeys (Roitman et al., 2007), we know that area LIP contains neurons that code for numbers. Previous fMRI multivariate decoding experiments showed that individual numbers can be discriminated in the human brain using activation patterns in the PPC (Eger et al., 2009) and in the functional equivalent of area LIP (Eger et al., 2013). In the present context, we used decoding to obtain an additional neuronal index of differential capacity limitation in the two tasks: we asked whether the number of items that can be accurately discriminated varied as a function of task demands. The confusion matrices in Figure 4 depict classification results from multiclass classification between the six numbers for both tasks and ROIs. A repeated-measures ANOVA of Cohen's κ with the factors ROI (PPC vs PVC) and task (enumeration vs vSTM) revealed higher classification performance for enumeration ($F_{(1,16)} = 42.16$, $p < 0.0001$) and in PPC ($F_{(1,16)} = 158.97$, $p < 0.0001$). No interaction was observed ($F_{(1,16)} = 0.19$, $p = 0.666$). This result was corroborated by significantly better classification accuracy (percentage correct) for enumeration ($F_{(1,16)} = 74.14$, $p < 0.0001$) and in PPC ($F_{(1,16)} = 10.997$, $p = 0.0044$). Individual classification accuracy was higher than chance (16.7%) in each ROI and task (PPC enumeration, 34%, $t_{(16)} = 10.96$, $p < 0.0001$; PPC vSTM, 30.3%, $t_{(16)} = 9.24$, $p < 0.0001$; PVC enumeration, 25.4%, $t_{(16)} = 9.1$, $p < 0.0001$; PVC vSTM, 21.1%, $t_{(16)} = 2.9$, $p = 0.0104$).

Paralleling the GLM results, classification performance for different numerosities was modulated by task, resulting in different classification profiles. For enumeration, classification was on average much higher than for vSTM, being best for values 1–2 and 5–6 and reaching lowest values for classification of three and four items. However, in vSTM, classification peaked at low numerosities and decreased with increasing numerosity, resulting in a broader confusion range for larger numerosities compared with enumeration. To further corroborate these impressions, we computed classification "tuning curves" for each task and numerosity in both ROIs. That is, for each predicted numerosity (e.g., one item), we computed the difference between the correct classification (i.e., the value on the diagonal in the confusion matrix) and the mean of the false classifications (i.e., values off the diagonal in the confusion matrix). We compared these classification profiles in a repeated-measures ANOVA with the factors task (enumeration and vSTM) and numerosity (one through six) for each ROI. In PPC, we observed a main effect of numerosity ($F_{(5,80)} = 37.54$, $p = 0.0001$, $\varepsilon = 0.73$). Task marginally influenced performance

($F_{(1,16)} = 3.62$, $p = 0.08$). Most importantly, an interaction between numerosity and task ($F_{(5,80)} = 16.49$, $p < 0.00001$, $\varepsilon = 0.8$) indicated that the confusion profile across numerosities was significantly modulated by task. For the early visual area, main effects of task and numerosity were significant ($F_{(1,16)} = 7.49$, $p = 0.0146$; $F_{(5,80)} = 2.57$, $p = 0.0443$, $\varepsilon = 0.83$) but did not significantly interact with each other ($F_{(5,80)} = 2.23$, $p = 0.0838$, $\varepsilon = 0.72$). In summary, these analyses based on decoding directly show that the number of items that can be accurately represented is lower in the vSTM task than the working memory task, confirming the prediction related to differential capacity limitations.

Finally, the notion of shared neural map architecture underlying enumeration and vSTM may suggest that numerosity coding should at least partially generalize across tasks. To test this idea, we trained a classifier to discriminate numerosity information in one task (e.g., enumeration) and tested it with numerosities from the other (e.g., vSTM). Overall classification rates were lower than in within-task classification ($t_{(16)} = -8.79$, $p < 0.0001$; Fig. 4*C*), and they were higher in PPC than in early visual area ($F_{(1,16)} = 19.35$, $p = 0.0005$), whereas no task difference or interaction between task and ROI was observed ($F_{(1,16)} = 0.009$, $p = 0.925$; $F_{(1,16)} = 0.124$, $p = 0.7292$). Classification rates were significantly better than chance in PPC for both vSTM to enumeration generalization (22.2%, $t_{(16)} = 4.928$, $p = 0.0002$) and enumeration to vSTM generalization (22.0%, $t_{(16)} = 3.76$, $p = 0.0017$). However, classification rates in early visual areas were at chance level (enumeration to vSTM, 17.8%, $t_{(16)} = 1.38$, $p = 0.187$; vSTM to enumeration, 17.5%, $t_{(16)} = 1.246$, $p = 0.2305$).

These results may lead to the conclusion that number is automatically extracted in this region independently of the tasks that subjects are performing. However, a closer exploration of the generalization confusion matrices suggested that this is not the case because we observe an asymmetric pattern of confusions between numerosities across tasks. In Figure 4*C*, classification errors in the top right part of the confusion matrix seemed to be more prominent compared with the bottom left part of the matrix. This would imply that a classifier that was trained to infer numerosities from PPC activation while subjects were performing the vSTM task would tend to "underestimate" numerosities when fed with PPC activation while subjects perform the enumeration task. This is compatible with the idea that numerosity might be encoded as the number of peaks in a saliency map and that, for any given number of objects during enumeration, there are systematically more active peaks than during vSTM. To test this hypothesis, we directly compared classification errors in

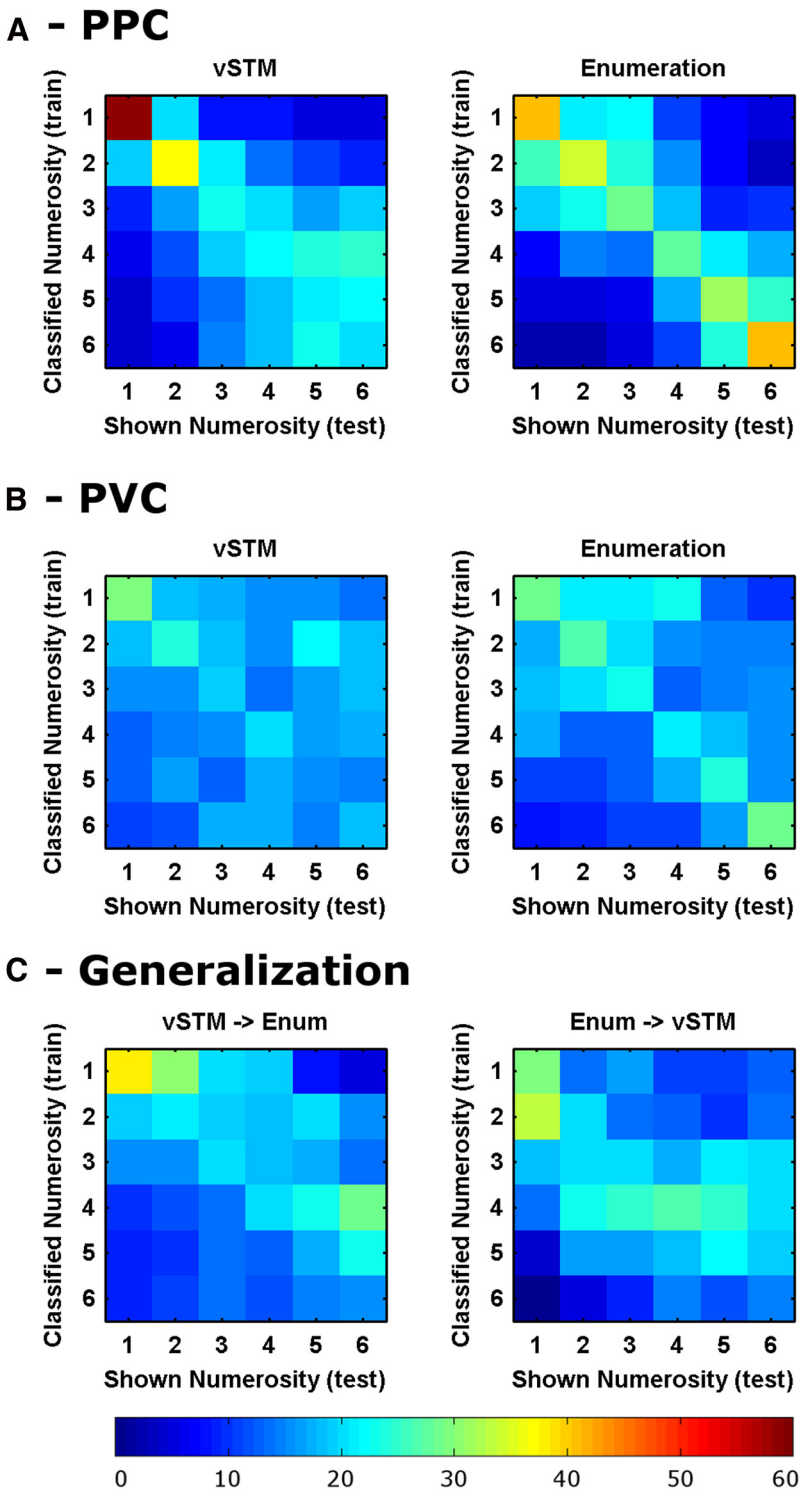

Figure 4. Classification results. Confusion matrices displaying the percentages of trials in which patterns from each of the six given numerosities were classified as the same or each one of the other numerosities. Values along the diagonal correspond to correct classifications, and off-diagonal values correspond to misclassifications. ***A***, Confusion matrices of the multiclass classification algorithm in PPC in the vSTM task (left) and the enumeration task (right). ***B***, Confusion matrices of the multiclass classification algorithm in PVC in the vSTM task (left) and the enumeration task (right). ***C***, Confusion matrices in PPC when the multiclass classification algorithm generalized from one task to the other. For example, on the left, the classifier was trained using numerosities from the vSTM task and tested with numerosities from the enumeration task (vSTM → Enum).

the top right to errors in the bottom left, separately for each ROI. Because generalization in early visual area was not above chance, we restricted this analysis to PPC. False classification rates in the vSTM to enumeration generalization were systematically biased ($F_{(1,16)} = 42.96$, $p < 0.0001$) and bias was modulated by direction of generalization, as indicated by the interaction ($F_{(1,16)} = 20.51$, $p = 0.0003$). We observed significantly higher false classification rates in the top right part of the confusion matrix compared with the bottom left part when generalizing from vSTM to enumeration ($t_{(16)} = 6.45$, $p < 0.0001$) but not when generalizing from enumeration to vSTM ($t_{(16)} = 0.81$, $p = 0.429$).

Discussion

Going beyond previous, separate studies showing different patterns of PPC activation for enumeration and vSTM tasks in different subjects, here we demonstrate that the same voxels in PPC are involved in both tasks. However, their response profiles were flexible and task dependent. For example, voxels defined by their flat activity profile up to approximately three items during enumeration exhibit a different response profile during vSTM, closely tracking vSTM capacity.

The current findings provide direct evidence that object enumeration and vSTM, previously studied separately, share a crucial neural mechanism that reflects modulations of their capacity limits. Starting from a map architecture in which representational precision could be varied according to task demands (Roggeman et al., 2010; Sengupta and Melcher, 2014), we predicted task-specific activation profiles reflecting previously reported performance patterns. In such a flexible saliency/priority map, a small number of items can be presented with high precision with minimal noise to allow for rich encoding of stimulus features, such as orientation and spatial position, that were required in our vSTM context. With lower precision, more items would be represented at the cost of lower feature resolution, albeit sufficient for mere enumeration of items in a given set (Melcher and Piazza, 2011). Such map architecture has a number of advantages. First, it provides a way to account for both evidence of discrete representations and also the fact that capacity limits change across context and task (Melcher and Piazza, 2011; Franconeri et al., 2013), providing a way forward from debates about slots versus resources (Franconeri et al., 2013). Second, these maps are biologically plausible models of the well studied behavior of neurons in PPC areas, such as the lateral intraparietal sulcus (Gottlieb, 2007).

Although a saliency/priority map in PPC that represents multiple objects with a variable degree of precision would parsimoniously account for the observed pattern of results, we cannot exclude the contribution of other regions to the observed capacity limits of visual perception. For example, in the case of vSTM, there are additional processes involved beyond individuating items (for review, see Melcher and Piazza, 2011; Wutz and Melcher, 2013). However, at the very least, our results indicate that PPC is part of a neural network that reflects capacity-limited information coding. Moreover, we show how a saliency/priority map model can account for previous results showing differential variation in activation profiles as a function of number of items for both enumeration and vSTM tasks by varying only a single parameter: inhibition between nodes.

To further validate the predictions of the saliency/priority map hypothesis, we used MVPA over the PPC regions to decode cortical activity patterns associated with the different set sizes in the two tasks. Distributed activity in PPC was differentially modulated by set size across the two tasks: numerosity was accurately decoded from PPC activation in the two tasks, but the overall decoding accuracy of numerosity was higher for the enumeration task than the vSTM task. Indeed, the tuning precision of the decoder was high across all numerosities in the enumeration regimen, whereas it was high only for numerosities 1 and 2 in the vSTM regimen. These data are in line with the idea that a common saliency/priority map in PPC codes for multiple objects with a different degree of precision in tasks differing by the amount of individual object resolution required. In particular, the low decoding accuracy for larger numerosities in the vSTM task supports the idea that activity of the saliency map is concentrated on a limited number of items, and, because of a lack of free resources, item information of additional items in a given set is lost when set size exceeds the given task-specific capacity. Additionally, beyond what was shown in the univariate analyses, we demonstrated that, in the enumeration regime, PPC was sensitive to number throughout the entire tested range of numerosities, including numerosities 1–3, which were impossible to differentiate based on the activation level in the univariate analysis. This finding speaks to a previously open question of whether small numerosities are coded using the same parietal cortex mechanisms as the ones involved in the coding of high numerosities or whether there is a separate neural system underlying subitizing (Vetter et al., 2011; He et al., 2013). Although overall activation was constant across numerosities in the 1–3 range, the information distributed over several PPC voxels was sufficient to discriminate number, in line with the existence of number neurons in the PPC of the macaque monkey specifically tuned to a broad range of numerosities (Roitman et al., 2007). The multivariate analysis also allowed us to directly compare the representations of items in PPC with those of PVC. Although some information related to numerosity was also present in PVC, numerosity coding was more reliable [enumeration, 34% vs 25% decoding accuracy in PPC and PVC, respectively (chance = 16.7%) vs vSTM, 30% vs 21%] and more precise (i.e., smaller width of the tuning profile in PPC vs PVC in both enumeration and vSTM) in PPC. These results are consistent with the idea that PPC is specifically sensitive to the individuation of specific items rather than just the total amount of visual stimulation. Consequently, the task-specific response profiles dismiss the idea that PPC activation merely reflects automatic extraction of numerical information regardless of task requirements.

Finally, the multivariate analysis allowed us to more directly compare the information in PPC and PVC during the two tasks. We found that numerosity information in the PPC (but not in PVC) generalized across tasks: here the decoder was trained on the data from the vSTM task and tested on the data from the enumeration task (and vice versa). Interestingly, however, training the decoder in a high inhibition, vSTM regime led to an underestimation of numerosity as encoded in PPC during the low inhibition, enumeration regime. For example, a PPC pattern evoked by six elements in the low inhibition task was more similar to a pattern evoked by four elements than to a pattern evoked by six elements in the high inhibition task. This finding provides additional support to the idea of a flexible representation system in which the same voxels change their response profile as a function of task.

Because of the very nature of our vSTM task and our neural measure (the BOLD signal extends over many seconds) in the present study, we cannot distinguish whether PPC activation is related to the perceptual encoding versus the memory maintenance stage of vSTM. Previous studies on brain activity during similar vSTM tasks have shown that the BOLD signal is similarly modulated by set size both at encoding and during memory

maintenance (Todd and Marois, 2004). Hence, it is reasonable to assume that both encoding and maintenance rely on activity in neural circuits in PPC that are organized in a map-based architecture. Maintaining the positions of the items, weighted for their behavioral and sensory relevance may indeed be one of the key functions of the map architecture in PPC.

The differential activation patterns reported for the enumeration and the vSTM task conditions may be seen as reflecting "task difficulty." Indeed, both the pattern of error rates and response times indicate that increasing set size differentially modulates the difficulty of the two tasks (it is equally easy to enumerate sets of one to three items, whereas the difficulty for encoding/maintaining the features of the objects increases with increasing set sizes from one to three). According to our model, this is attributable to the differential amount of lateral inhibition of the saliency map that is set by the task demands and that occurs right at the encoding stage in both tasks. Therefore, in this respect, our model precisely explains why set size differently modulates difficulty in the two tasks. Although we cannot completely exclude that response time may have affected the PPC BOLD activation amplitude, especially for the enumeration task, our idea of task-specific flexible representational precision in a map architecture is confirmed by MVPA, for which the BOLD amplitude is discarded. Nevertheless, it remains an interesting question for future research to investigate whether or not PPC activity for different set sizes would be similar to the currently observed ones after equating both the maintenance component and the response time components across tasks.

Conclusions

Overall, the current results suggest that previous reports of neural activity in parietal cortex during enumeration and vSTM tasks reflect a common, flexible system to represent multiple individual objects. This flexibility can be accounted for by a map-like architecture. Indeed, such a model is biologically plausible, given previous studies of PPC and would help to reconcile findings showing both discrete representations and variations in capacity and resolution of representations in different tasks across multiple experiments.

