WRITTEN REASONING IN PRIMARY SCHOOL

Silke Ruwisch, Astrid Neumann

Leuphana University, Germany

Currently, language competences in mathematics lessons gain more attention in Germany. The paper reports an interdisciplinary study of linguistics and mathematics education on reasoning. A model to rate the competences in arithmetic reasoning at primary level will be presented for discussion: mathematical reasoning is coded separately from its linguistic realization. In a pilot study, 243 students of 3rd, 4th, and 6th grade solved different arithmetic reasoning tasks. The results show a one-dimensional scale for the model of reasoning. Its specific components provide differentiated requirements, which are formulated concretely in the coding guidelines. They may unfold didactical potential for language support in mathematical reasoning as well as in mathematics lessons itself at primary level.

THEORETICAL BACKGROUND

Reasoning in mathematics and language learning

Mathematical argumentation can be divided into four steps: detecting mathematical regularities, describing them, asking questions about them and giving reasons for their validity (Meyer, 2010; Bezold, 2009). The content base of an argumentation is achieved by description of the detected structures or by reference to common knowledge (Ehlich & Rehbein,1986; Krummheuer, 2000); reasoning then is necessary to acknowledge the described regularities as true (Toulmin, 2003/1958; Schwarzkopf, 1999).

The didactical value of reasoning in mathematics learning is seen in gaining deeper insights into mathematical structures and thereby as a development of one's mathematical knowledge. In this sense, reasoning leads to ask questions about mathematical statements, to make sure they are right and to develop new mathematical connections (Steinbring, 2005). Two intertwined processes may be distinguished: one's own understanding and the process of sharing this understanding with others. Therefore, in its epistemic function mathematical reasoning may be monologic and lead to deeper individual understanding, in its communicative function it is dialogic and depend¬ent on other people if mathematical structures are explained and justified (Neumann, Beier, & Ruwisch, 2014).

Mathematical reasoning in this sense has to be distinguished from reasoning in language classes, especially at primary level. While both are seen as concepts which develop out of situated everyday ("vernacular") speech (Elbow, 2012), reasoning in language learning focusses much more on self-evident facts and personal meanings instead of provable structures in special content areas. So, argumentation in language learning leads to a more addressee-oriented cognitivization (Krelle, 2007); reasoning

2014. In Nicol, C., Oesterle, S., Liljedahl, P., & Allan, D. (Eds.) Proceedings of the Joint Meeting of PME 38 and PME-NA 36,Vol. 5, pp. 73-80. Vancouver, Canada: PME.

in this kind is much more persuasion than proving. Nevertheless, typical linguistic formats of reasoning are learned in these everyday situations and students have to learn how to use them in different content areas. So, in combining the mathematical and the linguistic view on early reasoning, we try to get a broader and deeper understanding of early reasoning, like it can be found in written argumentation of primary students.

Modelling written mathematical reasoning

Although mathematical reasoning is seen as a key issue for students already at the primary level, which for example can be seen in the National Mathematics Standards, there is only few reasoning requested. A textbook analysis showed that not more than 5-10% of all textbook tasks ask for reasoning (Ruwisch, 2012). As well, models which try to describe mathematical competences of this age regard reasoning as important but very specific and classify these competences only to the highest mathematical level (Roppelt & Reiss, 2012). This gap between importance for all and performance of only few was one reason for us to develop a model which may represent different stages of reasoning in early years.

DATA AND METHOD

Sample

The data include 477 written justifications of 243 students. 41 third-graders (♀21; ♂20), 96 fourth-graders (♀43; ♂53) and106 sixth-graders (♀52; ♂54) worked out two out of four designed arithmetic reasoning tasks (s. below).

Arithmetic reasoning tasks

All working sheets are divided into three sections (s. figure 1): In the first section given arithmetic tasks have to be solved and regularities have to be recognized and transferred to more tasks. Following this part of detection, the children are asked to describe their observations, before giving reasons for them.

