



**HAL**  
open science

# Mathematical reasoning in the written argumentation of primary students

Silke Ruwisch

► **To cite this version:**

Silke Ruwisch. Mathematical reasoning in the written argumentation of primary students. CERME 10, Feb 2017, Dublin, Ireland. hal-01865649

**HAL Id: hal-01865649**

**<https://hal.science/hal-01865649>**

Submitted on 31 Aug 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Mathematical reasoning in the written argumentation of primary students**

Silke Ruwisch

Leuphana University Lüneburg, Germany

[ruwisch@uni.leuphana.de](mailto:ruwisch@uni.leuphana.de)

*In Germany, there is increasing interest in language competences in mathematics lessons. Based on national curriculum standards, argumentation should also be strengthened in primary school mathematics classes (KMK, 2005). The reported interdisciplinary (linguistics and mathematics education) study on reasoning presents a model to rate arithmetic reasoning competences at primary level, in which mathematical reasoning and its linguistic realization are separately coded. In a pilot study, 243 third, fourth, and sixth grade students solved a number of arithmetic reasoning tasks. The results support a one-dimensional scale for the model of reasoning; its components identify differentiated requirements, which are formulated concretely in the coding guidelines and may point to didactical potential for language support in mathematical reasoning itself, as well as in mathematics lessons at primary level.*

*Keywords: Reasoning, written argumentation, primary school.*

## **Reasoning in early mathematics learning**

Early mathematical argumentation can be divided into four steps: detecting mathematical regularities, describing them, asking questions about them and giving reasons for their validity (Bezold & Ladel, 2014; Meyer, 2010; Bezold, 2009; Wittmann & Müller, 1990). The content base for argumentation is achieved by description of the detected structures or by reference to common knowledge (Krummheuer, 2000); reasoning is then needed to verify the described regularities as true (Toulmin, 2003/1958; Schwarzkopf, 1999).

The didactical value of reasoning in mathematics learning lies in gaining deeper insights into mathematical structures, so developing one's mathematical knowledge. In this sense, reasoning leads to questions about mathematical statements to ensure their correctness and to develop new mathematical connections (Steinbring, 2005). Two intertwined processes can be distinguished: one's own understanding and the process of sharing this understanding with others. In most cases, these processes don't occur separately, but are the response to cognitive-social needs (Harel & Sowder, 2007; Hersh, 1993). It follows that, in its epistemic function, mathematical reasoning may be monologic in leading to deeper individual understanding; in its communicative function, where mathematical structures are explained and justified, it is dialogic and dependent on other people (Neumann, Beier & Ruwisch, 2014; Ruwisch & Neumann, 2014).

In primary classrooms, mathematical reasoning usually occurs in the form of oral communication between pupils and in interactions with the teacher. These communicative processes have been widely studied. From an epistemological perspective, the emergence of shared knowledge and its structures has been described (e.g. Steinbring, 2005) while a more interactionist perspective traces the type and structure of argumentation in classroom interactions (e.g. Krummheuer, 2015).

Mathematical reasoning in this sense must be distinguished from reasoning in language classes, especially at primary level. While both are seen as concepts that develop out of situated everyday (“vernacular”) speech (Elbow, 2012), reasoning in language learning focuses much more on self-evident facts and personal meanings than on provable structures in special content areas. It follows that argumentation in language learning leads to more addressee-oriented cognitivization (Krelle, 2007), as reasoning of this kind is much more about persuasion than proving. Nevertheless, typical linguistic forms of reasoning are learned in these everyday situations, and students must learn how to use these in different content areas (e.g. Wellington & Osborne, 2001; Lemke, 1990). So, in combining mathematical and linguistic views of early reasoning, we can hope to gain a broader and deeper understanding of early reasoning.

While most age-related studies of primary students focus on oral communication, experts in language learning emphasise writing as an important instrument for deepening individual understanding (Becker-Mrotzek & Schindler, 2007; Pugalee, 2005; Galbraith, 1999; see also Wellington & Osborne, 2001; Morgan, 1998; Miller, 1991). Although primary school children are not yet expert in writing, fourth-graders are capable of constructing expository texts with a relevant number of causes in elaborating a topic (Hayes, 2012; Krelle, 2007). It may therefore be fruitful to look at their written argumentations, and especially at how they offer reasons for mathematical regularities (Ruwisch & Neumann, 2014; Fetzer, 2007).