References

Alvarez GA, Cavanagh P (2008) Visual short-term memory operates more efficiently on boundary features than on surface features. Percept Psychophys 70:346–364. CrossRef Medline

Anobile G, Cicchini GM, Burr DC (2012) Linear mapping of numbers onto space requires attention. Cognition 122:454–459. CrossRef Medline

Bays PM, Singh-Curry V, Gorgoraptis N, Driver J, Husain M (2010) Integration of goal- and stimulus-related visual signals revealed by damage to human parietal cortex. J Neurosci 30:5968–5978. CrossRef Medline

Bisley JW, Goldberg ME (2003) Neuronal activity in the lateral intraparietal area and spatial attention. Science 299:81–86. CrossRef Medline

Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. Annu Rev Neurosci 22:319–349. CrossRef Medline

Connolly JD, Goodale MA, Menon RS, Munoz DP (2002) Human fMRI evidence for the neural correlates of preparatory set. Nat Neurosci 5:1345–1352. CrossRef Medline

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24:87–114; discussion 114–185. CrossRef Medline

Dempere-Marco L, Melcher DP, Deco G (2012) Effective visual working memory capacity: an emergent effect from the neural dynamics in an attractor network. PLoS One 7:e42719. CrossRef Medline

Eger E, Michel V, Thirion B, Amadon A, Dehaene S, Kleinschmidt A (2009) Deciphering cortical number coding from human brain activity patterns. Curr Biol 19:1608–1615. CrossRef Medline

Eger E, Pinel P, Dehaene S, Kleinschmidt A (2013) Spatially invariant coding of numerical information in functionally defined subregions of human parietal cortex. Cereb Cortex. Advance online publication. Retrieved June 8, 2014. doi:10.1093/cercor/bht323. CrossRef Medline

Franconeri SL, Alvarez GA, Cavanagh P (2013) Flexible cognitive resources: competitive content maps for attention and memory. Trends Cogn Sci 17:134–141. CrossRef Medline

Gottlieb J (2007) From thought to action: the parietal cortex as a bridge between perception, action, and cognition. Neuron 53:9–16. CrossRef Medline

Greenhouse SW, Geisser S (1959) On methods in the analysis of profile data. Psychometrika 24:95–112.

He L, Zuo Z, Chen L, Humphreys G (2013) Effects of number magnitude and notation at 7T: separating the neural response to small and large, symbolic and nonsymbolic number. Cereb Cortex. Advance online publication. Retrieved June 8, 2014. doi:10.1093/cercor/bht074. CrossRef Medline

Kawasaki M, Watanabe M, Okuda J, Sakagami M, Aihara K (2008) Human posterior parietal cortex maintains color, shape and motion in visual short-term memory. Brain Res 1213:91–97. CrossRef Medline

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19:1233–1239. CrossRef Medline

Melcher D, Piazza M (2011) The role of attentional priority and saliency in determining capacity limits in enumeration and visual working memory. PLoS One 6:e29296. CrossRef Medline

Piazza M, Mechelli A, Butterworth B, Price CJ (2002) Are subitizing and counting implemented as separate or functionally overlapping processes? Neuroimage 15:435–446. CrossRef Medline

Piazza M, Giacomini E, Le Bihan D, Dehaene S (2003) Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. Proc Biol Sci 270:1237–1245. CrossRef Medline

Piazza M, Fumarola A, Chinello A, Melcher D (2011) Subitizing reflects visuo-spatial object individuation capacity. Cognition 121:147–153. CrossRef Medline

Roggeman C, Fias W, Verguts T (2010) Salience maps in parietal cortex: imaging and computational modeling. Neuroimage 52:1005–1014. CrossRef Medline

Roitman JD, Brannon EM, Platt ML (2007) Monotonic coding of numerosity in macaque lateral intraparietal area. PLoS Biol 5:e208. CrossRef Medline

Sengupta R, Surampudi BR, Melcher D (2014) A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. Brain Res, in press. CrossRef

Silk TJ, Bellgrove MA, Wrafter P, Mattingley JB, Cunnington R (2010) Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. Neuroimage 53:718–724. CrossRef Medline

Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428:751–754. CrossRef Medline

Van Essen DC (2005) A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. Neuroimage 28:635–662. CrossRef Medline

Van Essen DC, Dierker DL (2007) Surface-based and probabilistic atlases of primate cerebral cortex. Neuron 56:209–225. CrossRef Medline

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001) An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assn 8:443–459. CrossRef Medline

Vetter P, Butterworth B, Bahrami B (2011) A candidate for the attentional bottleneck: set-size specific modulation of the right TPJ during attentive enumeration. J Cogn Neurosci 23:728–736. CrossRef Medline

Wutz A, Melcher D (2013) Temporal buffering and visual capacity: the time course of object formation underlies capacity limits in visual cognition. Atten Percep Psychophys 75:921–933. CrossRef Medline

Recruitment of an Area Involved in Eye Movements During Mental Arithmetic

André Knops,[1,2,3]* Bertrand Thirion,[2,4] Edward M. Hubbard,[1,2,3] Vincent Michel,[2,3,4] Stanislas Dehaene[1,2,3,5]

Throughout the history of mathematics, concepts of number and space have been tightly intertwined. We tested the hypothesis that cortical circuits for spatial attention contribute to mental arithmetic in humans. We trained a multivariate classifier algorithm to infer the direction of an eye movement, left or right, from the brain activation measured in the posterior parietal cortex. Without further training, the classifier then generalized to an arithmetic task. Its left versus right classification could be used to sort out subtraction versus addition trials, whether performed with symbols or with sets of dots. These findings are consistent with the suggestion that mental arithmetic co-opts parietal circuitry associated with spatial coding.

The human species is unique in its capacity to create revolutionary cultural inventions such as writing and mathematics, which dramatically enhance its native competence. From a neurobiological standpoint, such inventions are too recent for natural selection to have dedicated specific brain mechanisms to them. It has therefore been suggested that they co-opt or "recycle" evolutionarily older circuits with a related function (*1*), thus enriching (without necessarily replacing) their domain of use. For instance, learning to read recruits a left inferotemporal area originally engaged in object recognition, and even the seemingly arbitrary shapes of our letters may originate in a neural repertoire of junction detectors that are useful for scene recognition and available to all primates (*2*). In the case of mathematics, although foundational intuitions such as number sense (*3*) and spatial maps (*4*) are present in many animal species and in humans before their education, mathematical constructions vastly exceed these initial domains of inherited competence. It has been argued that analogies between number and space play a crucial role in the expansion of mathematical concepts (*5*). We investigated the role of brain areas for spatial coding in mental arithmetic.

Many behavioral experiments have demonstrated automatic links between number and space. Even young children and uneducated adults readily conceive of numbers as forming an internal spatial continuum or "mental number line" (*6*). Merely perceiving an Arabic digit suffices to elicit a spatial bias in both attentional orienting (*7*) and manual responses (*8*), with small numbers inducing a left-sided and large numbers a right-sided advantage in left-to-right readers. When adults perform approximate additions and subtractions, they overshoot toward larger numbers for addition and toward smaller numbers for subtraction, as if carried along by spatial momentum (*9*). Perhaps the most conclusive evidence for numerical-spatial links comes from the syndrome of spatial hemineglect, in which brain-lesioned patients fail to attend to one side of space, usually the left side. When such patients attempt to bisect a numerical interval, their responses are shifted toward larger numbers, as if neglecting the left half of the numerical segment, where small numbers are represented (*10*).

The brain mechanisms of these numerical-spatial interactions, however, remain largely unknown. In both monkeys and humans, number processing recruits a brain area deep within the horizontal aspect of the intraparietal sulcus (hIPS) (*11*, *12*). This site partially overlaps with the ventral intraparietal cortex (VIP), an area coding for multimodal spatial movement and tightly interconnected with the nearby lateral intraparietal cortex (LIP), which is involved in saccadic and attention control (*13*–*15*). A model of the VIP-LIP circuitry proposes that it implements a form of vector addition of eye and retinal

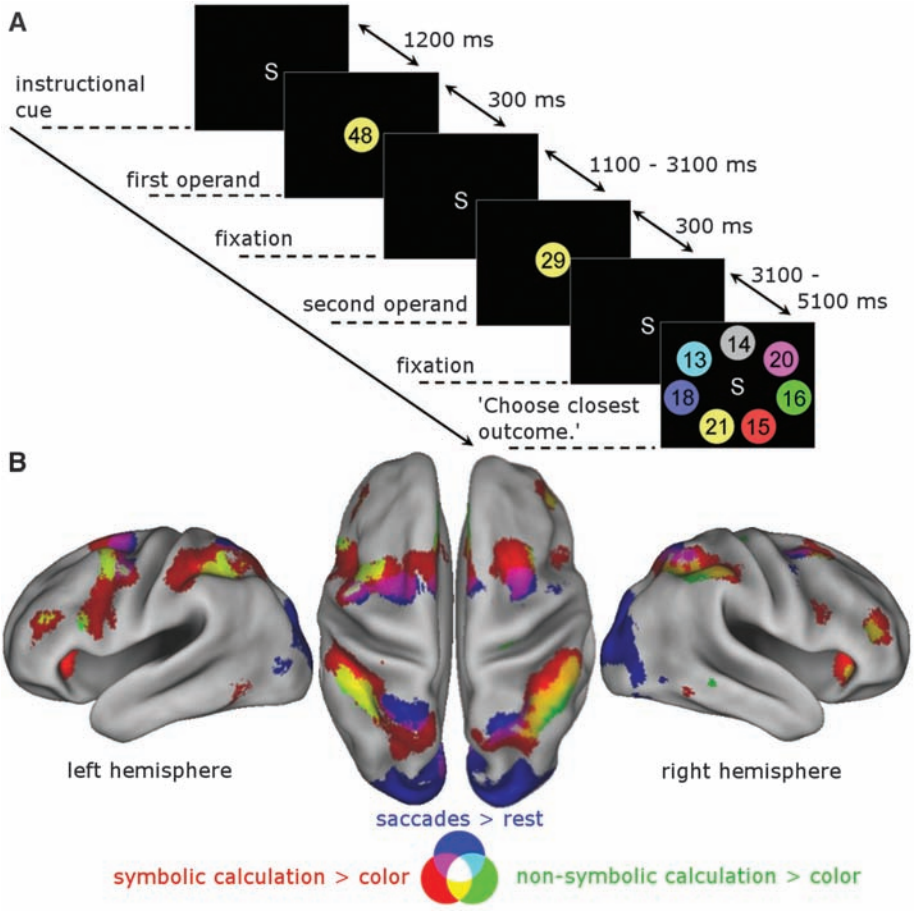

Fig. 1. (**A**) Schematic depiction of a calculation trial. After the initial presentation of an instructional cue [A, S, or C for addition, subtraction (shown here), or color task, respectively], two quantities were presented successively, either as dot patterns or Arabic digits. After a variable delay period, seven response alternatives appeared on the screen and participants had to choose the alternative closest to the actual outcome. (**B**) Brain activation in the calculation task and the saccades localizer task projected on lateral and top views of the brain. The images shown result from contrasting symbolic (red) or nonsymbolic (green) calculation to the color task and from contrasting saccades to rest (blue) ($P = 0.005$, uncorrected).

[1]INSERM, Cognitive Neuroimaging Unit, F-91191 Gif-sur-Yvette, France. [2]Commissariat à l'Energie Atomique (CEA), I2BM, NeuroSpin, F-91191 Gif-sur-Yvette, France. [3]Université Paris-Sud, F-91405 Orsay, France. [4]Institut National de Recherche en Informatique et en Automatique Saclay–Île de France, Orsay, France. [5]Collège de France, Paris, France.

*To whom correspondence should be addressed. E-mail: knops.andre@gmail.com

position information (*16*). We therefore reasoned that this circuit might be co-opted for a similar function in the arithmetic domain. Given the cultural link in left-to-right readers between small numbers and the left side of space, and large numbers and the right side of space, we predicted that mental addition, which increases number size, would be associated with a rightward shift of attention and subtraction with a leftward shift. Hence, the activation pattern in the parietal cortex during addition would resemble the activation pattern associated with a rightward eye movement, whereas subtraction would resemble a leftward eye movement.

In a 3 Tesla functional magnetic resonance imaging (fMRI) scanner, participants first performed a localizer task for eye movements. By contrasting eye movements against fixation, we isolated a set of six cortical regions that are classically associated with saccades, and we used them in all subsequent classifier-based analyses (*17*): the bilateral posterior superior parietal lobule (PSPL), at a site overlapping with the proposed human homolog of monkey area LIP (*18*); the bilateral frontal eye fields (FEF) proper; and two clusters of activation lateral to the FEF (lFEF; Fig. 1B).

Participants performed a second set of fMRI runs during which they moved their eyes either rightward or leftward on randomly intermixed trials. We adopted a machine learning approach to search for a linear combination of these voxel-based activation signals that reliably separated leftward and rightward saccades (*19*). We trained a linear support vector machine as a classifier, using a 10-fold cross-trial validation approach in which the classifier is first trained on a random subset of 90% of activation images (one image per trial), and then performance is evaluated on the remaining 10% of trials. The process was repeated 100 times, each time with a new random assignment of trials. Using only voxels from the bilateral PSPL region, we obtained a mean accuracy across all participants of 70.3 ± 2.4% (1 SE), which is significantly above the chance level of 50% [Student's t test $t(14) = 8.39$, $P < 0.001$]. Analysis by signal detection theory gave similar results [average sensitivity index d' across participants = 1.1 ± 0.15, $t(14) = 7.58$, $P < 0.001$]. Thus, saccade direction, which is known to be coded by neurons in monkey area LIP, could be inferred from fMRI of the human posterior parietal cortex.

We then examined whether the same classifier, without further training, would generalize to approximate arithmetic. In new fMRI runs, participants saw two successive numbers (presented as Arabic numerals or as sets of dots), mentally calculated their approximate sum or difference, and subsequently chose the closest number among seven possible outcomes. We concentrated on brain activation just after the presentation of the second operand, at which time the participants performed the calculation (Fig. 1A). Calculation activated a network of brain areas comprising the bilateral hIPS, prefrontal and premotor areas, with considerable overlap between both notations (Fig. 1B). Calculation overlapped only partially with saccades in the bilateral PSPL, but, as predicted, the classifier trained with bilateral PSPL activations during saccades generalized to calculation images. Equating addition with rightward saccades and subtraction with leftward saccades, the mean accuracy for inferring whether an addition or subtraction was performed, averaged over all participants, was 55.0 ± 1.8%, which is significantly greater than chance [$t(14) = 2.78$, $P = 0.015$; $d' = 0.31 \pm 0.10$, $t(14) = 2.85$, $P = 0.013$].

Further analyses showed that when the saccades classifier was tested with addition images, it classified them as rightward saccades 61% of the time (Fig. 2D), which is above the chance level [$t(14) = 2.35$, $P = 0.03$]. For subtraction, however, only 49.1% of images were classified as leftward saccades [$t(14) = -0.16$, n.s.]. This asymmetry, although unexpected, is congruent with earlier reports of larger rightward saccades in response to large numbers, relative to leftward saccades in response to small numbers (*20*), and might reflect reading habits in Western cultures.

A key aspect of the cortical recycling view is that saccadic areas of the posterior parietal lobule should contribute to calculation, not only when performed with concrete sets of objects but even with Arabic numerals, which are a recent product of human culture. We therefore tested the generalization from saccades to calculation in each notation separately. The saccade-trained classifier could distinguish addition from subtraction with an average accuracy of 54.3 ± 2% for Arabic numerals [$t(14) = 2.26$, $P = 0.02$; $d' = 0.38 \pm 0.11$, $t(14) = 2.1$, $P = 0.054$] and an average accuracy of 55.8 ± 2% for nonsymbolic notation [$t(14) = 2.93$, $P = 0.005$; $d' = 0.38 \pm 0.14$, $t(14) = 2.74$, $P = 0.016$]. Thus, both symbolic and nonsymbolic calculations rely in part on brain circuits for saccadic eye movements.

