Cracking percent problems in different formats: The role of texts and visual models for students with low and high language proficiency

Birte Pöhler, Susanne Prediger, Henrike Weinert

Cracking percent problems in different formats: The role of texts and visual models for students with low and high language proficiency

Birte Pöhler, Susanne Prediger and Henrike Weinert

Technische Universität Dortmund, Dortmund, Germany, birte.poehler@math.uni-dortmund.de

Students with low academic language proficiency are often considered to meet specific comprehension challenges with word problems. But how do conceptual and reading challenges interfere in these situations? We approach this question by investigating how performance depends on the problem format for different problem types. A test with N=250 students investigates cracking percent problems in pure, text, and visual format. The results show that text formats are most difficult for elementary problem types, whereas (con-)text can enhance the accessibility for more complex problem types. Item difficulties differ similarly for students with high and low language proficiency, hence reading challenges seem as crucial as conceptual challenges. A deeper analysis shows specific linguistic challenges with the expressions reduced to and reduced by.

Keywords: Percentages, word problems, visual models, language proficiency.

BACKGROUND

Word problems in mathematics tests – reading challenges or conceptual challenges for students with low language proficiency?

Secondary students' academic language proficiency is a crucial factor for their performance in mathematics tests, as has been found in various empirical studies (Abedi, 2006; OECD, 2007; Prediger et al., 2013). Especially in the context of high stakes testing in the US, language biases in mathematical word problems have been investigated for students with low language proficiency (in brief: low LP) (Abedi, 2006). Many researchers emphasize that students with low LP have specific *reading comprehension difficulties with word problems* (e.g., Duarte et al., 2011, for an overview) whereas test items with less text are often assumed to be "language fairer".

However, at least for the German high stakes test ZP10 NRW, this assumption turned out to apply only partially (Prediger et al., 2014): In nearly all items, language proficient students outperformed their less language proficient peers. But the items in which the former had even more difficulties than expected due to their general performance (the items with significant DIF-values) could *not* be characterized by reading challenges, but by conceptual or process-oriented challenges. We tentatively concluded that reading might not be the main obstacle for students with low LP to crack problems, but their restricted conceptual understanding accumulated in ten years of schooling with language disadvantages. However, the phenomenon requires a deeper investigation.

In many different mathematical topics and for students of all levels of language proficiency, word problems have proved to be more difficult than pure items (e.g., Kouba et al., 1988). For students with low LP, these text formats seem to pose specific challenges (cf. Duarte et al., 2011) that require further exploration. However, some studies (especially in primary schools) have shown that using contexts in problems can also support students' performance since a context can enhance the accessibility of the problem and the underlying mathematical concepts (van den Heuvel-Panhuizen, 2005).

Whereas the role of text and context is discussed incoherently in mathematics education research throughout age levels, there is a consensus on the role of visual models as having the potential to facilitate the accessibility of a test item (shown, for example, in Walkington et al., 2013). One could even assume that

students with low LP do equally well as their more language proficient peers in visually presented items if there were no problems in conceptual understanding, only in text comprehension.

These different considerations motivated our research interest on comparing difficulties in *different problem formats*. We treated it for the exemplary mathematical topic of percentages with the following research questions:

- How do students perform in parallel test items on percentages with text format, visual format and pure format?
- How does the role of texts and visual models in test items on percentages differ for students with high versus low language proficiency?

Cracking percent problems as identifying different problem types

The mathematical topic percentages was chosen because it is important in many everyday contexts, and percent problems in assessments bear various difficulties for students (Parker & Leinhardt, 1995; Prediger et al., 2013). Compared to other areas of arithmetic and proportions, relatively few recent studies exist that explore students' competencies and difficulties, (historical exceptions are named in Parker & Leinhardt, 1995; recent exceptions are Dole et al., 1997; Jitendra & Star, 2012 and Walkington et al., 2013).

Typical for percent problems is that students' mathematizing process is shaped by one core step, *identifying the problem type* (Dole et al., 1997). Classically, *three elementary problem types* are distinguished (ibid., with different names): 'Find the amount (if rate and base are given)', 'Find the rate (if amount and base are given)', and 'Find the base (if amount and rate are given)'. Empirical studies show different success rates for different problem types, often 'Find the amount' is easier than the two others (e.g., Kouba et al., 1988, p. 17), and this type being overgeneralized to 'Find the base'.

Beyond these three elementary problem types, *more complex problem types* pose even bigger challenges for students, for example 'percentage growth', 'percentage comparison' or 'Find the base after reduction (if discount and reduced amount are given)' (Parker & Leinhardt, 1995, p. 439). These complex problem types bear reading challenges as well as conceptual challenges and are therefore interesting to compare to elementary types in this study.