## Modelling written mathematical reasoning

To investigate children’s written reasoning, we developed a theoretical model that combines mathematical and linguistic aspects of reasoning (Ruwisch & Neumann, 2014; Neumann, Beier & Ruwisch, 2014).

### Arithmetic reasoning tasks

Following the four steps of argumentation in primary mathematics (see above), we decided to give the children an already structured situation (see Figure 1), which explicitly requires detection, transfer and description before offering reasons for the validity of their suggestion.

<p>a) <math>18 + 10 = \underline{\quad}</math> b) <math>36 + 20 = \underline{\quad}</math> c) <math>52 + 40 = \underline{\quad}</math> d) <math>87 + 30 = \underline{\quad}</math>  <math>8 + 20 = \underline{\quad}</math>    <math>26 + 30 = \underline{\quad}</math>    <math>42 + 50 = \underline{\quad}</math>    <math>77 + 40 = \underline{\quad}</math></p> <p>Erfinde zwei weitere Päckchen, die zu den anderen passen.</p> <p>e) <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math>                      f) <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math>  <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math>                      <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math></p> <p>Vergleiche die Aufgaben im Päckchen. Schreibe auf, was dir auffällt.</p> <p>Begründe die Auffälligkeiten!</p>	<p>a) <math>18 + 10 = \underline{\quad}</math> b) <math>36 + 20 = \underline{\quad}</math> c) <math>52 + 40 = \underline{\quad}</math> d) <math>87 + 30 = \underline{\quad}</math>  <math>8 + 20 = \underline{\quad}</math>    <math>26 + 30 = \underline{\quad}</math>    <math>42 + 50 = \underline{\quad}</math>    <math>77 + 40 = \underline{\quad}</math></p> <p>Invent two further task-packages, which fit to the others.</p> <p>e) <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math>                      f) <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math>  <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math>                      <math>\underline{\quad} + \underline{\quad} = \underline{\quad}</math></p> <p>Compare the tasks in the package. Describe what you notice.</p> <p>Give reasons for your observations.</p>
---	---

**Figure 1: Complex addition task (CA) as a sample item  
(left: original version; right: English translation)**

For the purposes of this study, four different arithmetic tasks were designed. Although differing in complexity of regularities, all of these tasks focused on detection and reasoning and were easy to compute. Format ZF involved three number sequences to be continued:  $+9$ ,  $+7$ , and  $+2n$ . Format EA asked the children to continue a given additive structure by increasing all three summands by one so that the sum increases by three. In solving formats CA and CM, the children had to identify two

structures at the same time. To answer the complex addition task in Figure 1, the children had to find two tasks with the same sum. At the same time, they had to take into account that the summands must be changed by 10 in opposite directions. The multiplication tasks (CM) showed a constant difference in the product caused by the difference between the multipliers while the multiplicands remained constant.

### Sample

The data include 477 justifications written by 243 students. In total, 41 third-graders (♀21; ♂20), 96 fourth-graders (♀43; ♂53) and 106 sixth-graders (♀52; ♂54) worked out two of the four arithmetic reasoning tasks.

### Data analysis: Theoretical model of rating scales

The separate evaluation of mathematical and linguistic aspects of reasoning is fundamental to our model, which we assume allows differentiated exploration of the sub-skills of reasoning. We also wish to check whether experts in either domain (mathematics teachers and German language teachers) differ in their evaluations. As our tasks demanded both computing and continuation of a given structure (see Figure 1), competencies involving detection of a mathematical structure are distinguished from ability to offer reasons for its validity. Students' writings are rated by one detection scale and two reasoning scales (see Table 1).

*Mathematical detections:* Children were required to compute the arithmetic tasks on the sheet to identify the underlying structure and transfer it to two further packages of tasks. This process might be realised fully or only partly; sometimes, only irrelevant aspects were used to create new tasks. If the structure is transferred fully, the results of the given tasks are also correct, and three levels of this rating scale therefore seemed sufficient. This scale will not be discussed in the following application of the model, as it provides little information about reasoning skills.