As a further test of this sharing of resources for nonsymbolic and symbolic arithmetic, we also

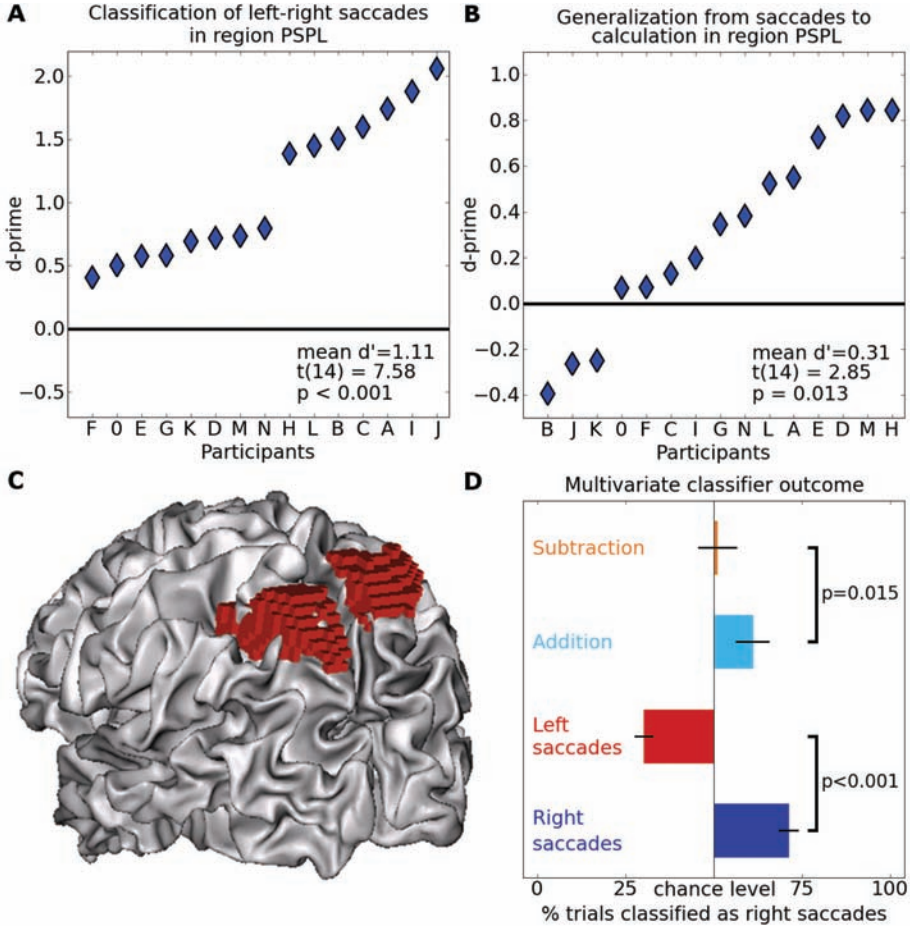

Fig. 2. (**A**) Classification performance (d') for each participant in the saccades task (participants are sorted according to d'). (**B**) Classification performance (d') per participant for generalization of the classifier trained on left/right saccades to subtraction/addition trials. (**C**) Voxel clusters in left and right PSPL region that resulted from the saccade localizer task and served as ROI for the classifier, rendered on the white matter/gray matter boundary. (**D**) Percentages of trials classified as right saccades for subtraction (orange), addition (light blue), and left and right saccades (red and dark blue, respectively).

examined the ability to predict which operation was being performed in one notation, on the basis of a classifier trained to sort additions from subtractions in the other notation. This cross-notation generalization yielded good results, both for the prediction of nonsymbolic calculation from the symbolic notation [mean accuracy 60.7 ± 2.5%, $t(14) = 4.37$, $P < 0.001$; $d' = 0.53 \pm 0.16$, $t(14) = 3.32$, $P = 0.005$] and vice versa [mean accuracy 62.2 ± 2.1%, $t(14) = 5.71$, $P < 0.001$; $d' = 0.75 \pm 0.14$, $t(14) = 5.39$, $P < 0.001$]. This finding indicates that the PSPL region is comparably involved in solving mental arithmetic problems in both notations. Approximate arithmetic with sets of dots is part of an inherited number sense available to infants (*21*) and nonhuman primates (*22*), but the cross-notation generalization proves that the corresponding brain circuitry is also used by arithmetic with culturally specific Arabic numerals.

Given the observed parietal cross-talk, one may wonder whether the arithmetic task, although involving only central visual presentations and a constantly present fixation point, led to overt eye movements. Eye position was continuously monitored throughout fMRI, and we found no detectable change in horizontal fixation at or around the time of the arithmetic calculation (*17*). Furthermore, the observed cross-talk was specific to the posterior parietal cortex. Activation patterns in FEF and lFEF could be reliably used to classify left versus right saccades [56.9 ± 2.4%, $t(14) = 2.87$, $P = 0.012$, and 57.8 ± 2.4%, $t(14) = 3.3$, $P = 0.005$, respectively], but this classification did not generalize to addition versus subtraction {49 ± 5.3% [$t(14) = –0.18$, $P = 0.86$] and 49.2 ± 6.3% [$t(14) = –0.12$, $P = 0.9$, respectively]}. The absence of decodable FEF activation during arithmetic confirms that calculation specifically engages parietal rather than frontal spatial mechanisms and involves covert visuospatial mechanisms, not overt eye movements. As a final test of the specificity of our results to the PSPL area, we repeated the major analyses with two control regions (hand motor area M1 and hIPS). None of these regions yielded better-than-chance generalization from saccades to calculation (*17*).

We demonstrated that a multivariate classifier can distinguish between brain activations during mental addition and subtraction, after having been trained on images from a separate experiment requiring saccades to the right or left. This generalization was observed with numbers presented either as Arabic symbols or as nonsymbolic sets of dots, which implies shared cognitive processes between both notations. The observed generalization goes beyond previous demonstrations of classifier-based decoding of line orientation and other pictorial contents from early visual areas (*23*–*26*); object identity and category from the ventral visual cortex (*27*); noun identity from distributed cortical regions (*28*); or intentions from premotor, prefrontal, and striatal sites (*29*). Although generalization was found across different image sizes (*27*), from real to imagined images (*26*), or from trained nouns to novel nouns (*28*), inference remained confined to the trained domain. In contrast, the present research demonstrates generalization from a low-level sensorimotor task to a high-level cognitive task involving learned cultural symbols.

Our results confirm a prediction first made by Hubbard *et al*. (*13*) that mental calculation can be likened to a spatial shift along a mental number line. In a certain sense, when a Western participant calculates 18 + 5, the activation moves "rightward" from 18 to 23. This spatial shift relies on neural circuits in the PSPL shared with those involved in updating spatial information during saccadic eye movements. The findings are reminiscent of the "embodied cognition" perspective, which stipulates that perceptual and action mechanisms lie at the core of human abstract thinking (*30*). However, the recycling view that we propose does not imply that all concepts originate in sensorimotor learning. Indeed, there is ample evidence that numerical concepts have a long evolutionary history and a dedicated neuronal circuitry in the intraparietal cortex, partially distinct from neighboring visuospatial circuits (*31*). Our proposal is that human mathematics builds from foundational concepts (space, time, and number) by progressively co-opting cortical areas whose prior organization fits with the cultural need. The PSPL area, perhaps because of its capacity for vector addition during eye movement computation (*16*), appears to have a connectivity or internal structure relevant to arithmetic.

The contribution of the PSPL appears to be fundamentally different from the function of other regions such as the FEF or hIPS, where no generalization from saccades to calculation was found. The PSPL is active not only during saccades but during a broad variety of tasks involving as a common denominator the representation, updating, or attention to spatial locations. This makes it an ideal site for explaining the broad variety of numerical-spatial interactions that have been observed behaviorally with eye, hand, or attention movements (*13*).

Like any fMRI study, the present work is correlative and cannot establish whether the observed PSPL activation plays a causal role in calculation. One interpretation is that the PSPL is causally recruited during the actual computation of the result of arithmetic operations. Another is that calculation is effected by other means and that the PSPL activation merely reflects a subsequent spread of activation to visuospatial areas, perhaps because the final numerical result attracts attention on the mental number line. To separate those alternatives, future work should evaluate the impact of temporary or permanent lesions; for instance, using transcranial magnetic stimulation of dorsal parietal areas, which has already been shown to cause joint impairments in attentive visual search and arithmetic (*32*).

References and Notes

1. S. Dehaene, L. Cohen, *Neuron* **56**, 384 (2007).
2. M. A. Changizi, Q. Zhang, H. Ye, S. Shimojo, *Am. Nat.* **167**, E117 (2006).
3. S. Dehaene, *The Number Sense: How the Mind Creates Mathematics* (Oxford Univ. Press, New York, 1997).
4. C. R. Gallistel, *The Organization of Learning* (MIT Press, Cambridge, MA, 1990).
5. G. Lakoff, R. E. Nuñez, *Where Mathematics Comes From* (Basic Books, New York, 2000).
6. S. Dehaene, V. Izard, E. Spelke, P. Pica, *Science* **320**, 1217 (2008).
7. M. H. Fischer, A. D. Castel, M. D. Dodd, J. Pratt, *Nat. Neurosci.* **6**, 555 (2003).
8. S. Dehaene, S. Bossini, P. Giraux, *J. Exp. Psychol. Gen.* **122**, 371 (1993).
9. K. McCrink, S. Dehaene, G. Dehaene-Lambertz, *Percept. Psychophys.* **69**, 1324 (2007).
10. M. Zorzi, K. Priftis, C. Umilta, *Nature* **417**, 138 (2002).
11. S. Dehaene, E. Spelke, P. Pinel, R. Stanescu, S. Tsivkin, *Science* **284**, 970 (1999).
12. S. Dehaene, M. Piazza, P. Pinel, L. Cohen, *Cogn. Neuropsychol.* **20**, 487 (2003).
13. E. M. Hubbard, M. Piazza, P. Pinel, S. Dehaene, *Nat. Rev. Neurosci.* **6**, 435 (2005).
14. J. R. Duhamel, C. L. Colby, M. E. Goldberg, *Science* **255**, 90 (1992).
15. C. L. Colby, M. E. Goldberg, *Annu. Rev. Neurosci.* **22**, 319 (1999).
16. A. Pouget, S. Deneve, J. R. Duhamel, *Nat. Rev. Neurosci.* **3**, 741 (2002).
17. Materials and methods are available as supporting material on *Science* Online.
18. M. I. Sereno, S. Pitzalis, A. Martinez, *Science* **294**, 1350 (2001).
19. K. A. Norman, S. M. Polyn, G. J. Detre, J. V. Haxby, *Trends Cogn. Sci.* **10**, 424 (2006).
20. M. H. Fischer, N. Warlop, R. L. Hill, W. Fias, *Exp. Psychol.* **51**, 91 (2004).
21. K. McCrink, K. Wynn, *Psychol. Sci.* **15**, 776 (2004).
22. J. F. Cantlon, E. M. Brannon, *PLoS Biol.* **5**, e328 (2007).
23. J. D. Haynes, G. Rees, *Nat. Neurosci.* **8**, 686 (2005).
24. Y. Kamitani, F. Tong, *Nat. Neurosci.* **8**, 679 (2005).
25. K. N. Kay, T. Naselaris, R. J. Prenger, J. L. Gallant, *Nature* **452**, 352 (2008).
26. B. Thirion *et al*., *Neuroimage* **33**, 1104 (2006).
27. E. Eger, J. Ashburner, J. D. Haynes, R. J. Dolan, G. Rees, *J. Cogn. Neurosci.* **20**, 356 (2008).
28. T. M. Mitchell *et al*., *Science* **320**, 1191 (2008).
29. J. D. Haynes *et al*., *Curr. Biol.* **17**, 323 (2007).
30. R. L. Goldstone, L. W. Barsalou, *Cognition* **65**, 231 (1998).
31. O. Tudusciuc, A. Nieder, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14513 (2007).
32. S. M. Göbel, M. F. Rushworth, V. Walsh, *Cortex* **42**, 774 (2006).

33. The fMRI experiment was part of a general research program on functional neuroimaging of the human brain that was sponsored by the Atomic Energy Commission (D. Le Bihan). We thank L. Hertz-Pannier, A. Jobert, and the NeuroSpin staff for their help. Supported by INSERM, CEA, a McDonnell Foundation centennial fellowship (S.D.), and a Volkswagen Foundation grant (S.D. and A.K.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1171599/DC1
Materials and Methods
Figs. S1 to S3
References

29 January 2009; accepted 27 April 2009
Published online 7 May 2009;
10.1126/science.1171599
Include this information when citing this paper.

Attention, Perception, & Psychophysics
2009, 71 (4), 803-821
doi:10.3758/APP.71.4.803

Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect

André Knops and Arnaud Viarouge
INSERM, Unité 562, Gif-sur-Yvette, France
CEA, I2BM, NeuroSpin, Gif-sur-Yvette, France
and Université Paris-Sud, Orsay, France

and

Stanislas Dehaene
INSERM, Unité 562, Gif-sur-Yvette, France
CEA, I2BM, NeuroSpin, Gif-sur-Yvette, France
Université Paris-Sud, Orsay, France
and Collège de France, Paris, France

When we add or subtract, do the corresponding quantities "move" along a mental number line? Does this internal movement lead to spatial biases? A new method was designed to investigate the psychophysics of approximate arithmetic. Addition and subtraction problems were presented either with sets of dots or with Arabic numerals, and subjects selected, from among seven choices, the most plausible result. In two experiments, the subjects selected larger numbers for addition than for subtraction problems, as if moving too far along the number line. This operational momentum effect was present in both notations and increased with the size of the outcome. Furthermore, we observed a new effect of spatial–numerical congruence, related to but distinct from the spatial numerical association of response codes effect: During nonsymbolic addition, the subjects preferentially selected numbers at the upper right location, whereas during subtraction, they were biased toward the upper left location. These findings suggest that approximate mental arithmetic involves dynamic shifts on a spatially organized mental representation of numbers. Supplemental materials for this study may be downloaded from app.psychonomic-journals.org/content/supplemental.

Humans are able to represent and process approximate numerosities in a language-independent, analogue fashion (Dehaene, 1997). This capability, referred to as *number sense*, can be demonstrated already very early in life. In particular, Xu and Spelke (2000) showed that 6-month-old infants could discriminate between large numerosities of 8 or 16 elements, and McCrink and Wynn (2004) showed that 9-month-old infants expected the approximate outcome of operations such as $10 - 5$ or $5 + 5$ performed on sets of dots.

Humans seem to share this capability with many other species, including rats, pigeons, parrots, dolphins, lions, and primates (Dehaene, 1997). The variability of the responses increases linearly with the size of the numerosities involved, thus conforming to Weber's law. Similar numerical competence has been observed in monkeys that could discriminate between visual sets of objects on the basis of numerosity alone (Brannon & Terrace, 1998, 2000). The tight psychophysical parallels between these data and those observed in humans (Cantlon & Brannon, 2006) suggest a common basis for *number sense* in phylogeny.

Indeed, at the neuroanatomical level, converging evidence from electrophysiological studies in the macaque monkey (Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2004), human neuroimaging studies (Dehaene, Piazza, Pinel, & Cohen, 2003; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Piazza, Pinel, Le Bihan, & Dehaene, 2007), and patient studies (Dehaene & Cohen, 1997) points to the bilateral horizontal segment of the intraparietal cortex as being crucial for representing numerical magnitudes. In this area, numerosity is encoded by neurons tuned to approximate number: A quantity such as 4 is represented by the distributed activity of overlapping neurons that prefer about three, about four, or about five objects (Nieder et al., 2002; Nieder & Miller, 2004). Importantly, in humans at least, this representation seems to be abstract, in the sense that both symbolic (i.e., Arabic numerals) and nonsymbolic (e.g., dot patterns) numerosities access this representation (Piazza et al., 2007).

A. Knops, knops.andre@gmail.com

This is in accordance with the view that during development, number symbols such as number words and Arabic numerals are being mapped onto the existing analogue magnitude representation (Gallistel & Gelman, 1992). An important difference between symbolic and nonsymbolic arithmetic, however, is the higher precision that can be achieved with number symbols (Pica, Lemer, Izard, & Dehaene, 2004). At the neurophysiological level, a theoretical model has been proposed that relates this higher precision to narrower tuning curves of the neurons coding for symbolic magnitudes, so that some neurons, at least, would become sharply tuned to precisely four objects, not three or five (Verguts & Fias, 2004).

In contrast to the vast knowledge about the representation and neural correlates of numerical magnitude, little is known about the exact mechanisms and neural structures that combine these magnitudes during mental calculation. Characterizing the psychophysical laws of approximate mental arithmetic, with both symbolic and nonsymbolic numerals, is the main goal of the present study. Nonsymbolic calculation in this context means the mental manipulation (addition and subtraction) of quantities presented as dot patterns. Delineating the properties and mechanisms of nonsymbolic arithmetic may help us to understand the nature and development of symbolic arithmetic. At present, only a few experiments are available on this topic. In a series of experiments, Barth et al. (2006) showed that human adults and 5-year-old children were capable of mastering basic arithmetic operations (i.e., addition and subtraction) with nonsymbolic stimuli (dot patterns). In particular, children showed above-chance performance in an addition task, although they performed no better than chance on a symbolic version of this task. A detailed mathematical theory of the results was developed, suggesting that variability arose both from the imprecise representation of the operands and from the representation of the computed result. However, the task (decide whether the outcome of an operation is larger or smaller than a specified numerosity) did not afford any possibility of revealing putative biases in the computed addition or subtraction results.

Recently, Gilmore, McCarthy, and Spelke (2007) showed that 5-year-old children's performance in symbolic calculation tasks is marked by characteristics similar to those of their nonsymbolic arithmetic system, suggesting that they relied on their nonsymbolic number knowledge to solve approximate symbolic calculation problems.

Cordes, Gallistel, Gelman, and Latham (2007) investigated the contribution of different sources of variation in a task in which subjects were told to compute the sum or the difference of two nonsymbolic quantities (arhythmic series of brief flashes) and to produce the result *n* by pushing a button *n* times. The authors compared the normalized residuals of several regression models that took into account different factors, such as the variability of the individual operands of a given problem or the variability due to the sum or the difference between the operands—that is, the outcome of a problem. The major determinant of variability in nonverbal arithmetic arose from the representation of the individual numerical magnitudes entering into the calculation process, rather than from mental computation or motor production processes.

Using a different task (decide whether a proposed numerosity is or is not the correct result of an operation), McCrink, Dehaene, and Dehaene-Lambertz (2007), too, found evidence that mental calculation with nonsymbolic numerosities follows Weber's law. They showed that for additions and subtractions with dot patterns matched for physical factors other than numerosity, both the mean number chosen by the participants and the variability of these chosen numbers increased with the correct outcome. Surprisingly, however, the values chosen by the subjects were not centered on the correct result but were influenced by the arithmetic operation that had to be carried out. With mental addition, the subjects' estimated outcomes tended to be larger than the actual outcomes, whereas they tended to be smaller than the actual outcomes with subtraction.