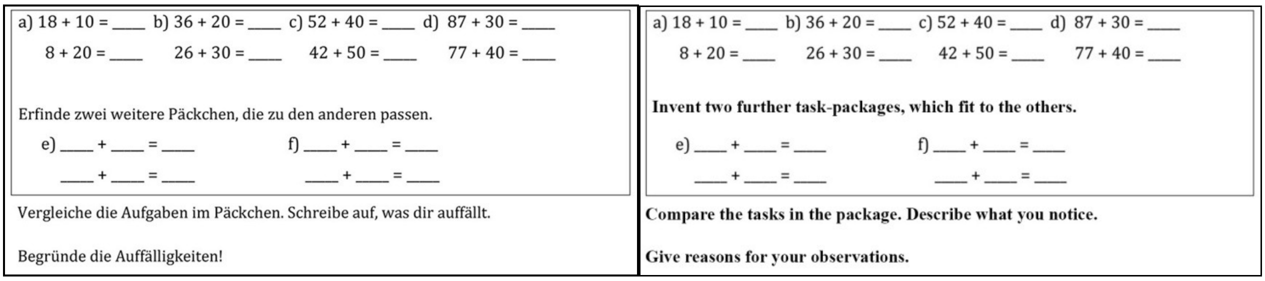

Figure 1: Complex addition tasks (CA) as a sample item.
(on the left: original version; on the right: English translation)

Four different arithmetic tasks were designed for this study. Although the tasks differ in the complexity of regularities, all of them are easy to compute and focus on detection and reasoning. In format ZF three number sequences need to be continued: +9, +7, and +2n. The format EA asks to continue a given additive structure in increasing all three summands by one, so the sum increases by three. In solving formats CA and CM the children need to recognize two structures at the same time. To answer the complex addition task which is given in figure 1 children need to find two tasks with the same sum. At the same time they had to take into account that the summands have to be changed by 10 in opposite directions. The multiplication tasks CM show a constant difference in the product, caused by the difference between the multipliers while the multiplicands remain constant.

Data analysis

Rating scales

Fundamental for our data analysis is the separate evaluation of detecting the mathematical structure and giving reasons for its validity. The argumentation itself is distinguished as well: we separate mathematical from linguistic aspects of reasoning. So, students' writings are rated by one detection-scale and two reasoning-scales (see table 1, explanations below). This separation allows a differentiated grasping for sub-skills of reasoning.

<table>
<tr><th>Mathematical detections</th><th>Mathematical aspects of reasoning</th><th>Linguistic aspects of reasoning</th></tr>
<tr><td rowspan="2">irrelevant aspects as regularities</td><td>regularities (partially) described</td><td>indicators without reason-effect-structure</td></tr>
<tr><td>rudimentary reasoning</td><td>reason-effect structure</td></tr>
<tr><td>regularities partly transferred</td><td>reasoning through examples</td><td>explicit linguistic reference to the task</td></tr>
<tr><td rowspan="2">regularities totally transferred</td><td>partially generalized reasoning</td><td>completeness and consistency</td></tr>
<tr><td>generalization / formal reasoning</td><td>use of math. terminology / decontextualization</td></tr>
</table>

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

Mathematical detections: Children have to compute the arithmetic tasks given on the sheet to find out the underlying structure and transfer it to two more packages with tasks. This process may be realised fully or only partly; sometimes only irrelevant aspects are used to create new tasks. If the structure is transferred fully, the results of the tasks given are also correct, so three stages of this rating scale seem sufficient.

Mathematical aspects of reasoning: Reasoning needs a description of mathematical aspects as basis. If only some regularities are described without giving reasons this leads to stage 1. If a rudimentary reasoning is given despite a description, the work is coded by stage 2. To be rated by stage 3 to 5 all relevant aspects have to gain attention in the argumentation. If this is done by examples, the work is rated by stage 3, if it is already partly generalized, it is rated by 4, and if it is totally general or a formal proof, by 5.