Mathematical detections	Mathematical aspects of reasoning	Linguistic aspects of reasoning
irrelevant aspects as regularities	regularities (partially) described	indicators without reason-effect structure
.....	rudimentary reasoning	reason-effect structure
regularities partly transferred	reasoning through examples	explicit linguistic reference to the task
.....	partially generalized reasoning	completeness and consistency
regularities totally transferred	generalization/ formal reasoning	use of math. terminology/ decontextualization

**Table 1: Rating-scales to evaluate written mathematical reasoning**

*Mathematical aspects of reasoning:* Mathematical reasoning must be based on a description of mathematical elements. If only some regularities are described without giving reasons, this is coded as level 1. If rudimentary reasoning is given in addition to a description, the work is coded as level 2. To be rated as level 3 to 5, all relevant aspects must feature in the argumentation. If this is done

by use of examples, the work is rated as level 3; if already partly generalized, it is rated as level 4; and if it is totally general or constitutes a formal proof, it is rated as level 5.

*Linguistic aspects of reasoning:* Realisation of a mathematical argument by written language is also rated in terms of five levels, defined in terms of use of connectors and identifiable coherence of the text. If explicit linguistic indicators are used without any structured reasoning, the text is classified as level 1. If the text shows a reason-effect structure, it is coded as at least level 2. If explicit linguistic reference to the tasks is also included, the text is classified as level 3. A level 4 text shows consistent and complete argumentation. To achieve level 5, there must also be use of mathematical terminology for identifiable decontextualization.

### **Process of coding**

Each written argumentation was assessed by at least four raters; while preservice mathematics teachers concentrated on the mathematical scales, German language teachers rated the linguistic aspects. There was 62% absolute agreement in the judgments across all tasks and scales. Deviations of more than one stage occurred in 8% of cases—mainly for linguistic ratings, which were reported as more difficult. Coding quality could be seen to increase during the course of the project. Although there were acceptable internal consistencies across all tasks (Cronbach's  $\alpha = .80$ ), these values increase if only ZF ( $\alpha = .82$ ) and EA ( $\alpha = .84$ ) (which were used later in the project) are considered.

By excluding the multiplication task for the following overall scaling, an acceptable average internal consistency of individual scales was achieved for the remaining tasks ( $\alpha = .86$  for mathematical detections,  $\alpha = .81$  for mathematical aspects of reasoning and  $\alpha = .71$  for the linguistic aspects of reasoning).

### **First results**

Given the number of raters and in the interests of acceptable inter-rater-consistency ( $\alpha > .70$ ), the following results are based on the means of ratings.

### **Overall scale**

The IRT scale for the three tasks and all texts shows a common scale across all components (see Table 2). As items also conform to the model (WMNSQ .85-1.09), early mathematical reasoning in arithmetic as measured by the three tasks and ratings on our scales can probably be described as a one-dimensional construct. Looking at the three scales, it becomes clear that, as expected, it is easier to detect and transfer mathematical structures than to give reasons for their validity (negative deviation from zero). Comparing the two reasoning scales, it seems easier to realise mathematical aspects of reasoning than to find an appropriate linguistic structure. At the same time, the most stable dimension is *mathematical detections*, with a maximum difference of .783 as compared to 1.446 for the *linguistic aspects of reasoning* and 1.516 for the *mathematical aspects of reasoning*.

Comparing the three tasks, it seems that complex addition is the most difficult to transfer; simple addition and number sequences show almost no difference. The justifications show that it was easiest to realise both mathematical and linguistic aspects of reasoning in the number sequence tasks, followed by complex addition and then simple addition. Granted these differences, all tasks can be characterised as suitable for capturing mathematical reasoning in arithmetic.

Item	Mathematical detections		Mathematical aspects of reasoning		Linguistic aspects of reasoning	
	Estimate	WMNSQ	Estimate	WMNSQ	Estimate	WMNSQ
(ZF) number sequences	-1.556	1.02	-0.459	1.06	0.124	0.85
(EA) simple addition	-1.628	1.09	1.057	1.09	1.570	0.93
(CA) complex addition	-0.845	0.98	0.506	0.92	1.230	0.97

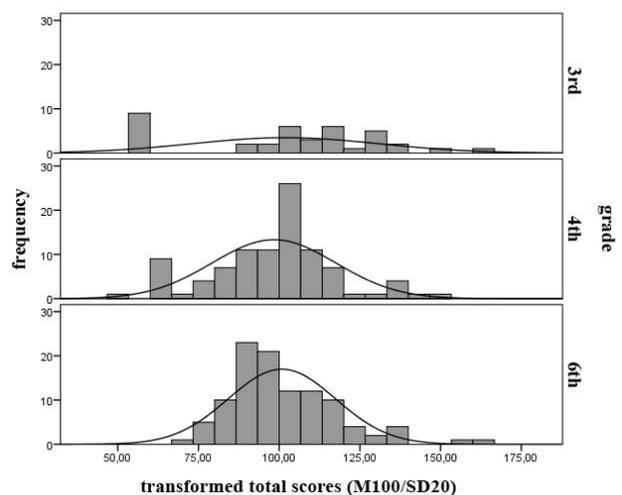
**Table 2: Item parameters (estimated) for IRT scaling**

### Student performance

Performance of the total sample is distributed normally to slightly right-shifted. On the raw scores level, 21.2% are one standard deviation above the mean; 9.6% are one standard deviation below; 6.2% are two standard deviations above; and 4.2% are two standard deviations below.