McCrink et al. (2007) argued that this bias showed similarity to a perceptual phenomenon called *representational momentum* (Freyd & Finke, 1984). When they watch a moving object suddenly disappear, subjects tend to misjudge its final position and report a position displaced in the direction of the movement (Halpern & Kelly, 1993; T. L. Hubbard, 2005; Kerzel, 2003). Analogously, McCrink and colleagues described their finding as an *operational momentum* (OM), since the misjudgment was related to the arithmetic operation carried out, and suggested that the subjects were moving "too far" on the number line.

The assumed parallels between representational momentum and the OM are in line with the hypotheses made by Hubbard and colleagues concerning the possible interplay of the posterior parietal lobes and the horizontal aspects of the intraparietal cortex during calculation (E. M. Hubbard, Piazza, Pinel, & Dehaene, 2005). Speculatively, mental calculation was proposed to correspond to a displacement on the spatially organized mental representation of numerical magnitude. This displacement might be mediated by the same parietal mechanisms as those involved in guiding saccadic eye movements. Since larger numbers are represented to the right side of space on the mental representation and smaller numbers to the left, this would yield rightward "movements" with addition and leftward "movements" with subtraction. Moreover, it has been shown that saccade latencies to the left or to the right covary parametrically with the numerical size of previously presented numbers (Fischer, Warlop, Hill, & Fias, 2004). Larger numbers were more rapidly followed by saccades to the right, and smaller numbers by saccades to the left. This points to a systematic and parametric relation between numerical magnitude and shifts of spatial attention, which, in turn, are thought to be mediated by oculomotor circuits (Sheliga, Riggio, & Rizzolatti, 1994).

Given the cortical parallels between circuits for eye movements and those for mental arithmetic, one might wonder what determines the OM effect. Sticking with the metaphor of perceiving calculation as motion along a mental number line, what determines this movement? Does the numerical magnitude of the operands determine the size of the OM effect? In the case of actual movement

and perceptual representational momentum, the effect is known to be modulated by several factors (for a review, see T. L. Hubbard, 2005): The amount of misjudgment is stronger for movements to the right (Halpern & Kelly, 1993), increases with increasing speed of the moving object (Freyd & Finke, 1985), and is larger with apparent, as compared with smooth, motion when the gaze is fixed (Kerzel, 2003). Interestingly, the effect can be increased by a secondary task that involves counting onward up to 30 in steps of one, two, or three (Halpern & Kelly, 1993).

Relative to this background, the present study had several goals. First, we systematically investigated the influence of the numerical magnitude of operands on the OM effect. As we described above, the representational momentum effect is influenced by several factors, such as speed of the moving object, the presence of landmarks, or representational gravity (T. L. Hubbard, 2005). In contrast, it is largely unknown what factors determine the OM effect. Since it has been found that saccadic eye movements are systematically related to the numerical magnitude of a previously presented number, the attentional shifts that accompany mental arithmetic might systematically covary with the numerical magnitude of the operands. This might, in turn, change the amplitude of the OM effect.

To study this influence in more detail, we introduced a method of assessing the psychophysical properties of symbolic and nonsymbolic calculation that was more efficient than the method used in McCrink et al. (2007). After presenting an arithmetic problem (e.g., 48 dots + 21 dots), instead of presenting subjects with a single proposed result (e.g., 69 dots) and asking them to evaluate its correctness with a yes/no answer, we increased the amount of information gained per trial by presenting subjects with seven closely spaced alternatives on each trial (e.g., 35, 41, 49, 58, 69, 82, and 98; note that numerical spacing between response alternatives is linear on a log scale) and directly recording their preferred outcome. This method presents the advantage of yielding, on each trial, an estimate of the subjects' arithmetic estimate, almost as if the subjects had given a spoken response, yet without requiring the actual production of any number words. Importantly, this response mode could be used identically for problems in both symbolic and nonsymbolic notation.

Second, we investigated whether OM is restricted to nonsymbolic calculation or can also be observed in symbolic calculation (i.e., using Arabic numbers). Finding similar effects for both notations would support the assumption that both nonsymbolic and symbolic magnitudes are represented on a common mental scale and that approximate calculation with symbolic operands also relies on this mental *number line* (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). The present data set allowed for a precise quantitative estimation of the size of the OM effect and of its variation with operation type, as well as with the size of the operands. Through such analyses, we hoped to delineate the determinants of the OM effect and gain a more thorough understanding of the underlying mechanisms, thus clarifying whether they are common to symbolic and nonsymbolic calculations.

Finally, a third goal was to probe putative associations between arithmetic operations and movements in space. E. M. Hubbard et al. (2005) speculated that shared parietal mechanisms could be involved in guiding eye movements and calculations on the mental number line, so that addition would be analogous to a rightward motion and subtraction to a leftward motion. If such an internal sense of spatial motion accompanies mental arithmetic, we might expect the subjects' responses to be spatially biased as a function of the type of operation they performed. In our experiment, the seven proposed choices were presented in a circular array, making it possible to measure any spatial preference, both on the left–right and on the top–down axes, as a function of problem type. If E. M. Hubbard et al.'s prediction holds, we should expect to find more rightward responses to addition problems and more leftward responses to subtraction problems.

EXPERIMENT 1

Method

Subjects. Sixteen volunteers (9 female; mean age = 23.5 years; *SD* = 3.6; range, 19–34 years) took part in the experiment after having given their written informed consent. All but one were right-handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971).

Stimuli. We first selected 18 calculation problems, 9 additions and 9 subtractions. The operands were identical for addition and subtraction (see Table 1). The first operand was 32, 48, or 60. The second operand was created in relation to the first operand—that is, as a fixed proportion of the first operand, being 31%, 45%, or 61% for small, medium, and large second operands, respectively. Together, the combinations of these three initial values (i.e., small, medium, and large first operand) and three amounts of change (i.e., small, medium, and large second operand) generated $3 \times 3 = 9$ problems. Apart from the correct result, eight deviant results were created for each arithmetic problem. These deviants were arranged as a geometric series (i.e., were linearly spaced on a logarithmic scale) and ranged from double the correct result to half the correct result [technically, they were generated as round($c \times 2^{i/4}$), where c is the correct result and i ranges from -4 to $+4$]. To avoid a strategy of always selecting the response falling in the middle of the proposed range, only seven out of those nine possible results were presented. On 50% of the trials, we presented the upper seven (high range), and thus the correct result was the third largest numerosity (although numerosities were randomly mixed with respect to both spatial position on the screen and temporal order of appearance [see the Procedure section below]). On the other 50% of the trials, the lower seven choices were shown (low range), and the correct result was therefore the fifth largest numerosity.

Because the experimental design was organized around a small number of arithmetic problems, it was important to prevent the subjects from memorizing them in symbolic form. To this aim, the problems and their proposed results were randomly "jittered," differently on each trial. First, the operands were jittered by a random value from 0 to ± 2, so that the actual outcome would remain unchanged (i.e., for a given task 48 + 18, the jittered operands could be 47 + 19). Second, all of the seven proposed results were jittered up or down by a random value (fixed for a given trial). This random value had a mean value of zero and was drawn from a flat distribution on a logarithmic scale, in the range plus or minus half the numerical interval between the correct result and the first deviant above or below it. Technically, this was achieved by drawing a random number r between -0.5 and 0.5 and defining the proposed results as round($c \times 2^{(r+i)/4}$), where i again ranges from -4 to $+4$.

Table 1
All Basic Arithmetic Problems Presented in Experiment 1 and Their Correct and Deviant Results

Operands		Results and Mean Deviants (Not Jittered)								
O1	O2	1/2	1/1.7	1/1.4	1/1.2	1/1	1.2/1	1.4/1	1.7/1	2/1
					Addition					
32	9	21	24	29	34	41	49	58	69	82
32	14	23	27	33	39	46	55	65	77	92
32	19	26	30	36	43	51	61	72	86	102
48	15	32	37	45	53	63	75	89	106	126
48	21	35	41	49	58	69	82	98	116	138
48	29	39	46	54	65	77	92	109	129	154
60	19	40	47	56	66	79	94	112	133	158
60	29	45	53	63	75	89	106	126	150	178
60	38	49	58	69	82	98	117	139	165	196
					Subtraction					
32	9	12	14	16	19	23	27	33	39	46
32	14	9	11	13	15	18	21	25	30	36
32	19	7	8	9	11	13	15	18	22	26
48	15	17	20	23	28	33	39	47	55	66
48	21	14	16	19	23	27	32	38	45	54
48	29	10	11	13	16	19	23	27	32	38
60	19	21	24	29	34	41	49	58	69	82
60	29	16	18	22	26	31	37	44	52	62
60	38	11	13	16	18	22	26	31	37	44

Note—The actual problems presented to the subjects were jittered by a small random amount (see the Method section), such that the correct outcome was never presented.

For example, for the problem $32 + 9 = 41$, with the closest deviants being 34 and 49, the closest proposed result could be jittered anywhere from $38 = \text{round}(41 \times 2^{-0.5/4})$ to $45 = \text{round}(41 \times 2^{+0.5/4})$. We ensured that the correct outcome would never appear as a response alternative. All the proposed results fell between 6 and 217.

All the problems were presented both in Arabic notation and as dot patterns (Figure 1 shows an example of a nonsymbolic trial). The notation of response alternatives in a given trial was always identical with the notation of the operands. Both notations were displayed in black within a colored circle that was presented on a black background. Each circle had a diameter of 120 pixels (3.55 cm) at a viewing distance of approximately 65 cm (no chinrest was used). Seven different colors were used for the results, whereas the operands' colors were identical (this color manipulation played no role in the present experiment but was designed as a control for a future neuroimaging experiment). The operands were presented successively in the center of the screen. The results were presented at seven locations arranged around the screen center in an ellipsoid fashion. The seven proposed outcomes were counterbalanced in a Latin square, so that the number closest to the correct result appeared once and only once at each location, for each notation and operation.

To prevent the use of nonnumerical cues, the sets of dots representing the nonsymbolic numerosities were designed and generated, using MATLAB, in such a way that dot size changed but total dot area in a given set was always fixed across stimuli. Thus, total occupied area could not serve as a cue for distinguishing between the different numerosities. As a result of this manipulation, average item size covaried inversely with numerosity during the presentation of the operands (i.e., sets with smaller numerosities had larger dots). To avoid memorization effects due to repetition of a particular stimulus, on each trial, the stimulus images were randomly chosen from a set of 10 precomputed images with the given numerosity.

Procedure. A total of 504 trials were presented in 14 blocks. After each block, the subjects were given a chance to rest. Trial structure is shown in Figure 1. Each trial started with the presentation of a fixation cross for 250 msec, which was followed by a blank screen for 150 msec and an uppercase letter ("A" for addition or "S" for subtraction; 1,400 msec) indicating the subsequent operation to be performed. After the instruction letter had disappeared, the two operands were presented for 1,000 msec each (Arabic notation) or 1,500 msec each (nonsymbolic notation). Finally, the screen was emptied, and seven proposed results appeared, one by one, every 190 msec, at one of seven possible locations and remained on the screen until response. The notation of the operands always was identical to that of the response alternatives in a given trial. The temporal and spatial order (and thus the numerical order of the response alternatives) in which the seven response alternatives appeared on the screen was randomized for each trial, with the constraint that each response alternative appeared only once at each position in the course of the experiment. After the appearance of the last response alternative, the mouse pointer appeared in the center of the screen, and the subjects had to indicate which numerosity was numerically closest to the actual result by clicking on the respective image. Speed was stressed over accuracy in order to maximize the use of approximation strategies and to avoid explicit calculation (Arabic numerals) or counting (dot patterns). During the experiment, no feedback was provided to the subjects.

The whole experiment lasted approximately 85 min. The paradigm was created using Python 2.4 software (python.org) run on a portable personal computer (Siemens/Fujitsu; 1.6 GHz).

Results

Because the subjects occasionally reported selecting the wrong operation (subtraction instead of addition), we first trimmed the data to exclude all responses outside a three standard deviations range around a subjects' overall mean of the difference between the logarithm of the subjects' choice and the logarithm of the correct value (see below). This procedure excluded 1.8% of all the responses.

Nonrandom distribution of responses. To start with, we checked whether the subjects chose among the proposed choices at random. On each trial, seven response alternatives were presented. The experimental design included two trial types, which were distinguished by whether the result closest to the correct outcome was the third or the fifth alternative in numerical order (naturally, these choices were

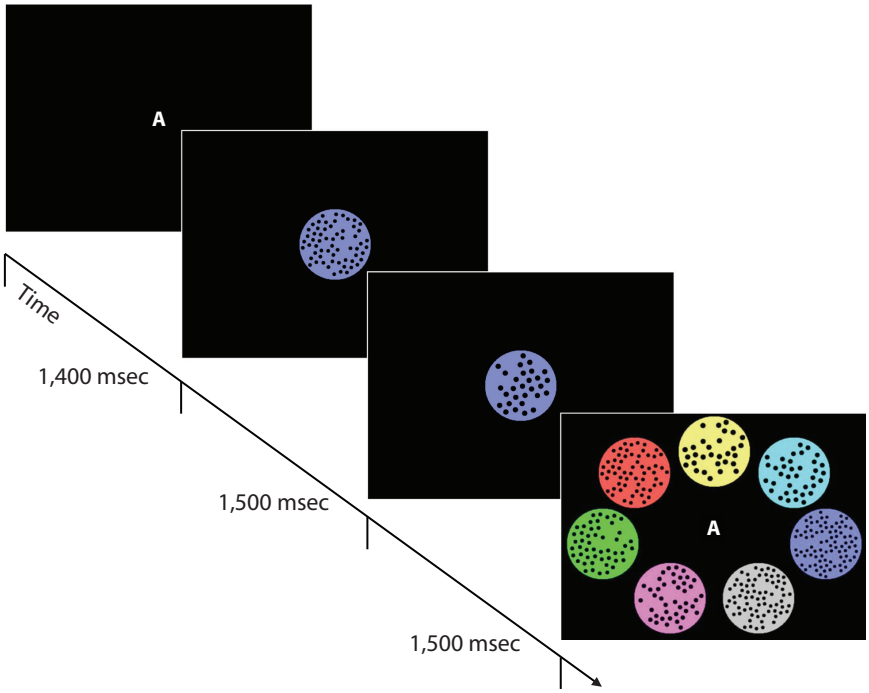

Figure 1. Screenshots taken from a trial with nonsymbolic addition to illustrate the task and the trial sequence. After an initial appearance of the letter "A" or "S," indicating addition and subtraction, respectively, the first and second operands successively appeared in the center of the screen. The response alternatives were presented on screen in random order, separated from each other by a delay of approximately 190 msec to direct the subjects' attention to each of the response alternatives. For trials in symbolic notation, the general layout of the trials was identical. Instead of presenting a set of black dots, black Arabic numbers were presented in the center of the colored circles.

presented to the subjects in random order). If the subjects were able to solve the arithmetic problems, their response choices should show a nonflat distribution, presumably centered close to the correct value (neglecting for the moment the OM effect) and, therefore, shifting across those two trial types. In contrast, if they responded randomly, we would not expect any differences in the frequency of choosing a particular response alternative. In Figure 2, we plot response frequency for each notation and each operation, separately for trials in which the correct answer was third or fifth. Responses were clearly distributed nonrandomly. For symbolic notation, the peak of the distribution was always centered on the response alternative closest to the correct outcome. This also held for nonsymbolic addition, although the distributions were broader, implying a larger variability for nonsymbolic than for symbolic notation. For nonsymbolic subtraction, the subjects always preferred the smallest possible outcome, suggesting an underestimation bias that will be analyzed further below.

These conclusions were supported by an ANOVA over the different response categories, with percentage of choice as the dependent variable and rank of the subject's choice (one to seven) and trial type (third or fifth value correct) as factors. Only for nonsymbolic addition was there no effect of rank [$F(6,90) = 2.02$, $p = .141$]. For all the other operations in either notation, a main effect of rank significantly influenced the subjects' choices (all $ps < .001$). Most important, however, for each operation and in each notation, a significant rank $\times$ range interaction was observed (all $ps < .001$).

Linear increase of response and response variability with the correct result. We next examined how the subjects responded to our different arithmetic problems. The left column of Figure 3 shows the subjects' mean responses (chosen values) as a function of the size of the correct result, separately for the two notations (Figures 3A and 3B) and the two operations. If the subjects were able to solve the arithmetic problems, the chosen value should increase as a function of the correct outcome for both nonsymbolic and symbolic notation. With increasing numerical magnitude, theory predicts an increasing variability of the chosen values (see the appendix in Barth et al., 2006). Finally, according to Weber's law, the increase in the chosen values should be paralleled by a proportional increase in response variability, as expressed in terms of their respective standard deviation, resulting in a constant coefficient of variation (CV, the ratio of the standard deviation and mean of the subjects' responses) across arithmetic problems of different numerical magnitude.

