Strategy variability in numerosity comparison task: a study in young and older adults

Angélique Roquet, Patrick Lemaire

Research Article

Angélique Roquet, Patrick Lemaire*

Strategy variability in numerosity comparison task: a study in young and older adults

https://doi.org/10.1515/psych-2018-0011
Received July 29, 2018; accepted December 9, 2018

Abstract: We investigated strategies used by young and older adults in dot comparison tasks to further our understanding of mechanisms underlying numerosity discrimination and age-related differences therein. The participants were shown a series of two dot collections and asked to select the largest collection. Analyses of verbal protocols collected on each trial, solution times, and percentages of errors documented the strategy repertoire and strategy distribution in young and older adults. Based on visual features of dot collections, both young and older adults used a set of 9 strategies and selected strategies on a trial-by-trial basis. The findings also documented age-related differences (i.e., strategy preferences) and similarities (e.g., number of strategies used by individuals) in strategies and performance. Strategy variability found here has important implications for understanding numerosity comparison and contrasts with previous findings suggesting that participants use a single strategy when they compare dot collections.

Keywords: strategic variations, aging, numerosity comparison, dot comparison task, congruency effects

How do participants estimate numerosities (i.e., find the approximate number of elements in a collection without exactly counting them)? Numerosity comparison is accomplished in many everyday situations (e.g., when estimating how many people are waiting in a queue, how much time we need to go from home to work, when estimating whether the available place to park our car is large enough). It is also an important skill that is suggested to underlie a number of other mathematics skills (e.g., Chen & Li, 2014). The present study adopted a strategy perspective to further our understanding of mechanisms underlying numerosity comparison. Examining strategies helps to better describe and explain (in mechanistic terms) how participants accomplish numerosity comparison tasks. We examined strategy use on a trial-by-trial basis while participants accomplished a widely used task in previous research on numerosity comparison, namely a dot comparison task. Strategy variability found here has important implications for understanding numerosity comparison and it contrasts with previous findings suggesting that participants use a single strategy when they compare dot collections. Before outlining the logic of the present work, we review previous findings in numerosity comparison and age-related differences therein, as well as strategic variations during aging.

Article note: This article is a part of Topical Issue "Strategy Contributions to Cognitive Aging" edited by Beatrice G. Kuhlmann.

***Corresponding author: Patrick Lemaire,** CNRS & Aix-Marseille Université, 3 Place Victor Hugo, 13331 Marseille France, Email: patrick.lemaire@univ-amu.fr
Angélique Roquet, Aix-Marseille Université

Previous findings on numerosity comparison.

To determine how participants estimate numerosities, two types of tasks have been used: estimation and comparison tasks. In estimation tasks, participants are shown collections of dots and have to determine how many dots each collection approximately contains (e.g., Crites, 1992; Gandini, Lemaire & Dufau, 2008; Lemaire, Lecacheur, & Farioli, 2000; Luwel, Lemaire, & Verschaffel, 2005; Luwel, Verschaffel, Onghena, & De Corte, 2003a, 2003b; Siegel, Goldsmith, & Madson, 1982). In comparison tasks, participants are given a series of two dot collections and have to indicate which collection they believe has more dots (e.g., Clayton & Gilmore, 2015; DeWind & Brannon, 2016; Halberda, Mazzocco, & Feigenson, 2008; Gebuis & Reynvoet, 2011, 2012; Smets, Gebuis, Defever, & Reynvoet, 2014). Previous research based on these tasks has shown that numerosity comparison and estimation performance is influenced by a number of factors, including a stimulus, situation, and participants' characteristics (Cappelletti et al., 2014, Gandini, Lemaire, & Dufau, 2008; Gandini, Lemaire, & Michel, 2009; Gilmore et al., 2016; Halberda et al., 2012; Inglis & Gilmore, 2013; Lemaire & Lecacheur, 2007; Li et al., 2010; Norris, McGeown, Guerrini, & Castronovo, 2015 ; Price, Palmer, Battista, & Ansari, 2012; Trick, Enns, & Brodeur, 1996; Watson, Maylor, & Bruce, 2005; Watson, Maylor, & Manson, 2002). Most of the previous works concerning the influence of non-numerical factors have examined numerosity estimation tasks rather than numerosity comparison tasks. This is why the focus of this study is on the numerosity comparison task.

There are two types of stimulus characteristics crucially influence numerosity comparison performance: numerical and visual features of dot collections. Regarding numerical properties of dot collections, previous studies found that participants are faster and more accurate when they are asked to compare smaller relative to larger collections of dots (e.g., Clayton & Gilmore, 2015; Revkin et al., 2008) and to estimate collections with larger ratios than on smaller ratios collections (i.e., Ansari et al., 2005; Barth, La Mont, Lipton, & Spelke, 2005; Pica, Lemer, Izard, & Dehaene, 2004). Thus, participants are faster when they compare small collections of 1 and 4 dots than large collections including 10 and 40 dots. They are also faster at larger-ratio collections (e.g., comparing collections of 24 and 12 dots) than at smaller-ratio collections (e.g., comparing collections of 24 and 18 dots). Such findings have been explained as resulting from functioning characteristics of an "approximate number system" (ANS) or "number sense" (Dehaene 1997, p. 5). The ANS is a universal system present in animals, children, and adults that allows comparison, addition, and subtraction of quantities without counting them. Previous studies have shown that ANS performance is dependent on the ratio between the quantities to be compared. Indeed, participants are faster comparing smaller than larger collections and larger-ratio (i.e., 8 vs. 16 dots) than smaller-ratio collections (i.e., 8 vs. 10 dots; Emmerton 1998; Hauser et al. 2003; Pica et al., 2004).

Participants' performance is also influenced by visual features of stimuli when they compare two collections of dots, such as cumulative surface area covered by the dots, convex hull (i.e., smallest contour around the dot collection), average dot size, density of dots, distance between dots, etc. For example, participants are faster at collections with larger convex hull than at collections with smaller convex hull (e.g., Clayton, Gilmore, & Inglis, 2015; DeWind & Brannon, 2016; Gebuis & Reynvoet, 2012; Gilmore, Cragg, Hogan, & Inglis, 2016; Smets et al., 2015) or at collections of dots that occupy a larger area on computer screen than at collections displayed with smaller areas (e.g., Clayton & Gilmore, 2015; Gebuis & Gevers, 2011; Gebuis & Reynvoet, 2011, 2012; Halberda, Mazzocco, & Feigenson, 2008; Pica et al., 2004; Libertus, Feigenson, & Halberda, 2011; Mazzocco, Feigenson,& Halberda, 2011).