To facilitate comparison of the three groups of students, all scores were transposed to a scale with mean 100 and standard deviation 20. Figure 2 shows almost the same mean performance across third-graders ( $M = 102$ ,  $SD = 29$ ), fourth-graders ( $M = 98$ ,  $SD = 19$ ) and sixth-graders ( $M = 101$ ,  $SD = 17$ ).

Unexpectedly, reasoning competences seem not to increase over time. Interpreting the differences in standard deviations across the three groups, it seems that third-graders differ more within their group than fourth-graders, and both differ more than sixth-graders, suggesting homogenization during schooling. However, the relatively limited data and lack of comparative data means that any general conclusions remain speculative.



**Figure 2: Student performance by grade**

### Discussion of the model by application to examples

Five examples are presented here for deeper discussion of the model's adequacy (see Table 2). As it might prove difficult to discuss the argumentations only in the translated version, the original sentences are included on the left.

*Mathematical aspects of reasoning:* The child in example 1 has recognized that “something” is the same and “something” has changed. However, as he/she does not refer to any connection between the tasks or mention that the results are the same, this answer was coded as level 1. Examples 2 and 4 were coded as level 2; in example 2, it is clear that the child focused on only one of the two

relevant aspects. It is arguable whether child 4's argumentation is complete; as it is confined to one example in the task, it might be evaluated as level 3. In our opinion, the change of the summands in opposite directions is only implicit in "that's always 10 less". Answer 5 shows both connections clearly. Furthermore, the child is able to conclude (using an example) that the results must be the same, and so it is clearly to be coded as level 3. As child 3 is doing almost the same but also exhibits some generalization in using "always", we coded this as level 4.

1) es sind immer die gleichen Aufgaben nur umgedreht weil wenn man es rechnet merkt man das.	The tasks are all the same but vice versa because if you calculate, you'll realize it.
2) Es sind immer 10 mehr und 10 weniger.	It's always 10 more and 10 less.
3) Dass es die gleichen Ergebnisse sind, kommt davon, weil bei der einen Aufgabe immer 10 weniger sind als bei der anderen. Aber bei der Aufgabe wo 10 weniger sind, ist die Zahl die noch dazu gerechnet wird wiederum 10 größer als die über ihr.	The results are the same because in one task it's always 10 less than in the other one. But in the task that has 10 less, the number to be added is 10 bigger than the one above.
4) Das es immer 10 weniger sind. Zum beispiel $18+10=28$ aber wenn man 10 weg nimmt und in der mitte 10 dazu nimmt z.b. $8+20=28$ und dann kommt das gleiche ergebnis wie bei der 1. Aufgabe	That's always 10 less. For example, $18+10 = 28$ . But if you take away 10 and put 10 in the middle—for example, $8+20 = 28$ —then you'll get the same result as in the first task.
5) mir fällt auf das immer die Ersten 2 Ergebnisse gleich sind. Die Ersten zwei Ergebnisse sind gleich weil die bei zum beispiel a) $18+10=28$ und dann haben die bei $8+20$ einfach 18, 10 weniger 8, und bei 10 zehn mehr, $10-10$ ist 0, also bleibt das so	I notice that the first two results are always the same. The first two results are the same, because, for example a) $18+10 = 28$ , and then at $8+20$ it's simply 18, 10 less 8, and at 10 ten more, $10-10$ is zero, so it remains the same.

**Table 3: Examples of written argumentation for the arithmetic sample item  
(left: original version; right: English translation)**

*Linguistic aspects of reasoning:* Example 1 is coded as level 2 because as well as the comparative connector "because", a link between the sentences is also given. As example 2 includes only the indicator "always", without any link, it is coded as level 1. Examples 3 and 4 are coded as level 4 because there is a clear reasoning structure as well as a link to the tasks. As the argumentation in example 5 is ambiguous, and the language used is imprecise, it is coded as level 3.