As can be seen in Figure 3, the subjects' responses (depicted as circles in Figure 3) increased as a function of the correct outcome for both addition (black circles) and subtraction (gray circles) and in both notations. This impression was confirmed by one-way repeated mea-

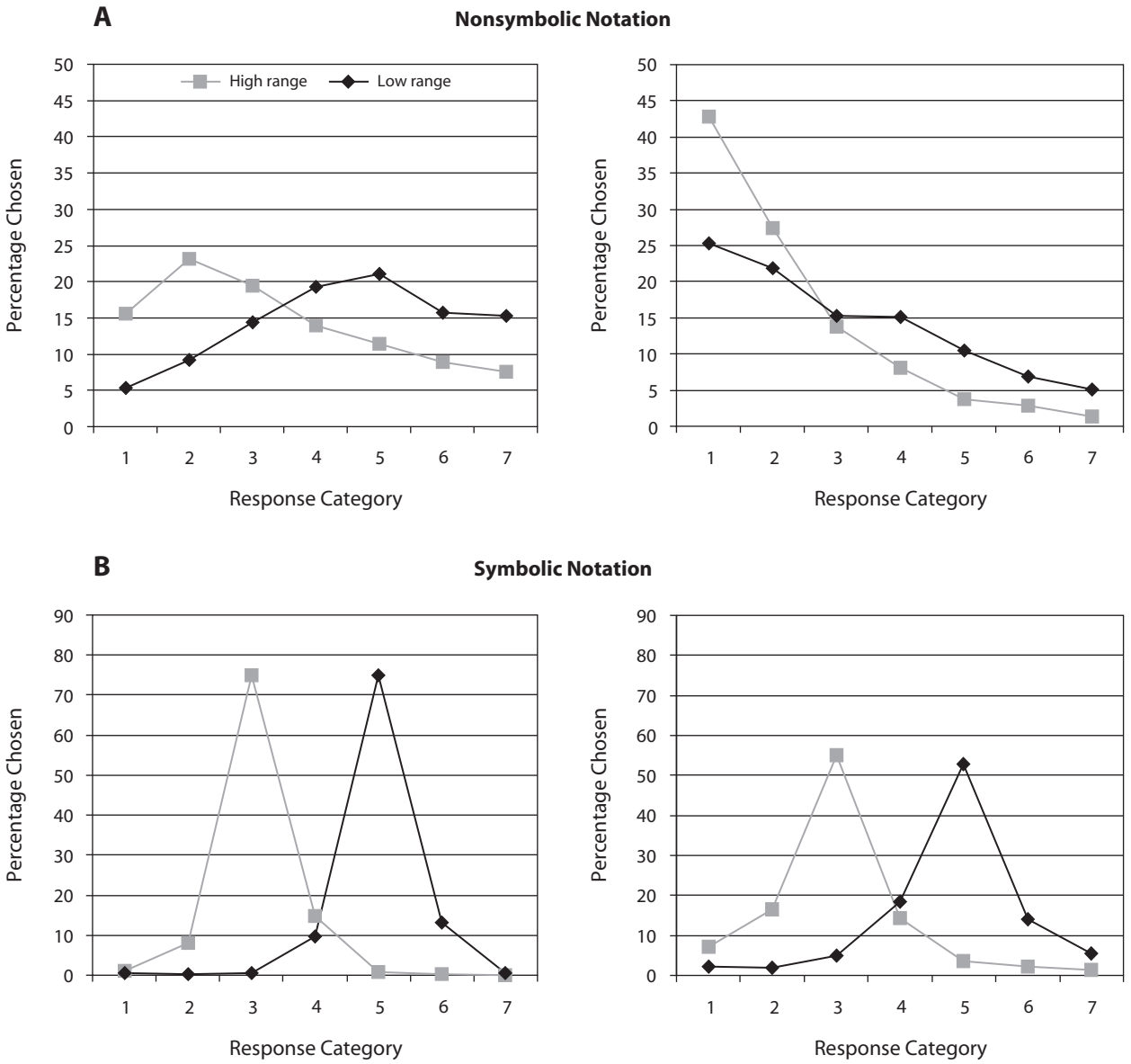

Figure 2. Distribution of the subjects' choices across the seven proposed results, averaged over all arithmetic problems, separately for addition (left column) and subtraction (right column) in nonsymbolic (A) and symbolic (B) notation. The subjects' responses were not distributed randomly but, rather, depending on the range of response alternatives presented (high or low range), were centered around the value that was closest to the correct outcome (fifth for low range and third for high range). Additional influences of operation (smaller choices for subtraction than for addition) and of notation (underestimation bias for nonsymbolic, as compared with symbolic, notation) are also visible.

sures ANOVAs, separately for each operation and each notation, with problem size as the only factor. A highly significant linear trend was observed[1]: for nonsymbolic addition [$F(8,120) = 180.52$, $p < .001$, $\varepsilon = .46$; ε denotes the Huynh–Feldt term (Huynh & Feldt, 1976) of correction for nonsphericity] and subtraction [$F(8,120) = 212.27$, $p < .001$, $\varepsilon = .73$], as well as for symbolic addition [$F(8,120) = 2{,}460$, $p < .001$, $\varepsilon = .98$] and subtraction [$F(8,120) = 558.51$, $p < .001$, $\varepsilon = .85$].

A similar repeated measures ANOVA over the standard deviations (squares) of the mean chosen values showed that the subjects' choices became more and more variable with increasing numerical size of the result. This was true for both addition (black) and subtraction (gray) with the nonsymbolic notation [Figure 3A; $F(8,120) = 42.57$, $p < .001$, $\varepsilon = .88$, and $F(8,120) = 23.02$, $p < .001$, $\varepsilon = .78$, respectively] and with the symbolic notation [Figure 3B; $F(8,120) = 27.01$, $p < .001$, $\varepsilon = .68$, and $F(8,120) = 15.61$, $p < .001$, $\varepsilon = .88$, respectively].

As can be seen in the lower left parts of Figure 3A and 3B, the CV was essentially constant across the whole range of outcomes for addition and subtraction with both notations. This was tested statistically with the same four repeated measures ANOVAs, with problem size as the

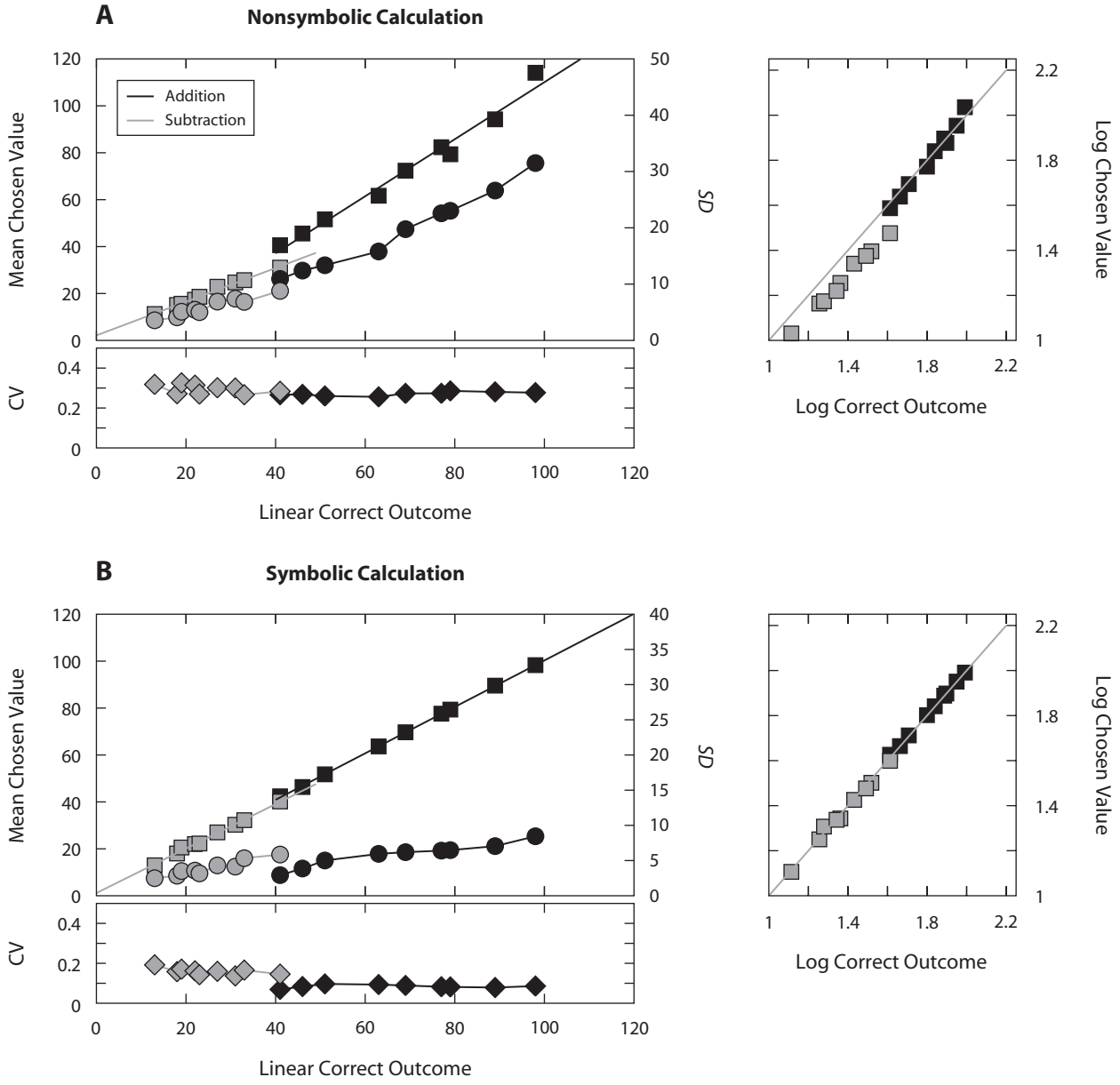

Figure 3. Left column: Mean responses (chosen values, squares) of the subjects and standard deviations (circles) plotted against the correct outcome for both nonsymbolic (A) and symbolic (B) addition and subtraction. The lower part of panels A and B depicts the coefficient of variation (CV, diamonds)—that is, the ratio of standard deviation and mean chosen value, plotted against the correct outcome. Right column: The logarithm of the correct outcome plotted against the logarithm of the mean value chosen by the subjects for nonsymbolic (top) and symbolic (bottom) notation. The gray line indicates a ratio of 1—that is, perfect performance.

only factor. For nonsymbolic notation, indeed, the CV did not change as a function of the correct outcome, either for addition ($F < 1$) or for subtraction [$F(8,120) = 1.65$, $p = .121$, $\varepsilon = .97$]. For symbolic notation, this was true only for subtraction [$F(8,120) = 1.27$, $p = .277$, $\varepsilon = .77$], whereas a significant main effect of problem size was obtained for addition [$F(8,120) = 2.90$, $p = .007$, $\varepsilon = .91$].

Constant response variability on a logarithmic scale. If Weber–Fechner's law holds for both symbolic and nonsymbolic calculation, the data should show constant response variability once plotted on a double-logarithmic scale (log of the subjects' responses as a function of the log of the correct results; Dehaene et al., 2003). Therefore, we calculated the difference between the correct outcome and the mean chosen value, once both of them had been transformed to a logarithmic scale, and calculated a repeated measures ANOVA on the standard deviations of these differences, with factors of notation (nonsymbolic and symbolic) and size of the correct result, separately for both operations (addition and subtraction).

Once transformed to a logarithmic scale, the size of the outcome did not systematically influence response vari-

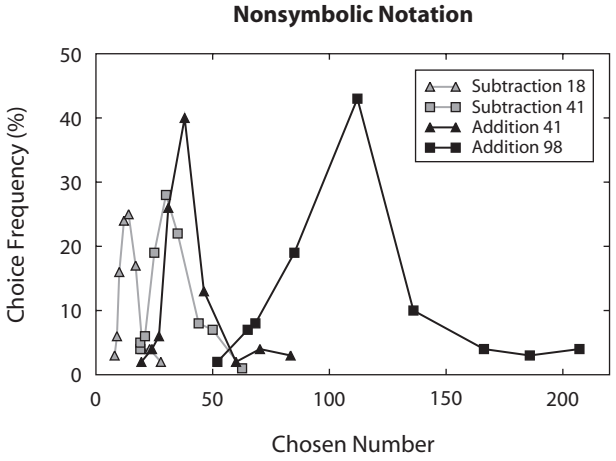

Figure 4. Sample distributions of responses to four nonsymbolic arithmetic problems: the nonsymbolic subtraction problem yielding the smallest value (18), the addition yielding the largest value (98), and the addition and subtraction problems resulting in the same numerical outcome of 41. In each case, quantiles are shown (1%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, and 99%). Note the clear difference between addition and subtraction problems with the same outcome, which is a manifestation of operational momentum.

ability. For both operations, the main effect of problem size failed to reach significance [$F(8,120) = 0.88$, $p = .52$, $\varepsilon = .78$, and $F(8,120) = 1.61$, $p = .14$, $\varepsilon = .87$, for addition and subtraction, respectively]. This analysis thus confirmed that Weber's law held and that performance was determined mostly by the ratio of true outcome and chosen value (or equivalently, by the difference of their logs). However, there were systematic differences between the notations for both operations [$F(1,15) = 94.80$, $p < .001$, and $F(1,15) = 325.82$, $p < .001$, respectively]: Variability was higher for dot patterns than for Arabic numerals.

Since problem size did not influence the results, we carried out a 2×2 repeated measures ANOVA with the factors of notation and operation to evaluate possible interactions between these factors. The two main effects indicated that response variability was higher for the nonsymbolic notation than for the symbolic notation [$F(1,15) = 284.73$, $p < .001$] and when the subjects were engaged in subtraction, as compared with addition [$F(1,15) = 12.56$, $p = .003$]. A significant interaction of both factors [$F(1,15) = 33.93$, $p < .001$] indicated that variability of the responses was higher for subtraction than for addition only for symbolic notation [$t(15) = 7.28$, $p < .001$], whereas for nonsymbolic notation, no significant difference was observed [$t(15) = 1.27$, $p = .23$]. This might point to different underlying processes when Arabic numerals are dealt with (exact calculation for addition vs. more approximate processes for subtraction), whereas no differences were observed for calculation with dot patterns.

Taken together, these results suggest, as predicted by Weber's law, that the logarithms of the numbers involved provide a more compact description of the data. Therefore, all the following analyses were carried out in a logarithmic scale, using as input the difference between the logarithm of the correct outcome and the logarithm of the chosen value. Such analyses also present the advantage of meeting the prerequisites of the ANOVA, which stipulates that all data have a fixed variability.

Operational momentum. Although the subjects' responses increased roughly linearly with the correct result, as is shown in Figure 3, the slope of increase tended to be shallower for subtraction than for addition (especially for nonsymbolic operations). Figure 4 shows the distribution of the subjects' responses for a few selected nonsymbolic addition and subtraction problems, including the case in which the two operations had an identical outcome (41). It can be seen that the subjects' responses tended to be higher for addition than for subtraction. The fact that, for equal objective outcomes, the subjects tended to anticipate a smaller outcome for a subtraction than for an addition problem (McCrink et al., 2007) might be taken as a manifestation of the OM effect.

To quantify this OM effect, we computed a simple estimate of response bias: the mean difference between the log of the subject's responses and the log correct result (see Figure 5). This value was first submitted to a simple ANOVA with notation and operation as factors. A main effect of notation [$F(1,15) = 51.39$, $p < .001$] indicated an overall tendency toward underestimation for the nonsymbolic notation, as compared with rather precise performance for the symbolic notation. Most important, a main effect of operation [$F(1,15) = 21.58$, $p < .001$] provided evidence for an OM effect—that is, a significant bias toward smaller responses for subtraction than for addition. The significant interaction [$F(1,15) = 24.92$, $p < .001$] was due to a larger difference between operations for nonsymbolic notation, as opposed to symbolic

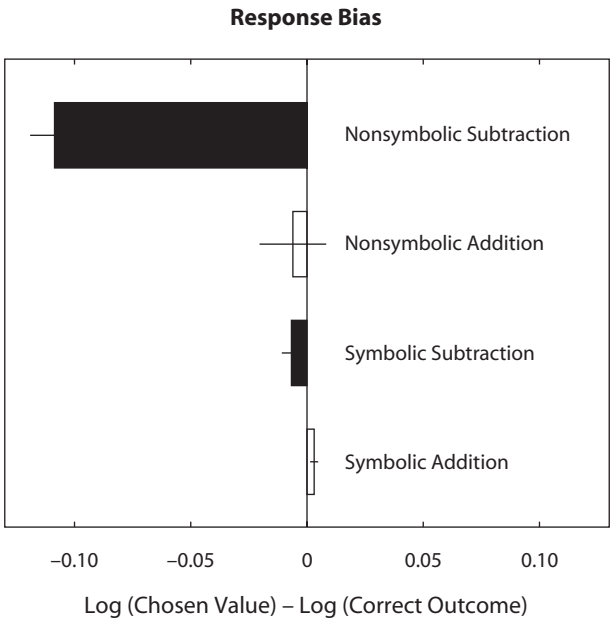

Figure 5. Mean response bias, defined as the difference between the chosen value and the correct outcome, both expressed on a log scale. A negative bias indicates underestimation, and a positive bias indicates overestimation.

notation [$t(15) = -4.99$, $p < .001$]. Crucially, however, subtraction yielded significantly smaller values than did addition with both notations [$t(15) = -4.84$, $p < .001$, for nonsymbolic notation, and $t(15) = 2.73$, $p < .001$, for symbolic notation]. McCrink et al. (2007) explored the OM effect only with nonsymbolic numerosities; the present results provide the first evidence that this effect holds also during approximate symbolic calculation with Arabic numerals.