The main objective of the present study is the interactions between numerical and visual features. For example, when participants are asked to select the collection of dots with the largest number of dots, they are faster at congruent items, when both numerical and visual features match (e.g., larger numerosity displayed with a larger convex hull) than at incongruent items, when both features mismatch (e.g., larger numerosity displayed with a smaller convex hull). Such congruency effects have been found in numerous studies (Barth et al., 2006; Cappelletti, Didino, Stoianov, & Zorzi, 2014; Clayton et al., 2015; Clayton & Gilmore, 2015; Gebuis & Reynvoet, 2012; Gilmore et al., 2013, 2016; Gilmore, Keeble, Richardson, & Cragg, 2015; Halberda et al., 2008; Hinault & Lemaire, 2016 a & b; Hinault, Dufau, & Lemaire, 2014; Hinault, Lemaire, & Philips, 2016; Hinault, Badier, Baillet, & Lemaire, 2017; Lemaire & Hinault, 2014; Inglis &

Gilmore, 2014; Fush & McNeil, 2013; Nys & Content, 2012). They suggest that dot comparison tasks involve domain-general mechanisms, such as inhibition. Indeed, participants have to ignore irrelevant dimensions (i.e., visual features) and focus on the relevant dimension (i.e., the numerosity) on incongruent items.

In addition to numerical and visual features of dot collections, participants' performance in dot comparison is also influenced by task or situation characteristics, such as the method of presentation (e.g., sequential vs. simultaneous; e.g., Price, Palmer, Battista, & Ansari, 2012), set size (e.g., Barth, Kanwisher, & Spelke, 2002; De Smedt et al., 2013; Revkin et al., 2008); or the display time (e.g., Inglis & Gilmore, 2013). For example, participants are faster when they compare two collections of dots displayed separately on a computer screen than when the two collections of dots are displayed in different colors but intermixed (e.g., Price, Palmer, Battista, & Ansari, 2012). They are also more accurate when collections of dots are displayed for longer than for shorter durations (e.g., Inglis & Gilmore, 2013; Gilmore et al., 2016).

Finally, performance in numerosity comparison has been found to either change with participants' age during adulthood (e.g., Halberda et al., 2012; Li et al., 2010; Cappelletti et al., 2014; Norris, McGeown, Guerrini, & Castronovo, 2015; Trick, Enns, & Brodeur, 1996) or remain age-invariant (e.g., Gandini, Lemaire, & Dufau, 2008; Gandini, Lemaire, & Michel, 2009; Lemaire & Lecacheur, 2007; Watson, Maylor, & Bruce, 2005; Watson, Maylor, & Manson, 2002). For example, Lemaire and Lecacheur (2007) asked young and older adults to estimate numerosity of dot collections that included 40—460 dots. Although the results showed similar accuracy in young and older adults, older adults made more numerous but shorter eye fixations than young adults. In a large-scale study, Halberda et al. (2012) asked more than 10,000 participants ranging from 11 to 85 years of age to accomplish a dot comparison task. They found that the ability to discriminate two numerosities decreases with age after age 30. Also, Cappelletti et al. (2014) asked young and older adults to compare collections of dots that included between 5 and 16 dots. Collections were congruent (i.e., larger collections had a larger cumulative area, and smaller collections were displayed with a smaller cumulative area) or incongruent (i.e., larger collections had a smaller cumulative area, and smaller collections were displayed with a larger cumulative area). The researchers found that older adults' performance was poorer than young adults' in incongruent items but both groups performed equally well in congruent items. This result indicates that older adults were specifically impaired on incongruent items where it is necessary to inhibit the irrelevant information (i.e., cumulative area) to focus on relevant information (i.e., numerosity). Cappelletti et al. (2014) proposed that age-related declines on incongruent items result from declined efficiency of inhibition mechanisms (see also Norris et al., 2015). These finding suggests that one of the potential sources of age-related differences in numerosity comparison performance may be age-related changes in general cognitive mechanisms (such as executive control).

In summary, previous studies showed that to select the more numerous of two dot collections, participants base their smaller/larger responses on the number of dots in each visual collection as well as on visual characteristics of stimuli. Participants use visual features of dot collections because one or several visual characteristics correlate with numerosity in each collection of dots. For example, larger collections of dots occupy larger dot area, and have a smaller dot density, a larger distance between dots, a larger convex hull, etc. Previous studies, however, suggest that convex hull may be the most crucial visual feature used by participants to make their smaller/larger responses (e.g., Gilmore et al., 2016). What still remains unknown is whether participants use different visual features for different items and whether they use several features on a given item. To discover this, it is important to assess how participants accomplish dot comparison tasks on a trial-by-trial basis. This is what we did in the present study to test the hypothesis that people use several strategies in dot comparison tasks, a hypothesis that is based on previous findings on strategic variations.

Previous findings on strategic variations in dot comparison tasks.

According to the strategy variability hypothesis tested here, participants may use several strategies to accomplish dot comparison tasks. A strategy is generally defined as a "procedure or a set of procedures to achieve a higher-level goal or task" (Lemaire & Reder, 1999, p. 365). In the present context of dot comparison tasks, strategies were distinguished on the basis of which visual features (e.g., distance between dots,

dot size, total surface occupied by dots) participants chose the more numerous of two dots collections. Lemaire and Siegler (1995) distinguished the following four strategy dimensions: strategy repertoire (i.e., which strategies are used?), strategy distribution (i.e., how often each available strategy is used?), strategy execution (i.e., how quickly and accurately each strategy is applied?), and strategy selection (i.e., how do people choose among available strategies on each item?). The present study aimed at testing the usefulness of this framework to study strategies in dot comparison tasks and age-related differences therein. Previous research on cognitive aging found significant differences between young and older adults in each of these strategy dimensions (see Lemaire, 2016, for an overview). Indeed, in a wide variety of cognitive domains, young and older adults tend to use different types (and number) of strategies and available strategies with different frequencies, and to have different strategy preferences. Moreover, older adults tend to execute strategies less efficiently (i.e., they tend to be slower and less accurate) and select strategies more poorly (e.g., they tend to choose the best strategy for each item less often than young adults). This has been found in a wide variety of cognitive domains, including pattern recognition, attention, memory, problem solving, decision making, reasoning, language, and arithmetic (see Lemaire, 2016, for an overview).

In the specific domain of numerosity estimation, previous data suggest that participants use different strategies and that young and older adults may differ in these strategies. Thus, Gandini, Lemaire, and Dufau (2008) asked young and older adults to accomplish numerosity estimation tasks. They collected verbal reports and eye-movements to investigate how participants provide a quick and rough estimate of the number of dots in collections. Results showed that both young and older adults used the same set of six different strategies, but varied in how well they executed each strategy (resulting in poorer performance in older adults), how often they used each strategy, and how the type of collections influenced their strategy use. Whether such strategy variability and systematic age-related differences in strategic variations can be observed in the numerosity comparison tasks is still unknown. In the present study, we collected verbal protocols on each trial in young and older adults to determine whether participants use several strategies, and whether young and older adults differ in dot comparison strategies.

The present study

The present study aimed to further our understanding of how participants accomplish numerosity comparison tasks. Young and older adults performed a dot comparison task in which they had to determine as quickly and accurately as possible which collection has more dots. After each item, we collected verbal protocols and asked participants to indicate on the basis of which visual feature they selected the more numerous of the two dot collections. On each trial, participants saw two dot collections that varied in numerosity as well as visual features. Thus, one collection included 24 dots and the other included 12—48 dots. Some items were congruent (e.g., the more numerous dot collection was displayed with a larger convex hull) whereas other items were incongruent (e.g., the larger collection was displayed with a smaller convex hull).