Linguistic aspects of reasoning: The realisation of a mathematical argumentation by written language is also rated by 5 stages which were gained theoretically, especially in focussing on linguistic categories like the use of connectors and identifiable coherence of the text. If explicit linguistic indicators are already used without any structure of reasoning, the text is classified in stage 1. If the text shows a reason-effect-structure it is coded at least as stage 2. If also an explicit linguistic reference to the tasks is visible, the text is classified in stage 3. A text of stage 4 shows a consistent and complete argumentation. To be assigned to stage 5, the use of mathematical terminology must be given in addition, so a decontextualization is identifiable.

Process of coding

14 raters which concentrate either on the mathematical or the linguistic scales were included in the coding process. This process ensured an independent coding by the two professions.

The raters found it easy to code the texts with respect to the detection scale. More difficulties were reported concerning the aspects of reasoning. So the decision between description and rudimentary reasoning was difficult for the mathematical raters. The trade-off between stage 2 and 3 (use of connectors without/with explicit reference to the tasks) as well as between 4 and 5 (use of mathematical terminology) was reported by the linguistic raters as difficult.

Despite the many rater-combinations high absolute agreement in judgments can be reported (62% across all tasks and scales). Deviations of more than one stage occurred in 8% of the cases and showed three important results:

- The multiplication task cannot be compared to the others, because up to now only 35 encodings made by only one pair of raters exist in the data.
- The linguistic scale is the most difficult. Throughout all tasks and raters deviations of more than one stage are observable.
- During the project an increase of coding quality can already be determined. Although acceptable internal consistencies exist across all tasks (Cronbach's α=.80), these values increase, if only ZF (α=.82) and EA (α=.84) which were used later in the project are considered. Nevertheless, large individual deviations can still be observed.

With respect to these results the multiplication task was excluded for the following overall scaling. Thereby, an acceptable average internal consistency of the individual scales over the remaining tasks was achieved: α=.86 for the mathematical detections, α=.81 for the mathematical aspects of reasoning and α=.71 for the linguistic aspects of reasoning.

RESULTS

Due to the great number of rating persons and on the basis of an acceptable inter-rater-consistency (α >.70) we worked on with the means of the ratings for reporting first results.

Overall scale

The IRT-scale of the three tasks and all texts shows a common scale over all components (see table 2). The items are conform to the model as well (WMNSQ .85-1.09). Therefore, early mathematical reasoning in arithmetic like it is measured by the three tasks and the ratings with our scales can be described as a one-dimensional construct.

	Mathematical detections		Mathematical aspects of reasoning		Linguistic aspects of reasoning	
Item	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) compl. addition	-0.845	0.98	0.506	0.92	1.230	0.97

Table 2: Item parameters (Estimate) in IRT scaling.

Looking at the three scales, it becomes obvious 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 scales of reasoning it seems to be easier to realise mathematical aspects of reasoning than to do this in an appropriate linguistic structure. At the same time, *mathematical detections* is the most stable dimension with a maximum difference of .783 compared to 1.446 for the linguistic and 1.516 for the *mathematical aspects of reasoning*.

Comparing the three tasks it seems as if the complex addition is the most difficult to be transferred whereas the simple addition and the number sequences show nearly no difference. The justifications show that it was most easy to realise mathematical as well as linguistic aspects of reasoning in the format number sequences, followed by the complex addition and then by the simple addition task. Despite these differences, all tasks can be characterized as well suited to capture mathematical reasoning in arithmetic.

Students' performances

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

All scores were transformed onto a scale with the mean of 100 and a standard deviation of 20 to make comparison between the three groups of students easier (see figure 2): 3rd graders (M=102/SD=29), 4th graders (M=98/SD=19) and 6th graders (M=101/SD=17) showed nearly the same mean performance.