There was, however, another important difference between notations (see Figure 5). For symbolic operations, the OM took the form of a full crossover effect, with positive deviations relative to the correct result for addition (overestimation) and negative deviations for subtraction (underestimation). For nonsymbolic operations, however, the OM effect was superimposed onto a general tendency to underestimate the correct result. This finding may relate to a general tendency to underestimate the number of dots in visual displays, as has been reported by others (Izard & Dehaene, 2008; Krueger, 1984).

Note that we propose to apply a definition of the momentum effect that differs from the more stringent one used in the visual domain. In the domain of representational momentum, to speak of a significant momentum effect requires the data to have the form of a full crossover effect—that is, values that significantly deviate from zero for both directions, positively and negatively, as a function of movement direction. In the numerical domain, however, we propose to speak of a momentum effect as soon as there is a difference in the mean responses to matched addition and subtraction problems, even if the responses for both operations differ from zero with the same sign, as is the case for nonsymbolic notation in the present study. This stance is useful because numerical data are often affected by an additional general tendency to underestimate the results (Izard & Dehaene, 2008). The key new finding is that this underestimation tendency is modulated by the arithmetic operation, yielding considerable underestimation for subtraction and near-correct values for addition. This interaction in itself can be interpreted as evidence for an OM effect. This point becomes most evident for those problems with identical results: As is shown in Figure 4, even for these objectively equal problems, addition resulted in larger subjective responses than did subtraction (this effect was explored further in Experiment 2).

Influence of the arithmetic operation on the spatial distribution of responses. Since each proposed result appeared equally frequently at each of the seven spatial positions on the screen, by chance alone, we would expect responses to be equally distributed over these different positions. However, if one assumes that calculation resembles a spatial displacement on the mental number line via shifts of spatial attention (E. M. Hubbard et al., 2005), one might expect this internal movement to influence the position of the subjects' responses. Our paradigm therefore offered a unique opportunity to directly test one of the assumptions made by Hubbard and colleagues (E. M. Hubbard et al., 2005): If addition induces a shift of spatial attention to the right and subtraction to the left, this might bias the subjects' clicking on the right or left side of the screen.

Figure 6 depicts the mean percentage with which each of the seven positions on screen was chosen, averaged over subjects, separately for the two types of notations for each type of operations. It is obvious that the data are not equally distributed over the seven positions for each of the four notation $\times$ operation combinations. Particularly in nonsymbolic notation, there is a clear spatial bias, with addition yielding more rightward responses and subtraction more leftward responses. We tested this by computing two 2×2 repeated measures ANOVAs with factors of operation and position for the two top right and top left positions only (separately for nonsymbolic and symbolic stimuli). For nonsymbolic stimuli, the critical interaction between operation and position was significant [$F(1,15) = 6.30$, $p = .024$, $\varepsilon = 1$], indicating a differential pattern of spatial responding as a function of operation. For symbolic stimuli, no such interaction was present [$F(1,15) = 1.88$, $p = .19$], although a trend in the same direction can be seen in Figure 6.

As an additional statistical test of this effect, a simple 2×2 contingency table served as input for computing a χ^2 test over the two positions (top left or top right) for each operation (addition or subtraction). For nonsymbolic notation, a significant χ^2 test [$\chi^2(1) = 5.60$, $p = .018$] confirmed that the spatial distribution of responses varied with the type of arithmetic operation. For symbolic notation, no significant effect was found [$\chi^2(1) = 0.29$, $p = .59$].

Discussion

We replicated the OM effect observed by McCrink et al. (2007) with a multiple choice paradigm. Subjects tended to misjudge the outcome of a mental calculation. The misjudgment was a function of the notation and the operation carried out. In accordance with earlier reports (e.g., Izard & Dehaene, 2008), the subjects generally tended to underestimate results presented in nonsymbolic form. However, this underestimation bias was modulated by the arithmetic operation: In accordance with the OM, the difference between the preferred and the correct results (response bias) was more positive for addition than for subtraction. This OM effect was very strong for nonsymbolic notation, but it was also significant for symbolic notation. The fact that the OM effect is smaller for symbolic stimuli is not surprising, since it merely reflects the subjects' higher precision with symbolic than with nonsymbolic stimuli. However, the mere presence of such an effect for calculation with Arabic numerals is an important novel finding, since it suggests that a similar analogue magnitude representation is used during symbolic and nonsymbolic approximate arithmetic.

Finally, a new finding was that arithmetic operation biased the spatial distribution of responses: The subjects preferentially selected values presented at the upper left position of the computer screen with nonsymbolic subtraction and values presented at the upper right position with nonsymbolic addition. For symbolic notation, a similar but smaller and nonsignificant trend was found. Again, this is presumably because there was not much room for such a bias, given that the subjects so frequently selected the

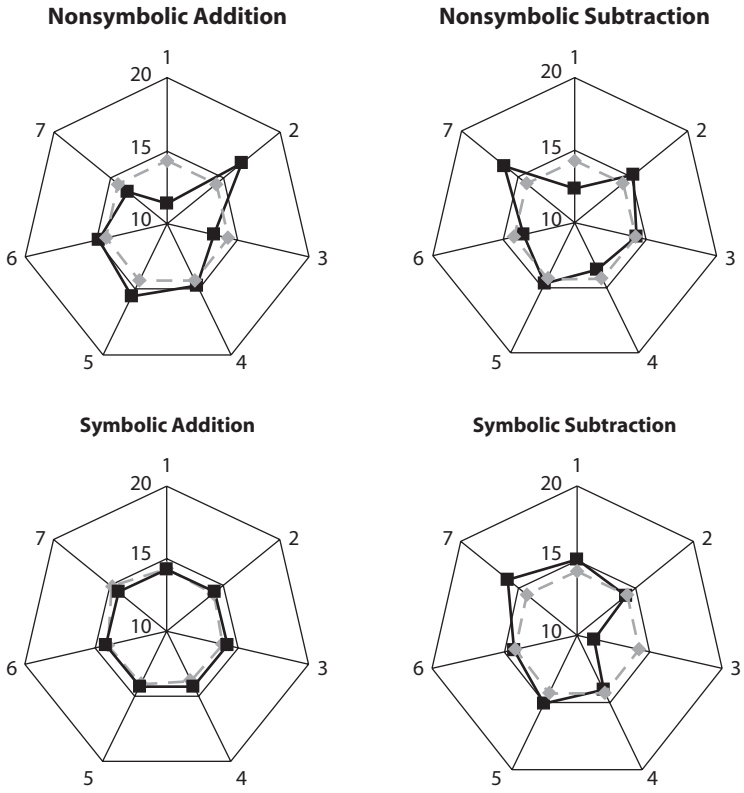

Figure 6. The space–operation association of responses effect. The graph shows the mean percentage with which a certain location was selected, separately for each notation and each operation. Gray dotted lines represent theoretical expectations under the assumption of perfect performance (100/7 ≈ 14%). The subjects were biased toward an upper right location for addition problems and toward the upper left location for subtraction problems, at least for nonsymbolic operations.

value closest to the correct result. Overall, the results are consistent with the notion of a cross-talk between spatial and arithmetic operations, as predicted by E. M. Hubbard et al. (2005). We propose to refer to this novel effect as the space–operation association of responses (SOAR) effect.

Although this finding may point to interactions between calculation and space, as predicted by E. M. Hubbard et al. (2005), the present experiment alone does not allow one to conclude whether this bias is truly due to the operation being carried out. An alternative possibility is that it arose from the numerical magnitude of the response choices presented on screen, because, in Experiment 1, the operands were identical for addition and subtraction and, as a consequence, the outcomes were systematically larger for addition than for subtraction. It is therefore possible that the subjects were spatially biased by the larger numerosities present on screen (a simple variant of the spatial numerical association of response codes [SNARC] effect [Dehaene, Bossini, & Giraux, 1993]), rather than by the preceding arithmetic operations.

In order to clarify this point, we conducted a second experiment very similar to the first one, but with constant outcomes. Instead of using identical operands for addition and subtraction, we manipulated the operands of the addition and subtraction problems so that their correct outcomes were identical. All other experimental details remained unchanged.

EXPERIMENT 2

Method

Subjects. Eighteen volunteers (15 female; mean age = 22.5 years, *SD* = 2.4 years; range, 19–26 years) took part in the experiment after having given their written informed consent. All but one were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971).

Stimuli. The basic stimulus set consisted of nine addition and nine subtraction problems. For addition (subtraction), the operands ranged from 14 to 56 (32 to 128) for the first operand and from 5 to 42 (7 to 59) for the second operand. As in Experiment 1, the second operand was a fixed proportion of the first operand: 28% (23%), 49% (34%), or 76% (44%). The correct results covered approximately the same numerical range as in Experiment 1 (21–98). Table 2 provides an overview of the problems used in Experiment 2.

All other details of creating the stimulus set were identical to those in Experiment 1.

Procedure. The procedure was identical to that in Experiment 1.

Results

The trimming procedure was identical to that in Experiment 1. Of all responses, 0.4% were excluded from further analyses.

Table 2
All Basic Arithmetic Problems Presented in Experiment 2 and Their Correct and Deviant Results

Operands		Results and Mean Deviants (Not Jittered)								
O1	O2	1/2	1/1.7	1/1.4	1/1.2	1/1	1.2/1	1.4/1	1.7/1	2/1
					Addition					
14	5	10	11	13	16	19	23	27	32	38
14	7	11	12	15	18	21	25	30	35	42
14	11	13	15	18	21	25	30	35	42	50
28	7	18	21	25	29	35	42	49	59	70
28	13	21	24	29	34	41	49	58	69	82
28	21	25	29	35	41	49	58	69	82	98
56	13	35	41	49	58	69	82	98	116	138
56	28	42	50	59	71	84	100	119	141	168
56	42	49	58	69	82	98	117	139	165	196
					Subtraction					
32	13	10	11	13	16	19	23	27	32	38
32	11	11	12	15	18	21	25	30	35	42
32	7	13	15	18	21	25	30	35	42	50
64	29	18	21	25	29	35	42	49	59	70
64	23	21	24	29	34	41	49	58	69	82
64	15	25	29	35	41	49	58	69	82	98
128	59	35	41	49	58	69	82	98	116	138
128	44	42	50	59	71	84	100	119	141	168
128	30	49	58	69	82	98	117	139	165	196

Note—The actual problems presented to the subjects were jittered by a small random amount (see the Method section), such that the correct outcome was never presented.

Nonrandom distribution of responses. Again, we first examined whether the subjects responded randomly to the choice screen only or genuinely took into account the arithmetic problems. In Figure 7, we plotted response frequency for each notation and each operation, separately for trials on which the correct answer was closest to the third or the fifth response alternative (see Table 2). An ANOVA over the different response categories, with percentage of choice as the dependent variable and rank of the subject's choice (one to seven) and trial type (third or fifth value correct) as factors, supported the notion that the subjects' choices were influenced by the arithmetic problem. For each operation in both notations, significant effects of rank and significant rank × range interactions were observed (all ps $< .001$).

Linear increase of response and response variability with the correct result. The subjects' mean responses (depicted as dots in Figure 8) again increased as a function of the correct outcome for both addition (black) and subtraction (gray) in both notations (Figure 8A; nonsymbolic notation; Figure 8B; symbolic notation). Four one-way repeated measures ANOVAs with result size as the only factor for nonsymbolic addition [$F(8,136) = 70.14$, $p < .001$, $\varepsilon = .39$] and subtraction [$F(8,136) = 664.60$, $p < .001$, $\varepsilon = .74$], as well as for symbolic addition [$F(8,136) = 4{,}135.88$, $p < .001$, $\varepsilon = .66$] and subtraction [$F(8,136) = 1{,}755.14$, $p < .001$, $\varepsilon = .50$] confirmed this impression.

Crucially, as can be seen in Figures 8 and 9, for nonsymbolic operations, although the correct results were now identical under addition and subtraction, the subjects' mean responses differed, a first clear indication of OM. The numbers selected were bigger under addition than under subtraction, and all the more so that the outcome increased. This resulted in a main effect of operation [$F(1,17) = 77.75$, $p < .001$] and an operation × result size interaction [$F(8,136) = 58.45$, $p < .001$, $\varepsilon = .27$] for nonsymbolic stimuli. For symbolic stimuli, no such main effect was found ($F < 1$), but the interaction [$F(8,136) = 3.53$, $p = .007$, $\varepsilon = .60$] indicated that the impact of the operation was not identical for all the problems.

An equivalent repeated measures ANOVA over the standard deviations of the chosen values (depicted as squares in Figure 8) showed that the subjects' choices became increasingly variable as the numerical size of the results increased. This was true for both addition (black) and subtraction (gray) with the nonsymbolic notation [Figure 8A; $F(8,136) = 3.91$, $p = .004$, $\varepsilon = .60$, and $F(8,136) = 29.85$, $p < .001$, $\varepsilon = .61$, respectively] and with the symbolic notation [Figure 8B; $F(8,136) = 22.27$, $p < .001$, $\varepsilon = .72$, and $F(8,136) = 14.64$, $p < .001$, $\varepsilon = .42$, respectively].

As is depicted in the lower parts of Figures 8A and 8B, the CV seemed to be constant or decreasing across the whole range of outcomes for addition and subtraction with both notations. Indeed, CV did not change as a function of the correct outcome for nonsymbolic subtraction [$F(8,136) = 1.66$, $p = .117$, $\varepsilon = .97$] or symbolic addition [$F(8,136) = 1.83$, $p = .091$, $\varepsilon = .83$]. However, it significantly increased with increasing problem size for nonsymbolic addition [$F(8,136) = 2.91$, $p = .015$, $\varepsilon = .68$] and significantly decreased for symbolic subtraction [$F(8,136) = 3.01$, $p = .009$, $\varepsilon = .76$], although both effects were quantitatively small.

Operational momentum. We again computed the difference between the log chosen and log correct outcome to obtain an index of the bias by which the subjects' choices deviated from the correct outcome (see Experi-

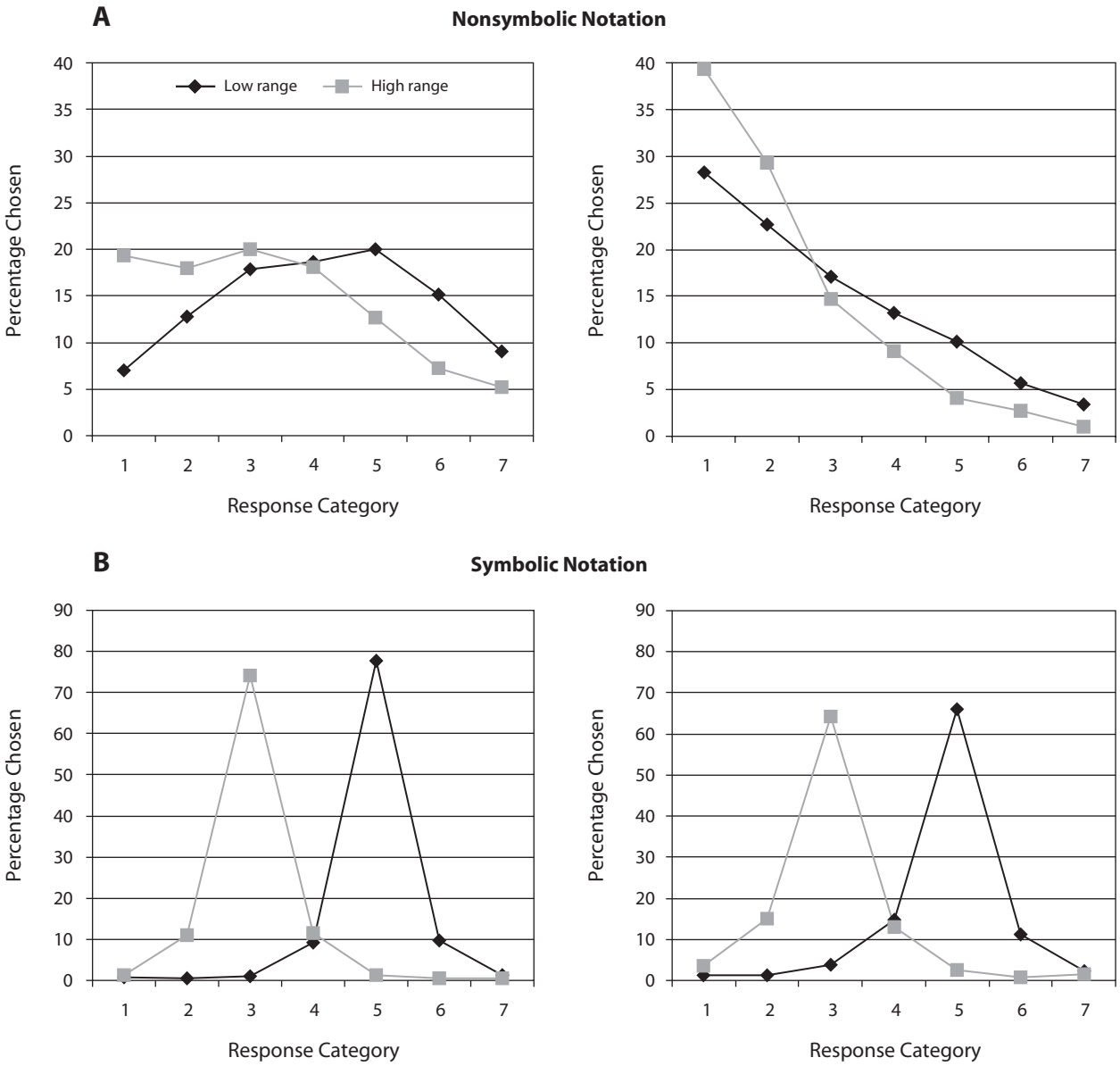

Figure 7. Distribution of the subjects' choices across the seven proposed results, averaged over all arithmetic problems, separately for each addition (left column) and subtraction (right column) problem in nonsymbolic (A) and symbolic (B) notation (for other information, see Figure 2).

ment 1). This measure of the OM effect is depicted in Figure 10 separately for each operation in both notations. To test the overall presence of an OM effect, we computed a 2 × 2 repeated measures ANOVA over the difference of the logarithm of the correct outcome and the log of the chosen value. The results resembled those from Experiment 1. Smaller values (indicating an underestimation) were obtained with nonsymbolic than with symbolic stimuli [$F(1,17) = 61.46$, $p < .001$]. Most important, the OM effect was reflected in smaller values for subtraction problems than for addition problems [$F(1,17) = 52.36$, $p < .001$]. A significant interaction [$F(1,17) = 45.87$, $p < .001$] indicated that the difference in response bias between addition and subtraction was larger for nonsymbolic than for symbolic problems, although it was significant in both cases [nonsymbolic, $t(17) = 7.11$, $p < .001$; symbolic, one-tailed $t(17) = 1.76$, $p = .048$].