Before investigating strategic variations, we determined whether collecting trial-by-trial verbal reports changed participants' approach to dot comparison tasks. We compared our participants' performance with the participants' performance in Gebuis and Reynvoet (2012)'s study, aa the researchers conducted the same experiment in young adults with the same items without collecting verbal protocols. Consistent with previous findings, we expected to observe congruency effects in both age groups (i.e., poorer performance on incongruent items relative to congruent items). We also expected to observe larger congruency effects in older adults than in young adults.

Next, we examined strategies used by young and older participants to accomplish dot comparison tasks, investigating which strategies they used, how many strategies individuals used, how often they used each strategy, and for which items they used each available strategy. Moreover, we analyzed age-related differences in these strategic variations. The objective was to determine whether older adults use fewer strategies than young adults, whether young and older adults exhibit different strategy preferences, and whether young and older adults' strategy selection is similarly influenced by item congruency in young and older adults.

Method

Participants. 73 participants took part in this study: 37 young adults (i.e., 21 men; 18-26 years of age; *mean age*: 21.1 years) and 26 older adults (i.e., 6 men; 65-93 years of age; *mean age*: 74.8 years).

Older adults were recruited from senior community centers and young adults were students at Aix-Marseille University (Marseille, France). Older adults accomplished the Mini Mental-State Examination (i.e., MMSE; Folstein et al., 1975), which is a clinical test that provides a global measure of cognitive impairment in older adults. No older adults were excluded because they all had scores larger than the usual cut-off score of 27 (i.e., *mean*: 28.0). All participants had normal or corrected-to-normal vision, and none of them had a prior experience with the task or the experimental apparatus.

Stimuli. Participants completed a dot comparison task. Each trial in this task included two collections of white dots. These two collections were simultaneously displayed on a black background, side-by-side on a 15″ computer screen. Participants were asked to select which collection was more numerous using left and right keys marked on the keyboard. There were eight practice trials followed by a total of 80 experimental trials divided into four blocks. One dot collection always represented 24 dots and the other 12, 16, 18, 20, 22, 26, 29, 32, 36, or 48 dots. This resulted in five ratio conditions (i.e., ratios of 0.5, 0.6, 0.7, 0.8, and 0.9 for numbers smaller or larger than 24). The dot collections were constructed following the method of Gebuis and Reynvoet (2011, using a freely available Matlab script provided online: http://titiagebuis.eu/Materials.html). This script controlled for cumulative surface area (i.e., area extended, item size, total surface, density) and convex hull, and generated four types of items (see Figure 1). The first (fully congruent) included pairs of collections for which the more numerous one had a larger cumulative surface area and a larger convex hull. The second (fully incongruent) included pairs of collections for which the more numerous one had a smaller cumulative surface area and a smaller convex hull. The third items (partially incongruent: cumulative surface area incongruent, convex hull congruent) included pairs of collections for which the more numerous one had a smaller cumulative surface area and larger convex hull. The fourth items (partially congruent: cumulative surface area congruent, convex hull incongruent) included pairs of collections for which the more numerous one had a larger cumulative surface area and a smaller convex hull.

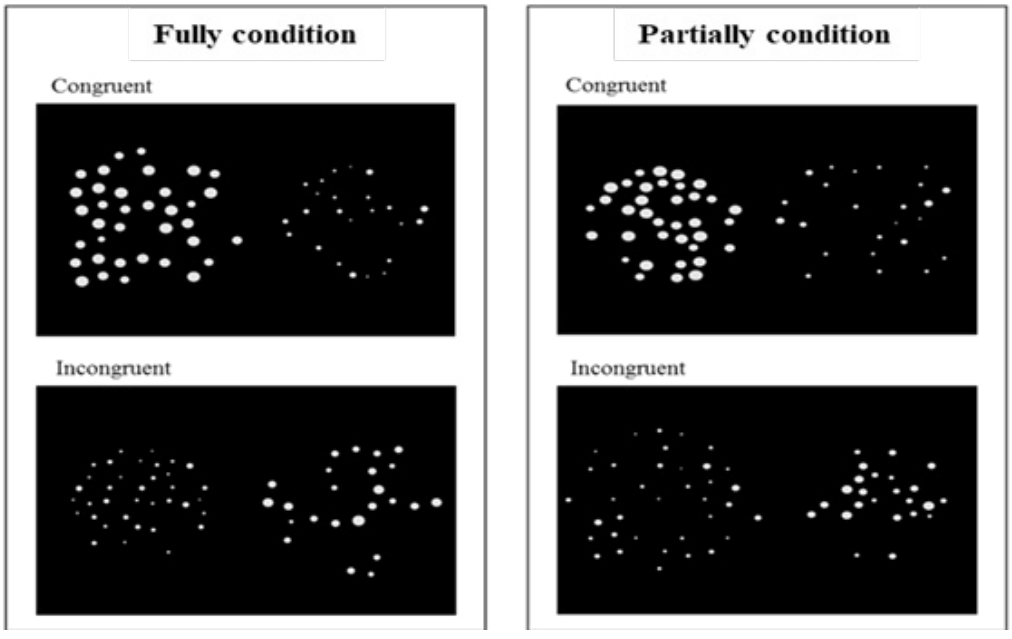

Figure 1: An example of the four item types created with the Gebuis and Reynvoet script. All items represent a 36 vs. 24 dots

Procedure and Design. The presentation of stimuli was controlled by the E-Prime Software. Each trial began with a 500-ms blank screen, followed by a warning signal ("*") displayed for 400 ms in the center of the screen. Then, each item was displayed for 1500 ms (see Figure 2). Participants were asked to indicate, as quickly and accurately as possible, which collection of dots included the largest number of dots by pressing the appropriate key (i.e., S or L key) on an AZERTY keyboard. Participants were asked to respond within 1500 ms before stimulus disappeared. If no response was given within 1500 ms, a blank screen appeared, and participants had to press a key (i.e., S or L).

After each response, participants saw a signal (i.e., "?") displayed for an unlimited time, asking them to indicate "Which visual features influenced you to select the largest collection of dots?" On each trial, the experimenter recorded participant's verbal protocols with a tape recorder. Those protocols were subsequently coded by an independent coder for each item based on the information recorded by the experimenter. Two raters who independently classified cue-based strategies on 100 randomly chosen trials agreed on 96% of them.