Existing empirical studies have compared students' performances on percent problems mainly with respect *to problem types* (Kouba et al., 1988; Dole et al., 1997). In contrast, the comparison of *problem formats* have been less considered (an exception is Walkington et al., 2013). Furthermore, little is known on difficulties with percent problems of students with varying language proficiency. Especially the more complex problem types seem to pose additional comprehension challenges that are worth being considered in more detail. These indications and the limited state of research on percent problems suggest the following additional research question:

- How successful do students of low / high language proficiency identify different percentage problem types? What supports them for more complex problem types?

RESEARCH DESIGN

The study presented in this paper was conducted as a mixed methods study with a paper and pencil test on percent problems for N=250 students (age 13 to 15) in two countries, Germany and German-speaking Switzerland, and complementary interviews. Here, we mainly focus on the tests.

Test design

Language proficiency. Students' language proficiency was assessed by a C-Test in German, an economical and reliable measure of a complex construct of general language proficiency of first- and second-language learners (Grotjahn et al., 2002). By its specific construction in a gap text receptive and productive skills in lexical and grammatical areas are addressed.

Three problem types for percent problems. Students' varying performances for percent problems were measured by a paper-and pencil test with 15 items, systematically constructed in three selected problem types 'Find the amount', 'Find the base' and 'Find the base after reduction' (Table 1 shows the 14 items, some taken from Hafner, 2012, which are relevant for this paper). The problem type 'Find the rate' was omitted as it is the easiest to distinguish from the others by merely considering the involved units.

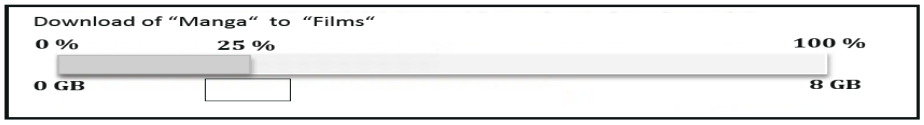
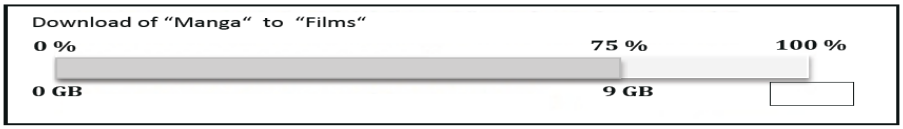
Item set for problem type "Find amount"		Frequencies
Pure format	(Item Find the Amount) What is 5 % of 400 €? Find the amount.	78 %
Visual format	(Item Bar Amount) How many GB have already been downloaded? Find the missing value. 	67 %
Text formats	(Item 1 Potatoes) Potatoes consist of 75 % water. How much water (in g) is contained in 1000 g potatoes? (Item 2 Foundation) A school transfers 60 % of the revenue earned with a school celebration to the "Aktion Mensch" (a foundation). The revenue was 1400 €. How much money does the school transfer? (Item 3 Sport Festival) 30 % of the revenue won at a sport festival in the context of a tombola, being at the amount of 700 €, were dedicated to philanthropy. How much was the donation?	46 % I1: 33 % I2: 61 % I3: 45 %
Item set for problem type "Find base"		
Pure format	(Item Find the Base) 30 % are 60 €. Find the base.	67 %
Visual format	(Item Bar Base) What is unknown here? Find the missing value. 	67 %
Text formats	(Item 4 Jeans) Mr. Koch pays 40 € for jeans in the summer sale. The jeans were reduced to 80 %. How much did the jeans cost before? (Item 5 Kitchen) When buying a new kitchen, Family Mays receives a discount of 250 €, that was 5 % of the regular price. What is the normal price of the kitchen? (Item 6 Holiday Trip) Mrs. Fuchs has prepaid 40 % of the price for her holiday trip. These were 800 €. How expensive is the trip?	53 % I4: 34 % I5: 62 % I6: 63 %
Item set for problem type "Find base after reduction"		
Pure format	(Item Find the Base after reduction) Calculate the former price (base). New price: 750 € Discount: 25 %	44 %
Visual format	(Item Bar Reduction) What is unknown here? Find the missing values. 	59 % 67 % for percent, 51 % for base
Text formats	(Item 7 Dress) Mrs. Schmidt pays 30 € for a dress in the summer sale. The dress was reduced by 40 %. How much did the dress cost before? (Item 8 Cross trainer) A customer buys a cross trainer in a shop. She pays 450 € for the equipment. As she is a member of a sports club, she receives a discount of 10 %. What is the normal price of the cross trainer?	41 % I7: 39 % I8: 43 %

Table 1: Item sets in three different formats for three problem types (translated) with frequencies of correct identification and mathematization in the whole sample

Three problem formats for percent problems. Each problem type was presented in three formats: For the *pure format*, exercises were given together with the technical terms (hence the decision of problem type is already explicit, see Table 1). The *visual format* roused the bar model, an established visual model for percentages (van den Heuvel-Panhuizen, 2003), here contextualized in download bars, a familiar everyday context for teenagers (Prediger, 2013). Three or two items for each problem type were constructed in *text formats* with varying language difficulties.