Influence of the arithmetic operation on the spatial distribution of responses. Again, we tested whether subjects preferentially selected values presented at certain screen positions as a function of the arithmetic operation. Figure 11 shows that the overall pattern of performance was quite comparable to that in Experiment 1.

We statistically tested this impression by computing two 2 × 2 repeated measures ANOVAs with factors of operation and position for the upper left and upper right positions only (separately for nonsymbolic and symbolic stimuli). The interaction between location and operation,

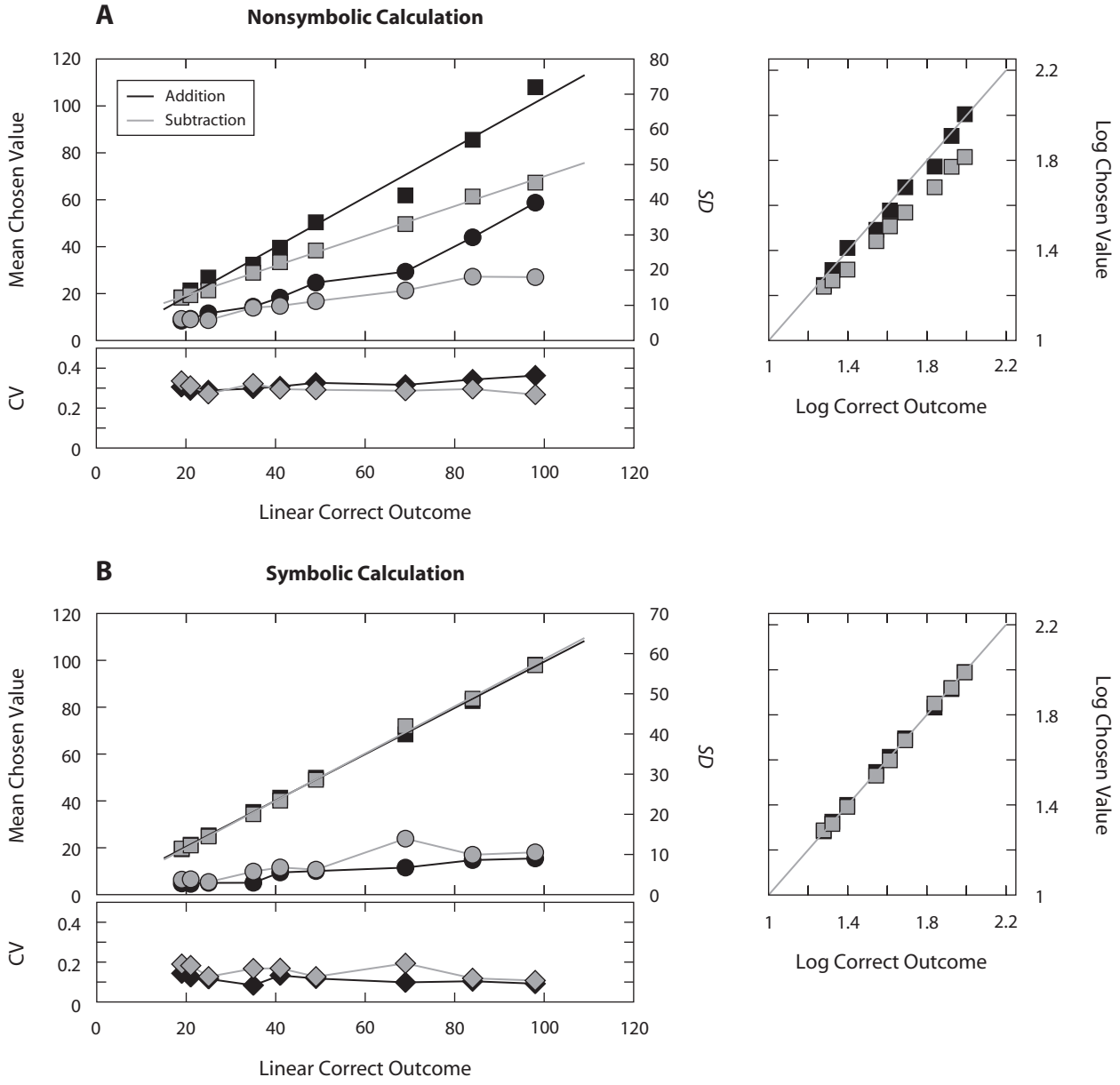

Figure 8. Left column: Mean responses (chosen values, squares) of the subjects and standard deviations (circles) plotted against the correct outcome for both nonsymbolic (A) and symbolic (B) addition and subtraction in Experiment 2. The lower part of panels A and B depicts the coefficient of variation (CV, diamonds) plotted against the correct outcome. Right column: The logarithm of the correct outcome plotted against the logarithm of the mean value chosen by the subjects for nonsymbolic (top) and symbolic (bottom) notation. The gray line indicates a ratio of 1—that is, perfect performance.

although in the same direction as in Experiment 1, was not significant [nonsymbolic notation, $F(1,17) = 2.26$, $p = .15$; symbolic notation, $F(1,17) < 1$]. This might be interpreted in terms of high intersubject variability preventing the effect from reaching significance in a test with subjects as the random factor. As in Experiment 1, we therefore also tested the deviation of the observed clicking preferences from the expected uniform distribution with a χ^2 test across trials. A 2×2 contingency table analysis showed that for nonsymbolic notation, the subjects did not choose each position on the screen equally often for the two operations ($p = .003$), whereas no such differential impact of operation on the spatial position chosen was present for symbolic notation ($p = .53$).

Note that in Experiment 2, contrary to Experiment 1, the choices presented on the screen had the same magnitudes for addition and subtraction. Thus, the results of Experiment 2 indicate that the spatial bias for results in certain locations on the screen was driven not only by the magnitude of the numerosities presented at choice time, but also, crucially, by the arithmetic operation that was carried out prior to the choice itself.

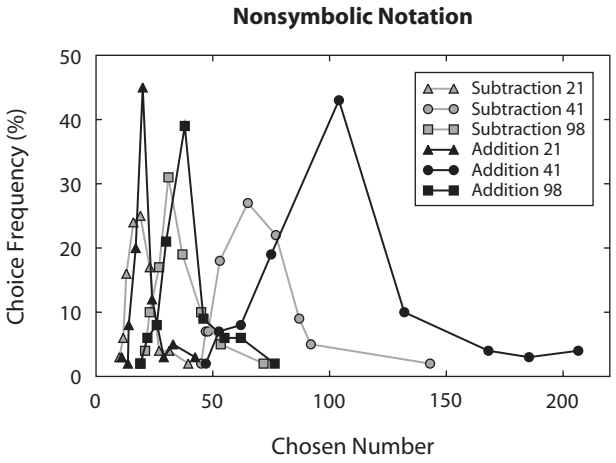

Figure 9. Sample distributions of the subjects' responses to three nonsymbolic addition and subtraction problems in Experiment 2. In each case, a small leftward displacement of subtraction, relative to addition problems, reflects the operational momentum effect.

Discussion

The second experiment was designed mainly to decide whether the SOAR effect (i.e., the spatial bias in response frequency as a function of the arithmetic operation) observed in Experiment 1 was due to the increasing magnitude of the outcomes or to the arithmetic operations that were deployed. Since we had used matched operands in Experiment 1, the outcomes were larger for addition than for subtraction, thus confounding these factors. In Experiment 2, we designed the operands so that the outcomes were identical (and therefore, the operands were smaller for addition than for subtraction). If the SOAR effect is due to the arithmetic operation, we should still observe it in this second experiment. Indeed, we replicated the spatial bias for the upper right and upper left positions of the screen as a function of the arithmetic operation. Again, this tendency was more pronounced for nonsymbolic notation than for symbolic arithmetic. Thus, we conclude that the SOAR effect is most probably due to the arithmetic operation, rather than to the numerical size of the outcome.

With regard to the OM effect, we also replicated the results from the first experiment. We observed a bias in the chosen outcomes as a function of both the notation and the operation. Most important, for equal objective outcomes, the subjects were biased toward smaller numbers for subtraction problems than for addition problems. As in Experiment 1, these effects were driven largely by the nonsymbolic notation. We still observed a significant OM effect with symbolic notation, albeit a very small one.

GENERAL DISCUSSION

We have reported the results from two experiments in which symbolic and nonsymbolic addition and subtraction were investigated in order to quantify the mechanisms underlying basic mental arithmetic. Two cognitive effects were investigated in more detail.

The OM effect refers to the fact that the subjects' responses were systematically biased by the arithmetic operation carried out. Responses were biased toward larger numbers for addition than for subtraction. Although this response bias was more pronounced for nonsymbolic problems, the present study proves for the first time that OM affects symbolic arithmetic. This aspect of our results strengthens the hypothesis that approximate arithmetic, even when the input numbers are presented as Arabic numerals, relies on magnitude representations and arithmetic procedures that are partially similar to those used for nonsymbolic calculation (Dehaene et al., 1999).

The second, new effect reflects a spatial bias of the subjects for preferring some locations on the screen over others, depending on the arithmetic operation that they just performed. With addition, the subjects preferentially selected numerosities displayed in the upper right location, whereas for subtraction, they preferred the upper left location. Experiment 2 showed that this effect occurs, although with high variability across subjects, even when addition and subtraction problems that yield the same numerical outcome are compared. Note that, in this case, the initial operands are *larger* for subtraction than for addition, and yet the observed effect associates subtraction with the *left* side of space. Thus, this new association between arithmetic operations and the left–right axis can be differentiated from the classical SNARC effect (Dehaene et al., 1993), whereby increasing number size causes an increasingly larger rightward spatial bias. Since arithmetic operations are associated with a distinct spatial bias, we coined the term *SOAR* (as an acronym for space–operation association of responses) for this new effect.

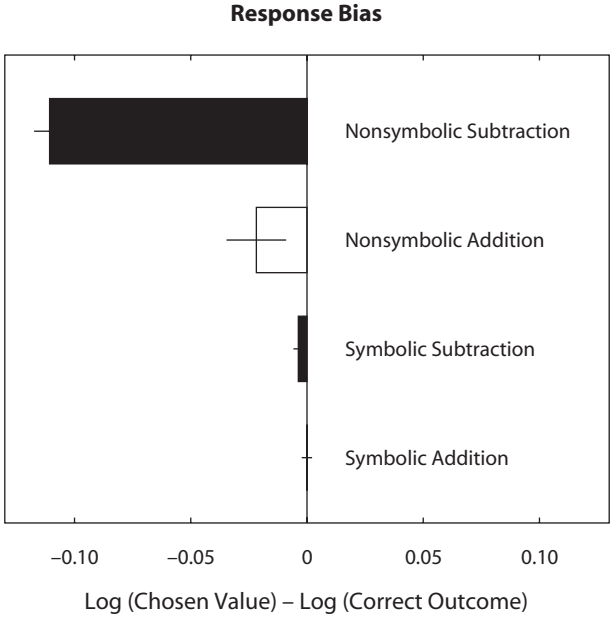

Figure 10. Mean response bias, defined as the difference between the chosen value and the correct outcome, both expressed on a log scale. A negative bias indicates underestimation, and a positive bias indicates overestimation.

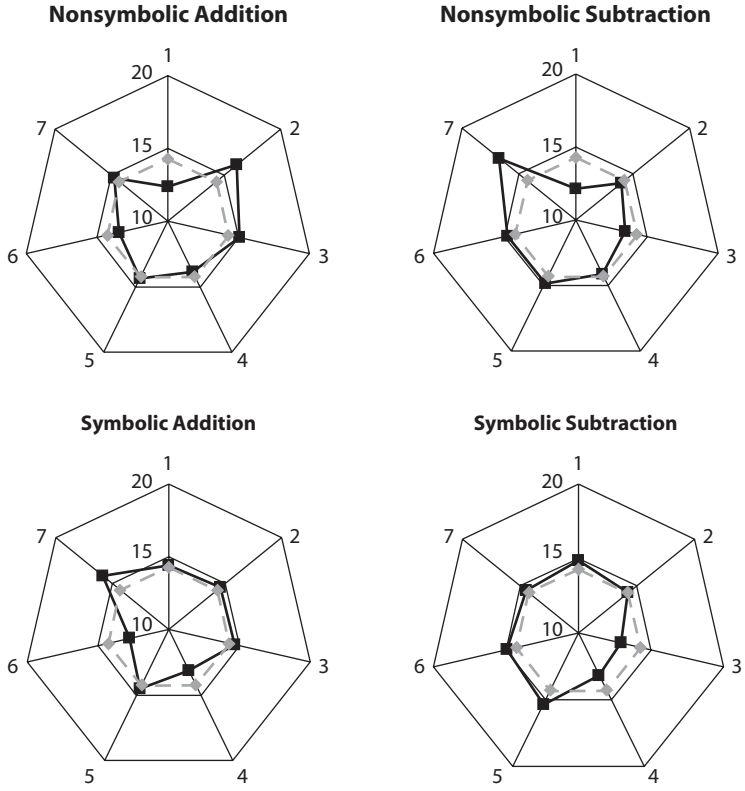

Figure 11. The space–operation association of responses effect for Experiment 2. Again, the subjects were clearly biased toward an upper right response to addition problems and an upper left response to subtraction problems for nonsymbolic operations. An unexpected bias toward upper left positions was observed for symbolic addition problems.

The SOAR effect may be interpreted in a framework where calculation is likened to a sort of movement or a shift of attention along the mental number line (Dehaene, 1992; E. M. Hubbard et al., 2005; Restle, 1970). This interpretation parallels the account of perceptual *representational momentum* in terms of dynamic mental representations (Freyd & Pantzer, 1995). Just as Freyd and Pantzer assumed that the mental representation of a moving object is permanently updated, the mental representation of a given numerosity in the course of being transformed by a mental calculation would be not fixed but dynamic, with a displacement in the distribution of activation on the mental number line representing the change from one value to another. The direction of change (addition/subtraction) would then determine the OM effect and, by congruity with space, the SOAR effect. Both the OM and SOAR effects add to previous evidence for number–space interactions arising from the SNARC effect and the numerical bisection task. They suggest that numerosity is internally mapped onto a spatially organized and dynamically updated mental representation and that this representation is not activated just in an epiphenomenal manner but is actively used and updated during mental arithmetic—at least when approximations are required.

Even with symbolic notation, we observed a differential bias due to the arithmetic operation in which the subjects were engaged (see Pinhas & Fischer, 2008, for similar results). The effect is no doubt very small and on the edge of detectability, but this should not be surprising, given that responses are so much more precise when the operation is presented in symbolic form and thus leave little room for errors and biases to emerge. The very fact that we observe an effect despite the use of verbally mediated, exact calculation lets us assume the presence of common nonverbal and, presumably, attentional mechanisms. It remains to be seen whether exact arithmetic, including more complex operations such as multiplication and division, would also yield such spatial biases. Under the view that exact arithmetic relies on a distinct system, based on a verbal or symbolic coding of the numbers involved (Dehaene et al., 1999), one might predict that no SOAR effect should be found. However, if the SOAR effect is determined solely by a representation of the magnitude of the final result on an internal number line, it might still be presented during complex exact arithmetic.

Over and above the biases perceptible in the OM and SOAR effects, the results of the present experiments indicate that, in the mean, the subjects were rather accurate in performing basic mental arithmetic, using nonsymbolic numerosities. On average, the subjects tended to underestimate the results of nonsymbolic problems, with addition being rather precise and subtraction largely underesti-

mated. It has been shown in previous studies that subjects generally underestimate the number of dots present in a given set of objects in transcoding tasks (e.g., from sets of dots to Arabic numerals; see Izard & Dehaene, 2008). Note, however, that the involvement of a general tendency to underestimate the number of items in a given set remains putative for the moment, since in theory no transcoding from one notation (e.g., nonsymbolic) to another (e.g., symbolic) is necessarily involved in the present paradigm. Nevertheless, subjects might engage in some sort of internal labeling of quantities that is then apt to undergo the same type of bias. The range of numbers used in the present study might as well have contributed to an overall bias to underestimate outcomes of nonsymbolic problems (yet, still resulting in larger values for addition than for subtraction). Here, we used larger numerosities than did McCrink et al. (2007), for example. Little is known so far about nonsymbolic calculation performance in the numerosity range we used in the present experiments.