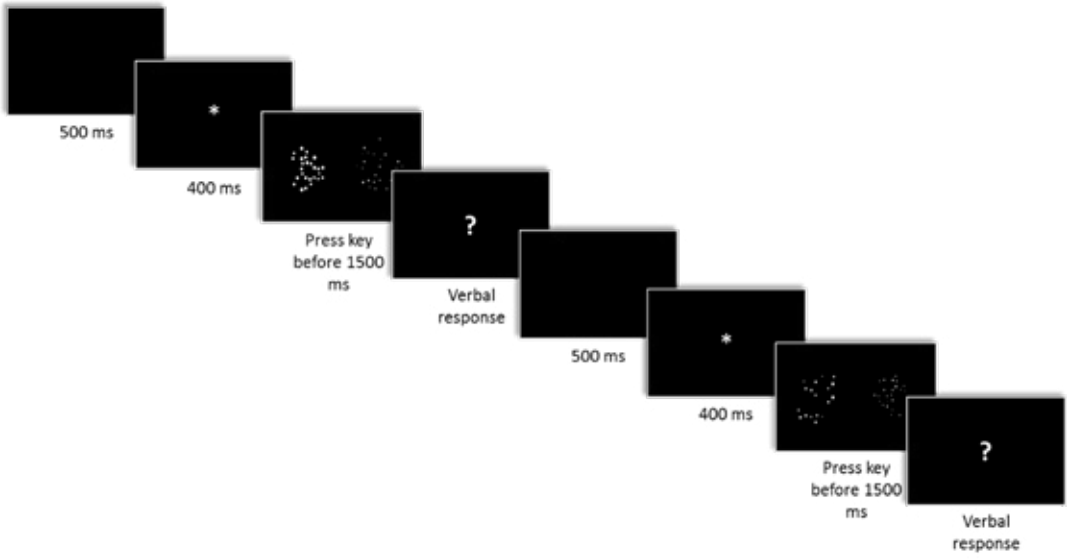

Figure 2: Sequence of events in a trial

Results

Results are reported in two main parts. First, we examined age-related differences in participants' performance (i.e., accuracy and response times). This also served the purpose of determining whether our procedure replicates findings previously reported by Gebuis and Reynvoet (2012) or whether collecting verbal protocols changed participants' approach to the task. Then, we investigated strategies used by young and older participants in dot comparison task. Preliminary analyses were run with the block factor (i.e., Block 1, 2, 3, and 4). No main or interaction effects involving this factor proved significant. In all the results, unless otherwise noted, differences are significant to at least $p<.05$.

Performance

Mixed-design ANOVAs were performed on mean percent errors and response times, 2 (Age: Young, older adults) × 2 (Congruency: Congruent, incongruent) x 2 (Conditions: Fully congruent, partially congruent), with repeated measures on the last two factors. Results showed main effects of age, congruency, and conditions on mean percentages of errors (See Table 1). Young adults made fewer errors than older adults

(i.e., 30.1% and 41.5% respectively; $F(1,61)=10.98$, $MSe=2003$, $\eta^2p=.15$). Participants made fewer errors while comparing congruent items (i.e., 29.0%) than while comparing incongruent items (i.e., 42.6%; $F(1,61)=24.86$, $MSe=11296$, $n^2p=.29$), and participants made fewer errors in partially congruent and incongruent conditions than in fully congruent and incongruent conditions (34.3% and 37.3%; $F(1,61)=5.67$ $MSe=533$, $n^2p=.09$). Finally, the Congruency x Condition interaction was significant ($F(1,61)=199.60$, $MSe=42276$, $n^2p=.77$). Participants made 39.9% more errors on fully incongruent than on fully congruent items ($F(1,61)=166.93$, $n^2p=.73$), and 12.7% more errors on partially congruent than on partially incongruent items ($F(1,61)=13.16$, $n^2p=.18$). Interestingly, these findings replicate results reported by Gebuis and Reynvoet (2012), suggesting that verbal protocols collected here did not change how participants approached the dot comparison task.

Table 1. Young and older adults' percentages of errors, and mean solution times (ms)

			Error rates			
	Fully congruent	**Fully incongruent**	**CE Fully**	**Partially congruent**	**Partially incongruent**	**CE Partially**
Young adults	12.1	52.1	40.0	35.8	20.3	-15.5
Older adults	22.6	62.4	39.8	45.6	35.6	-10.0
			Response times			
	Fully congruent	**Fully incongruent**	**CE Fully**	**Partially congruent**	**Partially incongruent**	**CE Partially**
Young adults	907	1030	123	963	951	-12
Older adults	1036	1131	95	1084	1135	51

Note: CE= Congruency Effects

Analyses of response times showed main effects of age and congruency (see Table 2). Young participants were faster (963 ms) than older participants (1096 ms; $F(1,61)=12.21$, $MSe=272872$, $n^2p=.17$). All participants compared congruent items more quickly than incongruent items (997 vs. 1062 ms; $F(1,61)=39.37$, $MSe=252174$, $n^2p=.39$). The Congruency x Condition interaction ($F(1,61)=22.77$, $MSe=121328$, $n^2p=.27$) revealed that participants were significantly faster at fully congruent items than at fully incongruent items ($F(1,61)=65.51$, $n^2p=.529$) and were equally fast at partially congruent and partially incongruent items ($F<1.0$). Finally, the Age x Congruency x Condition interaction was significant ($F(1,61)=5.94$, $MSe=31659$, $n^2p=.09$). Again, like on error rates, these findings on response times replicate those reported by Gebuis and Reynvoet (2012) in young adults. Congruency effects in older adults were larger when they compared fully incongruent (relative to fully congruent) items (95 ms; $F(1,31)=22.06$, $n^2p=.266$) than when they compared partially incongruent items (relative to partially congruent ; 51 ms; $F(1,31)=5.25$, $n^2p=.079$). These results show that congruency effects were found for both partially and fully congruent/incongruent conditions in older adults, in contrast to young adults (who showed congruency effects only in fully conditions). This suggests that when compared to young adults who are more sensitive to convex hull, older participants were influenced by all the visual features.

In sum, collecting verbal protocols did not change the way participants accomplished dot comparison task. Similar patterns of results were found both in Gebuis and Reynvoet (2012)'s study where no verbal protocols were collected and in the present study where we collected verbal protocols on each item.

Strategies

Examination of verbal protocols collected on each item for each participant revealed that participants used nine different strategies. These strategies were based on different visual features (see Table 2):

- *Distance*: Participants focused on the (larger or smaller) distance between dots.
- *Dot size*: Participants based their choices on whether dots were small or large.

- *Total surface*: Participants focused on the total surface covered by dots.
- *Shape*: Participants selected a collection with a recognizable shape (e.g., butterfly).
- *Dot size and distance*: Participants combined the size of dots and the distance between dots.
- *Dot size and shape*: Participant used the size of dots and assigned a specific shape to the collection at the same time.
- *Dot size and total surface*: Participants combined the size of dots and the total surface covered by dots.
- *Distance and total surface*: Participants combined the distance between dots and the total surface covered by dots.
- *Dot size, distance, and total surface*: Participants combined the size of dots, distance between dots, and total surface covered by dots.

Table 2. List of strategies (and examples of verbal report)

Strategies	Numbers of visual features	Example of verbal reports	Mean percent use
Distance	1	*"Dots were widely spaced from each other"*	32.2
Dot Size	1	*"Dots were big"*	19.4
Total surface	1	*"Dots occupied the large surface"*	13.2
Shape	1	*"Dots looks like a butterfly"*	0.6
Dot size + Distance	2	*"Dots were big and close from each other"*	18.8
Dot size + Shape	2	*"Dots were small and similar to a diamond"*	0.3
Dot size + Total surface	2	*"Dot were big and occupied a large space"*	5.3
Distance + Total surface	2	*"Dots were widely spaced and occupied the large surface"*	1.9
Dot size + Distance + Total surface	3	*"Dots were small, very close from each other, and occupied a reduce space"*	0.6
Others strategies	-	*"I don't know"*	1.8

While 3 young and 1 older individuals used only single cue on all items, 34 young and 25 older individuals used a single cue on some items and several cues on other items.