## Conclusion

The model was again presented for discussion here to improve its didactical value in evaluating the written reasoning of fourth-graders. Although our descriptions in the coding book have continued to improve over time, there are still deviations of more than one level between raters. While we wish of course to develop the model for its psychometric interest, the levels should also help teachers to evaluate written reasoning.

Although these tasks provide a good deal of information about children's written reasoning, we have to be aware that because they focus on products collected in a test situation, the argumentation was necessarily ad hoc. As requests of this type are not part of students' normal mathematics lessons, and they do not have time to restructure their texts, neither the requisite procedural knowledge for writing nor situated mathematical argumentation can be grasped in this way. It follows that competence in mathematical reasoning—even in written form—may be higher than is indicated by the results to date.

## References

- Becker-Mrotzek, M., & Schindler, K. (Eds.) (2007). *Texte schreiben*. Duisburg: Gilles & Francke.
- Bezold, A. (2009). *Förderung von Argumentationskompetenzen durch selbstdifferenzierende Lernangebote* (Doctoral dissertation). Hamburg: Kovač.
- Bezold, A., & Ladel, S. (2014). Reasoning in primary mathematics – an ICT-supported environment. *Bildung und Erziehung*, 67(4), 409–418.
- Elbow, P. (2012). *What speech can bring to writing*. New York: Oxford UP.
- Fetzer, M. (2007). “Why should I implement writing in my classes?” An empirical study on mathematics writing. In D. Pitta-Pantazi & G. Phillipou (Eds.), *Proceedings of the 5<sup>th</sup> Congress of the European Society for Research in Mathematics Education* (pp. 1210–1219). Cyprus: CERME5.
- Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance, & D. Galbraith (Eds.), *Knowing what to write. Conceptual processes in text production* (pp. 139–159). Amsterdam: UP.
- Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed), *Second handbook of research on mathematics teaching and learning*, (Vol. 2, pp. 805–842). Reston VA: NCTM.
- Hayes, J. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388.
- Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24(4), 389–399.
- Kultusministerkonferenz (KMK) (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4)*. München, Neuwied: Luchterhand.
- Krelle, M. (2007). Wissensbasierte Argumentation lehren und lernen. In H. Willenberg (Ed.), *Kompetenzhandbuch für den Deutschunterricht* (pp. 108–114). Baltmannsweiler: Schneider.
- Krummheuer, G. (2000). Mathematics learning in narrative classroom cultures: Studies of argumentation in primary mathematics education. *For the Learning of Mathematics*, 20(1), 22–32.

- Krummheuer, G. (2015). Methods for reconstructing processes of argumentation and participation in primary mathematics interaction. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education. Examples of methodology and methods* (pp. 51–74). Dordrecht: Springer.
- Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing Corporation.
- Meyer, M. (2010). Abduction—A logical view of processes of discovering and verifying knowledge in mathematics. *Educational Studies in Mathematics*, 74, 185–205.
- Miller, D. L. (1991). Writing to learn mathematics. *The Mathematics Teacher*, 84(7), 516–521.
- Morgan, C. (1998). *Writing mathematically. The discourse of investigation*. London: Falmer.
- Neumann, A., Beier, F., & Ruwisch, S. (2014). Schriftliches Begründen im Mathematikunterricht. *Zeitschrift für Grundschulforschung*, 7(1), 113–125.
- Pugalee, K. D. (2005). *Writing to develop mathematical understanding*. Norwood, MA: Christopher-Gordon.
- Ruwisch, S. & Neumann, A. (2014). Written reasoning in primary school. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 5, pp. 73–80). Vancouver: PME.
- Schwarzkopf, R. (1999). Argumentation processes in mathematics classrooms. Functional argumentation analysis: A method to describe orally developed arguments. In H.-G. Weigand et al. (Eds.), *Developments in mathematics education in German-speaking countries* (pp. 88–99). Hildesheim: Franzbecker.
- Steinbring, H. (2005). *The construction of new mathematical knowledge in classroom interaction—An epistemological perspective*. Berlin, New York: Springer.
- Toulmin, S. (2003). *The uses of argument* (first published 1958). Cambridge: Cambridge UP.
- Wellington, J. J., & Osborne, J. (2001). *Language and literacy in science education*. Buckingham: Open UP.
- Wittmann, E. C., & Müller, G. (1990). When is a proof a proof. *Bulletin de la Société Mathématique de Belgique Ser. A, 1*, 15–40.