We found that Weber's law clearly holds for both symbolic and nonsymbolic arithmetic, in agreement with previous results (Barth, Kanwisher, & Spelke, 2003; McCrink et al., 2007). The roughly constant coefficient of variation over different magnitudes suggests that a single underlying representation may be accessed whenever approximate arithmetic operations are carried out, whether in symbolic or nonsymbolic format. These results fit with those of other studies showing that basic numerical abilities are not restricted to humans but seem to be shared with other species (Beran, 2007). An important difference, however, is that humans possess symbolic codes for numbers, which, as was observed here, give them access to a much higher precision in calculation. In neural terms, it has been suggested that this effect could be explained by a sharpening of the tuning of number-coding neurons in the course of symbol acquisition (Verguts & Fias, 2004).

Alternative Strategies: Did the Subjects Engage in Mental Calculation?

The interpretation of the present findings relies on the idea that the present results really do give evidence for an arithmetical process of approximate calculation. Thus, we have to rule out the use of simpler response strategies.

An alternative interpretation of our results appeals to the anchoring and adjustment mechanisms put forward by Tversky and Kahneman (1974). The first operand might serve as an anchor (i.e., a numerical magnitude that subjects focus on), whereas the second operand would be used to "adjust" (i.e., to change the initial focus along the mental magnitude representation). We think that this anchor-and-adjust mechanism is not necessarily incompatible with our notion of approximate arithmetic but that the precise nature of our data requires a much more precise specification of this mechanism. As can be seen in Figures 3 and 8, the subjects' responses track very closely the correct exact arithmetic solution of each problem. Thus, the recombination of information gathered from each operand follows closely the laws of arithmetic (although with small spatial biases). The required subprocesses seem to go well beyond a mere *heuristic*, although they do not preclude its overall involvement.

Another objection might be that the subjects counted the dots to produce their responses. We analyzed the reaction times and did not observe any strong evidence that would imply a counting strategy. If the subjects had adopted a counting strategy, the reaction time should be linearly and strongly related to the chosen number in a given problem. However, the correlation between the numerical size of the chosen value (which, in turn, is linearly related to the different response alternatives) was $r = .08$ in Experiment 1 ($p = .001$) and $r = .03$ in Experiment 2 (n.s.). Two other observations make it rather unlikely that the subjects used a counting strategy. If we assume that the subjects can count individual items at a rate of about 250 msec per item (a time at the lower boundary of fixation times in simple visual search paradigms; cf. Hooge & Erkelens, 1998; Jacobs, 1986), the presentation time of 1,500 msec for the operands did not allow for counting more than six items in a given set of dots. This is way below the numerical range of the operands used here. A similar argument holds for the reaction times in response to the presentation of the response alternatives, which was theoretically unlimited in time. The median response times were 2,712 msec for addition and 3,011 msec for subtraction problems in Experiment 1, and 2,859 and 2,672 msec for addition and subtraction, respectively, in Experiment 2. Again, this is not long enough to count the number of elements for the numerosities used here, especially given that the subjects had to choose between seven response alternatives.

In combination with the fact that neither the subjects nor the experimenter reported the use of a counting strategy, we conclude that counting did not contribute to the pattern of results in the present study.

Beyond these alternative strategies, the subjects might have produced their responses on the basis of the individual item sizes, which covaried with numerical magnitude. If the subjects' performance was predominantly influenced by the individual item size, we would expect major differences between Experiment 1, in which the item size of the addition and subtraction problems differed due to different numerical magnitudes of the outcomes, and Experiment 2, in which the response alternatives were identical for both operations. The pattern of results in both experiments, however, was largely identical (see Figures 5 and 10), including a significant underestimation for both addition and subtraction problems, which was more pronounced for subtraction. This is not congruent with the idea that the results were due to the subjects' paying attention to individual item size.

Finally, one might wonder whether the subjects responded solely on the basis of the range of the presented response alternatives or on the basis of the size of one of the operands, without combining them in a mental calculation process. We carefully analyzed our data with respect to these and other alternative strategies. None of them could explain the present pattern of results (see the online supplemental materials).

In sum, the most plausible strategy seems to involve arithmetic processing, for both symbolic and nonsymbolic

quantities. The exact mechanism by which internal quantities are manipulated to support simple calculation remains unknown, however.

In the present study, we replicated the OM effect with numbers presented in nonsymbolic and symbolic notation. This allowed us to further analyze one putative determinant of the effect—that is, whether the OM effect is influenced by the numerical size of the operands (see the supplemental materials). For nonsymbolic notation, the results can be quickly summarized by stating that, for both addition and subtraction, whichever factor increased the numerical value of the outcome also increased the proportional size of the OM effect.

For numbers presented in symbolic notation (Arabic numerals), the subjects were much more precise in their judgments, suggesting either that they could not inhibit exact calculation or that they have an inherently more precise, but still analogue, representation of the quantity associated with Arabic numerals (Dehaene et al., 1999; Verguts & Fias, 2004). Both possibilities are likely to be correct. The second possibility is vindicated by the finding of Weber's law, but also of a smaller OM effect. The small size of the SOAR and OM effects in symbolic notation was probably due to the subjects' higher precision, which frequently led to the optimal choice and thus left little room for observing spatial or numerical errors and biases. Yet the fact that the observed effects were, overall, larger for nonsymbolic than for symbolic notation also suggests the involvement of partially different subprocesses in the course of solving these two kinds of problems. It seems very likely that the subjects engaged in some exact calculations when facing a symbolic arithmetic problem, plausibly involving memory recall of verbal representations and of rote arithmetic facts from long-term memory. With nonsymbolic stimuli, no such knowledge was available. Because of this major difference, it is all the more noteworthy that we did observe small yet significant indications of an OM effect even with exact and, presumably, verbally mediated calculation. In future work, a suggestion might be to use a narrower range of proposed outcomes for symbolic than for nonsymbolic stimuli, thus compensating for the subjects' higher precision. This method may have a greater chance of detecting the small biases that may characterize symbolic arithmetic.

Commonalities and Differences Between Operational and Representational Momentum

The OM effect was named by analogy with the representational momentum effect in the visual domain. This was meant to indicate that in both domains—numerical and visual—mechanisms may exist that operate on and involve the updating of a dynamic mental representation (be it of moving objects or of numerical magnitudes). Freyd (1993) postulated that the representational momentum effect arises from a nonstatic but dynamic underlying representation that is continuous and analogue in nature. Note, however, that in the present experiments, the input stimuli were discrete in nature (natural numbers, either in the form of dots in a set or as Arabic numerals). Nevertheless, many experiments suggest that the underlying representation of these numbers relies on analogue mental numerical magnitudes. When saying *analogue*, we mean that the internal representation of numerical magnitude is noisy and organized along a continuum in a similar way as are other (physical) quantities. As has been discussed in many studies since Moyer and Landauer's (1967) seminal finding, the brain represents discrete numbers by using continuous internal quantities (Dehaene et al., 1999; Gallistel & Gelman, 1992; Shepard, Kilpatrick, & Cunningham, 1975).

Whether the analogy between *physical* and *operational* momentum is equally applicable to all aspects and theoretical dimensions of both effects remains to be seen. For instance, in the domain of physical motion, representational momentum typically applies only if the perceived deviations in location occur in the direction of the movement. In our understanding, however, this is one of the points where the analogy between representational momentum and OM might have its limits. For the domain of numerical cognition, the most interesting analysis is probably not to look directly at whether the selected result exceeds the correct value or not, but rather to contrast directly addition and subtraction problems that are carefully matched either by operands (Experiment 1) or by outcome (Experiment 2). These analyses are useful in order to control for any overall bias that might affect the subjects' responses, particularly since it is well known that numerosity is frequently underestimated (e.g., Izard & Dehaene, 2008). These comparisons clearly indicate that under matched conditions, additions are overestimated, as compared with subtractions—even though addition estimates do not necessarily exceed the correct result.

It is also uncertain whether the underlying mechanisms are identical for both effects (i.e., representational momentum and OM). Following E. M. Hubbard et al. (2005) and Dehaene and colleagues (2003), we tentatively propose that OM may have its neural origin in parietal attentional mechanisms that operate on the internal mental number representation. These may share certain features with mechanisms involved in the phenomenon of representational momentum, such as a process of constant updating, but they may also differ with respect to other features and might even be dissociable by distinct brain lesions.

In fact, we observed two effects of "spatial" displacement in the present study. The OM effect suggests a displacement on the mental number line when a particular result is chosen. The SOAR effect describes a preference for certain positions on a screen as a function of the arithmetic process the subject is engaged in. In the domain of representational momentum, too, it has been shown that the final perceived spatial displacement reflects the influence of several sources of bias. In particular, E. M. Hubbard and colleagues demonstrated that the representational momentum effect decreases when objects appear to slide along a surface, thus indicating that humans incorporate and represent friction when estimating the trajectory of an object (T. L. Hubbard, 1995). In a similar vein, representational gravity and the presence of a landmark in the visual scene combine with other factors, such as velocity, to determine the final

spatial displacement of the subjects' judgments (T. L. Hubbard & Ruppel, 2000). Bridging the gap between cognitive domains, one might thus infer that OM, too, is a combination of several sources or forces—notably, the mere displacement along the mental number line and the SOAR effect. In the context of the present experiments, it is not possible, however, to disentangle these factors. By choosing a particular response alternative on a screen, the subjects simultaneously chose both a numerosity and a screen location. It remains to be seen in future experiments how these two biases interact to determine the final performance. To generalize this idea, it might be interesting in future experiments to investigate what sources of variability contributed to the OM effect. This question aims at differentiating whether the observed OM effect has its origin in the mental representation of the displayed numerosity, the mental calculation process, or the response selection stage (Cordes et al., 2007). In the study of Cordes and colleagues the major determinant of variability in nonverbal arithmetic was found to have its origin in the individual mental numerical magnitude representations entering the calculation process, rather than from mental transition or motor production processes. This is interesting in the context of the present experiments, since here we observed an additional source of variability that contributes to the outcome of a mental arithmetic operation—that is, the OM effect. At the same time, it might help to understand the results of the present experiments: In contrast to the procedure of Cordes and colleagues, in the present study, the subjects did not have to repeatedly push a button to indicate a certain numerosity but just could click on one of the quantities displayed on the screen. Since this procedure is less apt to be influenced by covarying factors such as the duration of the buttonpresses (the larger the numerosities, the longer it takes for subjects to respond), we may assume that the amount of variability introduced by the response stage is reduced. Therefore, the present methodological approach may be useful in future experiments to further delineate the laws that characterize nonverbal calculation.

AUTHOR NOTE

This work was first reported as an oral presentation at the workshop "Number, Space, and Action" (Aachen, Germany, March 23–24, 2007). It was supported by a grant from the Volkswagen Foundation to S.D., in support of A.K. We thank Charlie Chubb in his role as action editor and Timothy Hubbard and three anonymous reviewers for valuable suggestions on earlier versions of the manuscript. Correspondence concerning this article should be addressed to A. Knops, INSERM–CEA Cognitive Neuroimaging Unit, CEA/SAC/DSV/DRM/NeuroSpin, Bât 145, Point Courrier 156, F-91191 Gif-sur-Yvette, France (e-mail: knops.andre@gmail.com).

REFERENCES

BARTH, H., KANWISHER, N., & SPELKE, E. (2003). The construction of large number representations in adults. *Cognition*, **86**, 201-221.

BARTH, H., LA MONT, K., LIPTON, J., DEHAENE, S., KANWISHER, N., & SPELKE, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition*, **98**, 199-222.

BERAN, M. J. (2007). Rhesus monkeys (*Macaca mulatta*) enumerate large and small sequentially presented sets of items using analog numerical representations. *Journal of Experimental Psychology: Animal Behavior Processes*, **33**, 42-54.

BRANNON, E. M., & TERRACE, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, **282**, 746-749.

BRANNON, E. M., & TERRACE, H. S. (2000). Representation of the numerosities 1–9 by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, **26**, 31-49.

CANTLON, J. F., & BRANNON, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, **17**, 401-406.

CORDES, S., GALLISTEL, C. R., GELMAN, R., & LATHAM, P. (2007). Nonverbal arithmetic in humans: Light from noise. *Perception & Psychophysics*, **69**, 1185-1203.

DEHAENE, S. (1992). Varieties of numerical abilities. *Cognition*, **44**, 1-42.

DEHAENE, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

DEHAENE, S., BOSSINI, S., & GIRAUX, P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology: General*, **122**, 371-396.

DEHAENE, S., & COHEN, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*, **33**, 219-250.

DEHAENE, S., PIAZZA, M., PINEL, P., & COHEN, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, **20**, 487-506.

DEHAENE, S., SPELKE, E., PINEL, P., STANESCU, R., & TSIVKIN, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, **284**, 970-974.

FISCHER, M. H., WARLOP, N., HILL, R. L., & FIAS, W. (2004). Oculomotor bias induced by number perception. *Experimental Psychology*, **51**, 91-97.

FREYD, J. (1993). Five hunches about perceptual processes and dynamic representations. In D. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 99-119). Cambridge, MA: MIT Press.

FREYD, J., & FINKE, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 126-132.

FREYD, J., & FINKE, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society*, **26**, 443-446.

FREYD, J., & PANTZER, T. M. (1995). Static patterns moving in the mind. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 181-204). Cambridge, MA: MIT Press.

GALLISTEL, C. R., & GELMAN, R. (1992). Preverbal and verbal counting and computation. *Cognition*, **44**, 43-74.

GILMORE, C. K., MCCARTHY, S. E., & SPELKE, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, **447**, 589-591.

HALPERN, A. R., & KELLY, M. H. (1993). Memory biases in left versus right implied motion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 471-484.

HOOGE, I. T., & ERKELENS, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, **38**, 1295-1302.

HUBBARD, E. M., PIAZZA, M., PINEL, P., & DEHAENE, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, **6**, 435-448.

HUBBARD, T. L. (1995). Cognitive representation of motion: Evidence for friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 241-254.

HUBBARD, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, **12**, 822-851.

HUBBARD, T. L., & RUPPEL, S. E. (2000). Spatial memory averaging, the landmark attraction effect, and representational gravity. *Psychological Research*, **64**, 41-55.

HUYNH, H., & FELDT, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, **1**, 69-82.

IZARD, V., & DEHAENE, S. (2008). Calibrating the mental number line. *Cognition*, **106**, 1221-1247.

JACOBS, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Perception & Psychophysics*, **39**, 47-58.
KERZEL, D. (2003). Attention maintains mental extrapolation of target position: Irrelevant distractors eliminate forward displacement after implied motion. *Cognition*, **88**, 109-131.
KRUEGER, L. E. (1984). Perceived numerosity: A comparison of magnitude production, magnitude estimation, and discrimination judgments. *Perception & Psychophysics*, **35**, 536-542.
MCCRINK, K., DEHAENE, S., & DEHAENE-LAMBERTZ, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics*, **69**, 1324-1333.
MCCRINK, K., & WYNN, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, **15**, 776-781.
MOYER, R. S., & LANDAUER, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, **215**, 1519-1520.
NIEDER, A., FREEDMAN, D. J., & MILLER, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, **297**, 1708-1711.
NIEDER, A., & MILLER, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences*, **101**, 7457-7462.
OLDFIELD, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, **9**, 97-113.
PIAZZA, M., IZARD, V., PINEL, P., LE BIHAN, D., & DEHAENE, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, **44**, 547-555.
PIAZZA, M., PINEL, P., LE BIHAN, D., & DEHAENE, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, **53**, 293-305.
PICA, P., LEMER, C., IZARD, V., & DEHAENE, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, **306**, 499-503.
PINHAS, M., & FISCHER, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition*, **109**, 408-415.
RESTLE, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, **91**, 191-205.
SHELIGA, B. M., RIGGIO, L., & RIZZOLATTI, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research*, **98**, 507-522.
SHEPARD, R. N., KILPATRICK, D. W., & CUNNINGHAM, J. P. (1975). The internal representation of numbers. *Cognitive Psychology*, **7**, 82-183.
TVERSKY, A., & KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124-1131.
VERGUTS, T., & FIAS, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience*, **16**, 1493-1504.
XU, F., & SPELKE, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, **74**, B1-B11.

NOTE

1. One might argue that the ANOVA is not justified since the required prerequisite of homoscedacity is not met (the variance of the mean chosen value increases with increasing correct outcome). Note, however, that we report statistics that are corrected for nonsphericity, being a special case of heteroscedacity.

To protect us against the violation of the sphericity assumption, we additionally performed a log-transformation on the data and subsequently repeated the ANOVAs described here. We obtained identical results—that is, a significant increase of mean chosen value with increasing numerical size of the correct outcome in all cases (all $ps < .01$).

SUPPLEMENTAL MATERIALS

A discussion of alternative strategies and additional analyses of the determinants of the operational momentum effect may be downloaded from app.psychonomic-journals.org/content/supplemental.

(Manuscript received March 18, 2008;
revision accepted for publication December 2, 2008.)