We next analyzed the mean number of strategies used by young and older individuals and how often participants used each strategy. To increase the number of observations for each congruent and incongruent items, we disregarded whether items were fully or partially congruent/incongruent.

Mean number of strategies used by individuals were analyzed with an ANOVA with a 2 (Age: Young, older adults) x 2 (Congruency: Congruent, incongruent items), with age as the only between-participants factor. Results revealed that young and older individuals used an equal number of strategies (i.e., both groups used three strategies, $F<1.0$). Participants used significantly more strategies for congruent items than for incongruent items (i.e., 3.6 *vs.* 3.2 strategies: $F(1,61)=6.15$, $MSe=6$, $n^2p=.09$).

Mean percent use of the four strategies that were used on at least 5% of problems (i.e., Dot size, Total surface, Distance, Dot size, and Distance) were analyzed with an ANOVA involving a 2 (Age: young, older adults) × 2 (Congruency: congruent, incongruent items) × 4 (Strategy: Dot size, Total surface, Distance, Dot size and Distance) with repeated measures on the last two factors (see Table 3).

Results showed a main effect of strategy ($F(3,183)=8.89$, $MSe=7798$, $\eta^2p=.13$). The distance strategy was used most often, followed by the dot size strategy, the dot size and distance strategy, and finally the total surface strategy.

Moreover, the Congruency x Strategy interaction was significant, showing that participants used strategies in different proportions on congruent and incongruent items ($F(3,183)=60.49$, $MSe=14203$, $n^2p=.50$). This interaction was qualified by a significant Age x Congruency x Strategy interaction ($F(1,61)=5.94$, $MSe=31659$, $n^2p=.09$). As can be easily seen from Table 3, this interaction showed different

strategy distributions on congruent and incongruent items for young and older adults. Older adults used most often the dot size strategy (29.5%), followed by the distance strategy and dot size and distance strategy, and used the total surface strategy least often for congruent items. They used the distance strategy most often (39.9%), followed by the total surface strategy, the dot size strategy, and used the dot size and distance strategy least often for incongruent items. Young adults used mostly the dot size and distance strategy (31.6%), followed by the dot size strategy, the distance strategy, and the total surface strategy for congruent items. They used the distance strategy most often (47.2%), followed by the total surface strategy, the dot size strategy, and the dot size and distance strategy for incongruent items.

To summarize, both age groups used the same set of nine strategies and an equal number of strategies while accomplishing dot comparison task. However, strategy distributions differed between young and older adults when they compared congruent and incongruent items. Young adults tended to use strategies based on a single visual feature while comparing incongruent collections of dots and used strategies based on several visual features for congruent items. Older adults tended to use strategies based on a single visual feature for both congruent and incongruent items.

Table 3. Mean percentages use of each strategy on congruent and incongruent items by young and older adults

	Young adults			Older adults		
	Congruent	**Incongruent**	**Means**	**Congruent**	**Incongruent**	**Means**
Distance	17.4	47.2	32.3	24.1	39.9	32.0
Total surface	5.0	19.6	12.3	7.9	21.2	14.6
Dot size	29.7	7.8	18.8	29.5	11.2	20.4
Dot size + Distance	31.6	10.9	21.3	21.7	9.1	15.4

General discussion

The objective of the present study was to further our understanding of how young and older adults accomplish numerosity comparison tasks. In this context, we adopted a strategy perspective and examined age-related differences in strategic variations in dot comparison tasks. In addition to performance, we assessed strategies used by participants on each item, and determined whether young and older adults exhibit different strategy repertoires and preferences and whether strategy selection is similarly influenced by item characteristics in both age groups. First, consistent with previous findings reported by Gebuis and Reynvoet (2012), we observed congruency effects in both age groups (i.e., poorer performance in incongruent items relative to congruent items). Older adults were slower and less accurate overall and showed larger congruency effects than young adults. Secondly, and most importantly, we found that several strategies were used to compare two collections of dots. Also, young and older adults had different strategy preferences depending on item congruency. These findings have important implications for our understanding of mechanisms underlying numerosity comparison and age-related differences in these mechanisms.

How do participants accomplish numerosity comparison tasks?

Previous findings on numerosity comparison found that participants are influenced by visual features (e.g., convex hull, surface area) of stimuli, above and beyond numerical properties. Most importantly, previous research aimed at determining which visual feature is most crucial (e.g., Clayton et al., 2015; Clayton & Gilmore, 2015; Gilmore et al., 2016; Halberda et al., 2008; Mazzocco, et al., 2011). For example, in Gilmore et al.'s (2016) study, the influence of dot area and convex hull on participants' accuracy was tested. Results showed that convex hull was most influential on participant's performance (see also Clayton & Gilmore, 2015 for similar results). In this context, one issue is whether participants actually rely on only one visual feature (e.g., convex hull) and the same feature for all items, or whether they use different/several features

for one item and across all items, an issue that we refer to as strategy variability. Here, we addressed this issue by collecting comparison strategies on each item in the context of dot comparison task. We found that participants used at least nine different strategies to accomplish dot comparison task and selected strategies on a trial-by-trial basis. Such evidence of strategic variations in numerosity comparison tasks have important implications to further our understanding of mechanisms underlying numerosity comparison performance.

First, previous research suggested that numerosity comparison relies on an "approximate number system" (ANS), defined by Dehaene (1997) as a cognitive system that underlies number sense and skills at discriminating between different numerosities. Upon encoding collections of dots, participants visually scan the stimulus, retrieve a numerical representation in long-term memory, compare the difference between the encoded representation and the retrieved representation, and then adjusted their answer on the basis of this difference (e.g., Gandini et al., 2008; Siegel, Goldsmith, & Madson, 1982). Using several strategies, each of which relies on different visual features, does not mean that participants do not process numerical features of dot collections, and that ANS is not crucial for numerosity comparison. It means that, in addition to, and possibly correlated with numerosity of dot collections, participants are not making use of a single visual feature but use different visual features for different items and sometimes for a given item. Note that multiple-strategy use found here is not specific to dot comparison tasks. Indeed, Gandini et al. (2008) also found that participants used several strategies when shown dot collections that included 40—460 dots and asked to find as quickly as possible the approximate number of dots in each collection. Thus, multiple-strategy use is characteristics of numerosity comparison and contrasts with a view that emerged from previous research suggesting that participants use a single strategy.