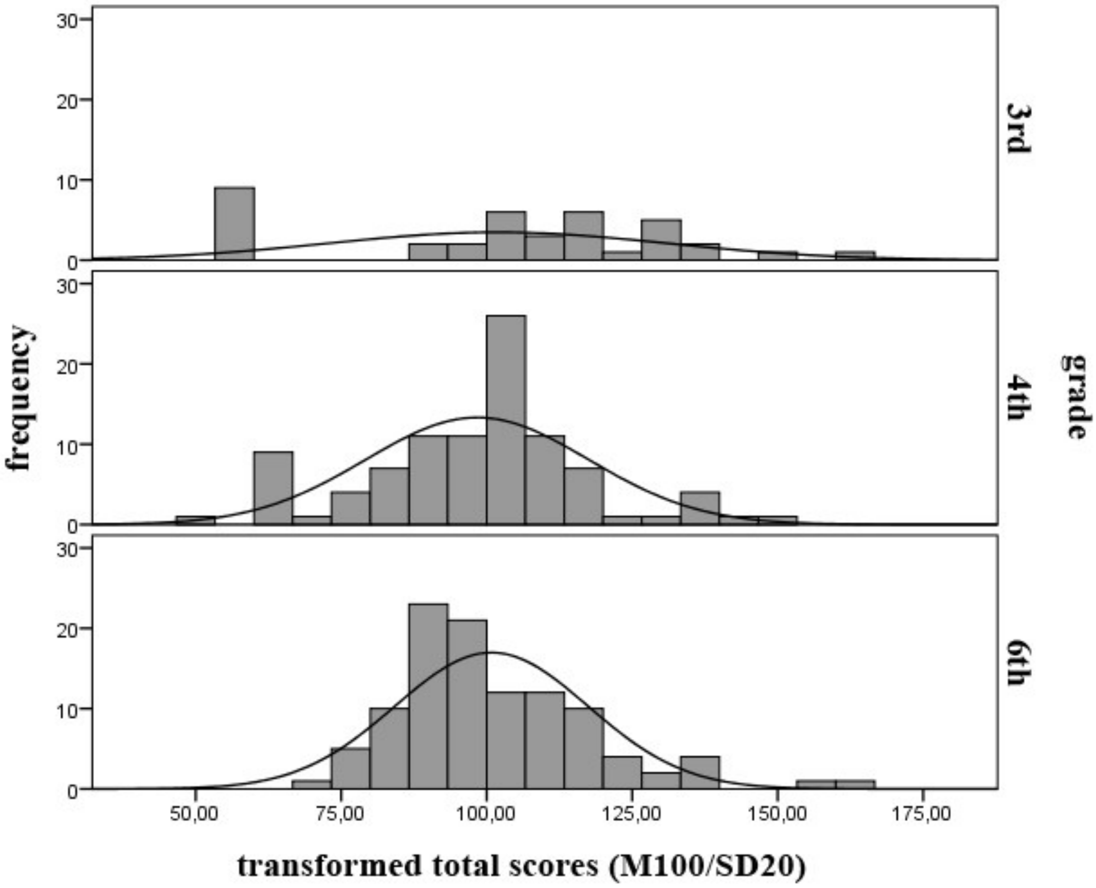

Figure 2: Students' performances by different grades.

Unexpectedly, reasoning competences as they were measured by our tasks and ratings do not increase over time. Even though our data were collected cross-sectionally and not longitudinally a significant increase of competences could have been expected. In interpreting the differences of standard deviations over the three groups, it seems as if 3rd graders differ more in their results than 4th graders and both more than 6th graders, so a homogenization seems to take place during schooling. But, due to the fact of missing comparative data and the small number of our data this remains speculative at the moment.