Since participants use several strategies to select the most numerous dot collection, we examined how often they use these strategies and whether their strategy use is influenced by item characteristics. Our findings revealed that participants used strategies with different proportions; they used most frequently the distance strategy, followed by the dot size strategy, the dot size and distance strategy, and finally the total surface strategy. Surprisingly, participants never mentioned that they relied on convex hull. This is in contrast to previous findings suggesting that convex hull is the most influential visual feature on which participants base their smaller/larger selection (e.g., Clayton & Gilmore, 2015; Gilmore et al., 2016). That participants did not refer to convex hull here does not mean they did not rely on it. It only means that this is not the way participants articulated their choices. As convex hull may be used less explicitly or consciously than other strategies, its use may be included in other strategies. For example, when participants mentioned the surface strategies, it is possible that they indeed used the surface strategy for some items and convex hull for other items. Yet, it was simpler for them to claim that they used the surface strategy. This does not undermine the validity of verbal protocols collected here as a number of previous studies in a wide variety of cognitive domains have shown the validity of verbal protocols to investigate strategies. It nevertheless suggests that future studies may examine whether participants use convex hull to compare collections of dots. This could be done in numerous ways, such as providing participants a list of strategies and asking them to select which strategy they used for each item from strategies proposed. Also, collecting eye movements may reinforce the validity of verbal protocols to assess strategies and refine determination of which strategy is used for each item, as previous studies have found, included in numerosity comparison tasks (e.g., Gandini et al., 2008).

Another interesting finding here was that strategy use depended on whether numerical and visual features matched (e.g., larger collections had larger convex hull) or mismatched (larger collections were displayed with smaller convex hull). Individuals tended to use a smaller number of strategies and strategies based on a single visual feature for incongruent items whereas they used a larger number of strategies and strategies based on several visual features for congruent items. One of the possible reasons for participants to use fewer strategies and strategies based on a single visual feature more often for incongruent items is possibly the fact that these items require more cognitive resources (i.e., they involve inhibiting irrelevant visual features to focus on the relevant numerosity dimension of the stimulus). Inhibiting a single visual feature requires fewer resources than inhibiting several visual features. This may have led participants to rely on single visual features for incongruent items. For congruent items, as several features match with the

numerosity of dot collections, it makes sense to rely on several visual features, as information provided by these features converge to the same smaller/larger decisions. Relying on several visual features increases evidence accumulation in favor of smaller/larger decisions, speeds up and increases confidence in these smaller/larger decisions.

At a very general level, it is important to note that the present findings of strategy variability extend to the numerosity comparison domain strategy variability previously found in many other cognitive domains (e.g., Campbell & Alberts, 2009; Duverne et al., 2004; Gandini et al., 2008; Hodzik & Lemaire, 2011; Hartley & Anderson, 1983). This is interesting because numerosity comparison could be viewed as one of these cognitive domains where strategy variability, if it exists, is less crucial than in many other domains, given how automatically people estimate numerosities of dot collections or are able to quickly see which of the two collections of dots has more dots. As in many other cognitive domains, multiple-strategy use suggests that investigating strategic variations in numerosity comparison should very useful to understand effects already documented in the subject literature, such as display time of collections (e.g., Inglis & Gilmore, 2013), sequential vs. simultaneous presentation of collections of dots (e.g., Price et al., 2012), or collections of dots displayed either intermixed or separately (e.g., Price et al., 2012). Note that here collections of dots were displayed for 1500 ms to both young and older adults. It could be argued that some age differences (e.g., larger use of multiple cues in young than in older adults) found here may be the result of the same presentation duration to both age groups. Given that older adults are known to process information more slowly than young adults (e.g., Salthouse, 1996), they may have not had enough time to use multiple cues. Although this is a possibility, in the present as well as in previous studies, older adults made their judgment in less than 1000 ms (e.g., Capelletti et al., 2014, found that older adults responded in 633 ms; see also Norris et al., 2015). Therefore, it is reasonable to think that 1500 ms was enough for older adults to accomplish our dot comparison tasks. Of course, given the differences in procedures across studies (e.g., we collected verbal protocols here, previous studies did not), it is possible that presentation duration of stimuli may result in age-related differences regarding strategic aspects of performance in dot comparison tasks, an issue that may be investigated in future studies.

Age-related differences in numerosity comparison strategies.

Strategy variability has important implications to further understand individual differences in numerosity comparison, such as aging effects. Here, we found both similarities and differences between young and older adults' strategies. First, young and older individuals used the same number and the same set of strategies. To our knowledge, the present study is the first to reveal no age-related differences in strategy repertoires. Previous works have reported age-related differences in strategy dimensions to accomplish cognitive tasks. Older adults tend to use simpler strategies and fewer strategies than young adults in a wide variety of cognitive domains, including arithmetic (e.g., Hodzik & Lemaire, 2011; Lemaire & Arnaud, 2008), reading (e.g., Shake, Noh, & Stine-Morrow, 2009), selective attention (e.g., Folk & Hoyer, 1992), episodic memory (e.g., Kuhlmann & Touron, 2012), and decision making (e.g., Mata & Nunes, 2010). The fact that strategy repertoires were the same in young and older adults here in dot comparison tasks suggests that age differences in strategy repertoire are not systematic and depend on cognitive domains and/or tasks. As discussed by Lemaire (2016), the subject literature does not provide sufficient data across multiple cognitive domains to know the defining features of cognitive domains or tasks in which young and older adults differ in the type and number of strategies, and of domains or tasks showing no age differences.

Interestingly, the present study revealed that age-related differences in how often strategies were used depend on item congruency. Although, young and older adults' strategy preference was the same while comparing incongruent items, these preferences differed while comparing congruent items. Such findings are important when analyzing age differences in effects of item characteristics (such as item congruency) on participants' performance. Indeed, previous findings revealed that young and older adults' performance was the same in congruent items but differed in incongruent items (e.g., Cappelletti et al., 2014; Norris et al., 2015), yielding larger congruency effects in older than in young adults. The present results suggest that older adults' poorer performance in incongruent items relative to young adults' does not stem from young

and older adults' using different strategies and/or from age differences in strategy preferences. Discarding age-related differences in strategies reinforces Cappelletti et al.'s proposal that age-related differences in performance in incongruent items are the result of age-related differences in inhibition.

One limitation of the present study is that with insufficient number of observations resulting from strategy selection biases, we could not examine whether congruency effects differed as a function of strategies. Future studies may further investigate whether age-related differences in congruency effects are modulated by strategies with procedures that enable to collect enough data points in each condition and that control for strategy selection biases. This is possible in experiments forcing young and older individuals to execute each available strategy for all items, as proposed by Siegler and Lemaire (2017)'s choice/no-choice method.

In conclusion, the present study is the first to examine multiple-strategy use in young and older adults while accomplishing dot comparison tasks, , and that strategy use depended on item characteristics. Such age-related and condition-related differences in strategy use are important to take into account in future works on numerosity comparison, given that strategies differ in relative efficiency. Future works may indeed adopt the present strategy approach to further investigate effects of stimulus, participants', and situations' characteristics on performance when participants accomplish numerosity comparison tasks.

Acknowledgments: This work was supported by the Centre National de la Recherche Scientifique (French National Science Foundation), by a grant from the Agence Nationale de la Recherche (Grant # ANR-17-CE28-0003-01-01) to PL, and by a doctoral fellowship to AR from Aix-Marseille University.

References

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B. (2005). Neural correlates of symbolic number processing in children and adults. *NeuroReport, 16*, 1769. DOI: 10.1097/01.wnr.0000183905.23396.f1.

Barth, H., Kanwisher, N., & Spelke, E. (2002). The construction of large number representations in adults. *Cognition*, *86*(3), 201-221. DOI: 10.1016/S0010-0277(02)00178-6.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the national academy of sciences*, *102*(39), 14116-14121.

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222. DOI: 10.1016/j.cognition.2004.09.011.

Bouazzaoui, B., Isingrini, M., Fay, S., Angel, L., Vanneste, S., Clarys, D., & Taconnat, L. (2010). Aging and self-reported internal and external memory strategy uses: The role of executive functioning. *Acta Psychologica*, *135*(1), 59-66. DOI: 10.1016/j.actpsy.2010.05.007.

Campbell, J. I., & Alberts, N. M. (2009). Operation-specific effects of numerical surface form on arithmetic strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), *999*. DOI: 10.1037/a0015829.

Cappelletti, M., Didino, D., Stoianov, I., & Zorzi, M. (2014). Number skills are maintained in healthy ageing. *Cognitive Psychology*, *69*, 25–45. DOI: 10.1016/j.cogpsych.2013.11.004.

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, *148*, 163–172. DOI: 10.1016/j.actpsy.2014.01.016.

Clayton, S., & Gilmore, C. (2015). Inhibition in dot comparison tasks. *ZDM: The International Journal on Mathematics Education, 47*, 759–770. DOI: 10.1007/s11858-014-0655-2.

Clayton, S., Gilmore, C., & Inglis, M. (2015). Dot comparison stimuli are not all alike: the effect of different visual controls on ANS measurement. *Acta Psychologica*, *161*, 177–184. DOI: 10.1016/j.actpsy.2015.09.007.

Craik, F. I., & Salthouse, T. A. (Eds.). (2011). *The handbook of aging and cognition*. Psychology press. DOI: 10.1002/acp.1505.

Crites, T. (1992). Skilled and less skilled estimators' strategies for estimating discrete quantities. *The Elementary School Journal, 92*(5), 601–619. DOI: 10.1086/461709.

Dehaene, S. (1997). *The number sense: how the mind creates mathematics* (New York: Oxford University Press).

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 in Neuroscience and Education*, *2*(2), 48-55. DOI: 10.1016/j.tine.2013.06.001.

DeWind, N. K., & Brannon, E. M. (2016). Significant Inter-Test Reliability across Approximate Number System Assessments. *Frontiers in Psychology, 7*. DOI: 10.3389/fpsyg.2016.00310.

Duverne, S., & Lemaire, P. (2004). Age-related differences in arithmetic problem-verification strategies. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *59*, P135-142. DOI:10.1093/geronb/59.3.P135.

Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning & Behavior, 26*(3), 243–256. DOI: 10.1037/0033-2909.92.3.547.

Folk, C. L., & Hoyer, W. J. (1992). Aging and shifts of visual spatial attention. *Psychology and aging*, *7*(3), 453. DOI: 10.1037/0882-7974.7.3.453.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. DOI: 10.1016/0022-3956(75)90026-6.

Fush, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: contributions of inhibitory control. *Developmental Science*, *16*, 136–148. DOI : 10.1111/desc.12013.

Gandini, D., Lemaire, P., Anton, J.-L., & Nazarian, B. (2008). Neural correlates of approximate quantification strategies in young and older adults: an fMRI study. *Brain Research*, *1246*, 144–157. DOI: 10.1016/j.brainres.2008.09.096.

Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavioural Research Methods*, *43*, 981–986. DOI: 10.3758/s13428-011-0097-5.

Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual features. *Journal of Experimental Psychology: General*, *141*, 642–648. DOI: 10.1037/a0026218.

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

Gilmore, C., Keeble, S., Richardson, S., & Cragg, L. (2015). The Role of Cognitive Inhibition in Different Components of Arithmetic. *ZDM: The International Journal on Mathematics Education*, *47*, 771–782. DOI: 10.1007/s11858-014-0659-y.

Gilmore, C., Cragg, L., Hogan, G., & Inglish, M. (2016). Congruency effects in dot comparison tasks: Convex hull is more important than dot area. *Journal of Cognitive Psychology*, *28(8)*, 923-931, DOI: 10.1080/20445911.2016.1221828.

Glisky, E. L. (2007). Changes in cognitive function in human aging. *Brain aging: Models, methods, and mechanisms*, 3-20. DOI: 10.1201/9781420005523.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668. DOI: 10.1038/nature07246.

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. *Proceedings of the National Academy of Sciences, 109*, 11116–11120. DOI: 10.1073/pnas.1200196109.

Hartley, A. A., & Anderson, J. W. (1983). Task complexity and problem-solving performance in younger and older adults. *Journal of Gerontology*, *38*(1), 72-77. DOI: 10.1093/geronj/38.1.72.

Hauser, M. D., Tsao, F., Garcia, P., & Spelke, E. S. (2003). Evolutionary foundations of number: spontaneous representation of numerical magnitudes by cotton–top tamarins. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(1523), 1441–1446. DOI: 10.1098/rspb.2003.2414.

Hinault, T., Dufau, S., & Lemaire, P. (2014). Sequential modulations of poorer-strategy effects during strategy execution: An event-related potential study in arithmetic. *Brain and Cognition*, *91*, 123–130. DOI: 10.1016/j.bandc.2014.09.001.

Hinault, T., Lemaire, P., & Phillips, N. (2016a). Aging and sequential modulations of poorer strategy effects: An EEG study in arithmetic problem solving. *Brain Research*, *1630*, 144–158. DOI: 10.1016/j.brainres.2015.10.057.

Hinault, T., Lemaire, P., & Touron, D. (2016b). Aging effects in sequential modulations of poorer-strategy effects during execution of memory strategies. *Memory*, 1–11. DOI: 10.1080/09658211.2016.1146300.

Hinault, T., Badier, J. M., Baillet, S., & Lemaire. P. (2017). The sources of sequential modulations of control processes in arithmetic strategies: A MEG study. *Journal of Cognitive Neuroscience.* DOI: 10.1162/jocn_a_01102.

Hodzik, S., & Lemaire, P. (2011). Inhibition and shifting capacities mediate adults' age-related differences in strategy selection and repertoire. *Acta Psychologica, 137*(3), 335-344. DOI: 10.1016/j.actpsy.2011.04.002.

Inglis, M., & Gilmore, C. (2013). Sampling from the mental number line: how are approximate number system representations formed? *Cognition*, *129*, 63–69. DOI: 10.1075/aicr.22.12hom.

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1352. DOI: 10.1037/a0027863.

Lemaire, P. (2010). Cognitive Strategy Variations During Aging, Cognitive Strategy Variations During Aging, *Current Directions in Psychological Science*, *19*, 363–369. DOI: 10.1177/0963721410390354.

Lemaire, P., & Arnaud, L. (2008). Young and older adults' strategies in complex arithmetic. *The American Journal of Psychology*, *121*, 1–16. DOI: 10.2307/20445440.