CONCLUSIONS

Our aim was to describe and report the competences of primary students in dealing with written arithmetic reasoning tasks by different aspects. The results show on the one hand one consistent scale as a one-dimensional construct from detecting and transferring mathematical structures to mathematical and linguistic aspects of reasoning. This one-dimensional construct confirms the approach of Roppelt and Reiss (2012) who assume that process-oriented mathematical skills at primary level are more or less interwoven, interdependent, and therefore one global construct, which will differentiate in higher mathematics learning.

On the other hand, the detailed descriptions of the three scales allow an awareness of different components of mathematical reasoning which will be missed by only one global scale (Neumann 2013). So, the described stages may help to understand which aspects have to be taken into account to be successful in written arithmetic reasoning tasks.

The internal relationships between mathematical and linguistic requirements in solving written reasoning tasks need further verifications and investigation. For instance, we cannot exclude that the difficulties during the coding process (see above) will have spilled over into the variance of the difficulty gradations in the students' results. It might also be that linguistic aspects of reasoning are such difficult, because students do not expect them in mathematics classes. This effect may be reinforced by our anticipation of a very explicit use of "reasoning language" as can be seen in the coding table. So maybe the tasks are too demanding concerning the use of appropriate language to reason in mathematics.

Another critical question concerns the multiplicative task, which did not fit into the model. This may be caused by a too small number of students solving this task (N=35) up to now. But we could also see that a more complex task produces more dropouts as well as more difficulties for the raters. Maybe, the multiplicative task is also too complex to gain information about written reasoning. This may lead to a deeper understanding of the critical aspects of a task to be a "good reasoning task" in mathematics classrooms. High complexity may require too much cognitive and motor capacity to assume a successful writing process (Hayes, 2012). As a consequence, we need more items to check which task is suitable to which function in reasoning processes.

An open question is the stagnation of the students' performance at the level of grade 4. This result may be caused by demotivation, because the sixth graders may think the tasks were too easy to give explicit reasons for the structures. Another argument could be that students still are not used to reasoning in mathematics lessons and competences do not increase by themselves without being taught.

The design of the tasks and the scales of rating show already that written reasoning processes in mathematics at the primary level may be challenged as well as described in more detail then by only a global measure. Hopefully, such interdisciplinary projects help to sharpen the construct and lead to criteria for teachers how to focus on the different aspects of reasoning as well as to unfold didactical potential for language support in mathematics lessons.

References

Bezold, A. (2009). *Förderung von Argumentationskompetenzen durch selbst-differenzierende Lernangebote* (Doctoral dissertation). Hamburg: Kovač.

Ehlich, K., & Rehbein, J. (1986). *Muster und Institution*. Tübingen: Narr.

Elbow, P. (2012). *What speech can bring to writing*. New York: Oxford University Press.

Hayes, J. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369-388.

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.

Meyer, M. (2010). Abduction – A logical view of processes of discovering and verifying knowledge in mathematics. *Educational Studies in Mathematics, 74*, 185-205.

Neumann, A. (2013). Diagnose von Schreibkompetenzen. In J. Fay (Ed.), *(Schrift-) Sprach-diagnostik heute* (pp. 65-83). Baltmannsweiler: Schneider.

Neumann, A., Beier, F., & Ruwisch, S. (2014). Schriftliches Begründen im Mathematikunterricht. *Zeitschrift für Grundschulforschung*, *7*(1), 113-125.

Roppelt, A., & Reiss, K. (2012). Beschreibung der im Fach Mathematik untersuchten Kompetenzen. In P. Stanat, H. A. Pant, K. Böhme, & D. Richter (Eds.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (pp. 34-48). Münster: Waxmann.

Ruwisch, S. (2012). Reasoning in primary school? An analysis of 3rd grade German text-books. In T. Y. Tso (Ed.), *Proc. 36th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, p. 267). Taipei, Taiwan: PME.

Schwarzkopf, R. (1999). Argumentation processes in mathematics classrooms - functional argumentation analysis: A method to describe orally developed arguments. In H.-G. Weigand (Ed.), *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*, *updated edition*. Cambridge: Cambridge University Press.