Lemaire, P., & Hinault, T. (2014). Age-related differences in sequential modulations of poorer-strategy effects: A study in arithmetic problem solving. *Experimental Psychology*, *61*(4), 253–262. DOI: 10.1027/1618-3169/a000244.

Lemaire, P., & Lecacheur, M. (2007). Aging and numerosity estimation. *The journals of gerontology. Series B, Psychological sciences and social sciences, 62*, 305-312. DOI: 10.1093/geronb/62.6.P305.

Lemaire, P., Lecacheur, M., & Farioli, F. (2000). Children's strategy use in estimation. Canadian *Journal of Experimental Psychology, 54*(2), 141–147.

Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition*, *27*(2), 364–382. DOI: 10.3758/BF03211420.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*, 83–97. DOI: 10.1037/0096-3445.124.1.83.

Li, R. W., MacKeben, M., Chat, S. W., Kumar, M., Ngo, C., & Levi, D. M. (2010). Aging and visual counting. *PLoS ONE, 5*, e13434. DOI: 10.1371/journal.pone.0013434.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental science*, *14*(6), 1292-1300. DOI: 10.1111/j.1467-7687.2011.01080.

Luwel, K., Lemaire, P., & Verschaffel, L. (2005). Children's strategies in numerosity judgment. *Cognitive Development, 20*, 448–471. DOI: 10.1016/j.cogdev.2005.05.007.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003a). Flexibility in strategy use: Adaptation of numerosity judgment strategies to task characteristics. *European Journal of Psychology of Education, 15*(2), 246–266. DOI: 10.1080/09541440244000139.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003b). Strategic aspects of numerosity judgment: The effect of task characteristics. *Experimental Psychology, 50*(1), 63–75. DOI: 10.1027//1618-3169.50.1.63.

Mata, R., & Nunes, L. (2010). When less is enough: Cognitive aging, information search, and decision quality in consumer choice. *Psychology and aging*, *25*(2), 289. DOI: 10.1037/a0017927.

Norris, J. E., McGeown, W. J., Guerrini, C., & Castronovo, J. (2015). Aging and the number sense: preserved basic non-symbolic numerical processing and enhanced basic symbolic processing. *Frontiers in Psychology, 6*. DOI: 10.3389/fpsyg.2015.00999.

Nys, J., & Content, A. (2012). Judgement of discrete and continuous quantity in adults: number counts! *Quarterly Journal of Experimental Psychology*, *65*, 675–690. DOI: 10.1080/17470218.2011.619661.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306,* 499–503. DOI: 10.1126/science.1102085.

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*, 50–57. DOI: 10.1016/j.actpsy.2012.02.008.

Reder, L. M., Wible, C., & Martin, J. (1986). Differential memory changes with age: Exact retrieval versus plausible inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 72. DOI: 10.1037/0278-7393.12.1.72.

Revkin, S. K., Piazza, M., Izard, V., Zamarian, L., Karner, E., & Delazer, M. (2008). Verbal numerosity estimation deficit in the context of spared semantic representation of numbers: a neuropsychological study of a patient with frontal lesions. *Neuropsychologia*, *46*, 2463–2475. DOI: 10.1016/j.neuropsychologia.2008.04.011.

Shake, M. C., Noh, S. R., & Stine-Morrow, E. A. (2009). Age differences in learning from text: Evidence for functionally distinct text processing systems. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(4), 561-578. DOI: 10.1002/acp.1494.

Siegel, A. W., Goldsmith, L. T., & Madson, C. R. (1982). Skill in estimation problems of extent and numerosity. *Journal of Research in Mathematics Education, 13*(3), 211–232. DOI: 10.1016/S0010-9452(08)70504-5.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM via the choice/no-choice method. *Journal of Experimental Psychology: General*, 126, 71-92. DOI: 10.1037//0096-3445.126.1.71.

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*, 310–325. DOI: 10.1080/20445911.2014.996568.

Smets, K., Gebuis, T., Defever, E., & Reynvoet, B. (2014). Concurrent validity of approximate number sense tasks in adults and children. *Acta Psychologica*, *150*, 120–128. DOI: 10.1016/j.actpsy.2014.05.001.

Trick, L. M., Enns, J. T., & Brodeur, D. A. (1996). Life span changes in visual enumeration: The number discrimination task. *Developmental Psychology*, *32*, 925–932. DOI:10.1037/0012-1649.32.5.925.

Watson, D. G., Maylor, E. A., & Manson, N. J. (2002). Aging and enumeration: a selective deficit for the subitization of targets among distractors. *Psychology and Aging, 17*, 496–504. DOI: 10.1037/0882-7974.7.4.536.

Watson, D. G., Maylor, E. A., & Bruce, L. A. M. (2005). Search, Enumeration, and Aging: Eye Movement Requirements Cause Age-Equivalent Performance in Enumeration but Not in Search Tasks. *Psychology and Aging*, *20*, 226–240. DOI: 10.1037/0882-7974.20.2.226.

Appendix

Table A. Mean percentages use of each strategy by young and older adults.

Strategies	Young adults	Older adults	Mean percentages
Distance	32.3	32	32.2
Dot Size	18.7	20.3	19.4
Total surface	12.3	14.5	13.2
Shape	0.2	1.2	0.6
Dot size + Distance	21.2	15.4	18.8
Dot size + Shape	0.4	0.1	0.3
Dot size + Total surface	5.8	4.5	5.3
Distance + Total surface	1.2	2.9	1.9
Dot size + Distance + Total surface	0.7	0.5	0.6
Others strategies	1.4	2.3	1.8

Table B. Mean percentages use of each strategy on congruent and incongruent items in fully and partially conditions by young and older adults.

Conditions	Congruency	Strategies	Older adults	Young adults
Fully	Congruent	Distance	24.1	14.0
		Total surface	12.7	6.8
		Dot size	41.8	44.5
		Dot size + Distance	21.4	34.7
	Incongruent	Distance	50.7	60.2
		Total surface	15.9	12.2
		Dot size	22.0	14.1
		Dot size + Distance	11.4	13.4
Partially	Congruent	Distance	32.2	25.7
		Total surface	10.1	8.1
		Dot size	27.6	25.4
		Dot size + Distance	30.0	40.7
	Incongruent	Distance	41.3	48.6
		Total surface	39.7	34.0
		Dot size	8.0	5.9
		Dot size + Distance	11.0	11.5

Table C. Young and older adults' mean response times (in ms) and percentages of errors in the dot comparison task for each strategy on congruent or incongruent items.

		Reaction times (ms)			Error rates (%)		
		Young adults	**Older adults**	**Total**	**Young adults**	**Older adults**	**Total**
Congruent	Distance	957	1071	1010	7.6	4.9	6.9
	Surface	1093	1091	1092	3.8	1.8	2.4
	Dot size	953	1018	979	0.9	1.1	1.0
	Dot size + Distance	900	1063	964	3.3	1.9	2.6
Incongruent	Distance	971	1128	1029	5.0	4.4	4.6
	Surface	903	1137	999	3.1	3.2	2.9
	Dot size	1123	1108	1112	6.8	12.9	8.8
	Dot size + Distance	938	1175	1068	7.3	3.3	6.3