Sampling and subsampling

The sample consisted of 15 classes in 7 schools, in sum N = 250 students (age 13 - 15). In order to investigate robustness of the findings to national curricular specificities, the test was conducted in two countries, Germany (in the metropolitan Ruhr region) and Switzerland (in the German-speaking region of Lake Constance). Due to differences in the school systems (streaming in Germany since Grade 5 versus a more comprehensive Swiss system), the selected Swiss schools have a higher general achievement level. For investigating differences in students' achievements with varied language proficiency (LP), the sample was split into three groups with high, medium and low LP (the medium group is not considered here); the cut-offs being set by the standardized test norms.

Complementary clinical interviews

In order to deepen the insights into students' difficulties and to gain explanations for difficulties, the quantitative data were triangulated by a small interview study in which students solved the problems in a thinking aloud format. So far, we analyzed interviews with forty students, in sum 594 minutes of video, which were completely transcribed. Due to space restrictions, we only briefly refer to the results of these analyses for selectively strengthen possible explanations even if there is only limited space for transcripts.

Data analysis and hypothesis

For the written tests, an evaluation of success in problem type identification and mathematization for each item was binary-coded in order to allow hypothesis-testing on achievements for different formats etc. with t-tests (Davison, 2003). More precisely, the following hypotheses were tested through attempting to falsify the corresponding null hypotheses:

(H1) Problems in text format are more difficult than in pure format due to comprehension difficulties for word problems (Kouba et al., 1988).

(H1*) Problems in text format are easier than in pure format since contexts can enhance students' accessibility of the problem (as shown for elementary arithmetic problems by van den Heuvel-Panhuizen, 2005).

(H2) Problems in visual format are easier than in text and pure format since visual models can enhance the accessibility of the problem (Walkington et al., 2013).

(H3) Students with low language proficiency have difficulties with *other* problem formats than students with high language proficiency; especially they have specific difficulties with problems in text format.

For testing hypothesis (H3), the data were treated in a Rasch-Model for identifying differential item functioning with respect to the students' language proficiency (Fischer & Molenaar, 1995). Additionally, a categorization of students' written solutions allowed deeper insights into students' challenges and resources, shown here for two similar items (in Section Reduce to versus reduce by: An example for linguistic challanges).

Language-driven subsampling **Regional subsampling**	Whole sample	Subsample with low LP	Subsample with high LP
Whole sample	N = 250	n = 60	n = 84
German Subsample	n = 98	n = 25	n = 33
Swiss Subsample	n = 152	n = 35	n = 51

Table 2: Sample and subsamples

Problem type Problem format	Find amount			Find base			Find base after reduction		
	Pure	Visual	Text	Pure	Visual	Text	Pure	Visual	Text
Whole sample	78 %	67 %	46 %	67 %	67 %	53 %	44 %	59 %	41 %
German subsample	70 %	60 %	36 %	51 %	47 %	40 %	15 %	38 %	21 %
Swiss subsample	82 %	72 %	53 %	78 %	80 %	61 %	63 %	73 %	53 %

Table 3: Frequencies of correct identification / mathematization in two regional subsamples

RESULTS

First results and their discussion

For each problem format, Table 3 shows the frequencies of successful problem type identification and mathematizations (here interpreted as finding adequate expressions but not necessarily successful calculations of the result). Frequencies are given for the whole sample and split for the German and Swiss subsample.

These empirical results show that *hypothesis (H1)* must be restricted to well known problem types: Problems in text format are significantly more difficult than in pure format only for the basic problem types "Find the amount" and "Find the base" ($p < 0.0009$). In contrast, the ranking of difficulty differs for the more complex and less acquainted problem type "Find the base after reduction" (no significance for higher difficulty with $p=0.175$). In the German subsample, this problem type is marginally easier in text format than in pure format ($p=0.097$), hence hypothesis ($H1^*$) tend to apply. The interviews strengthen our interpretation that the context of shopping discount can enhance the accessibility, and this role of context can compensate potential comprehension problems posed by the text format in this less known problem type (at least for the subsample of high language proficiency, see below). This is illustrated by the written comment of a student (see Figure 1) referring to Item Dress (printed in Table 1).

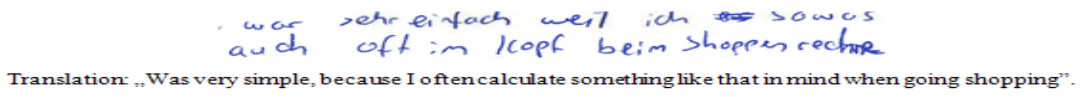


Figure 1: Statement of a student referring to Item Dress

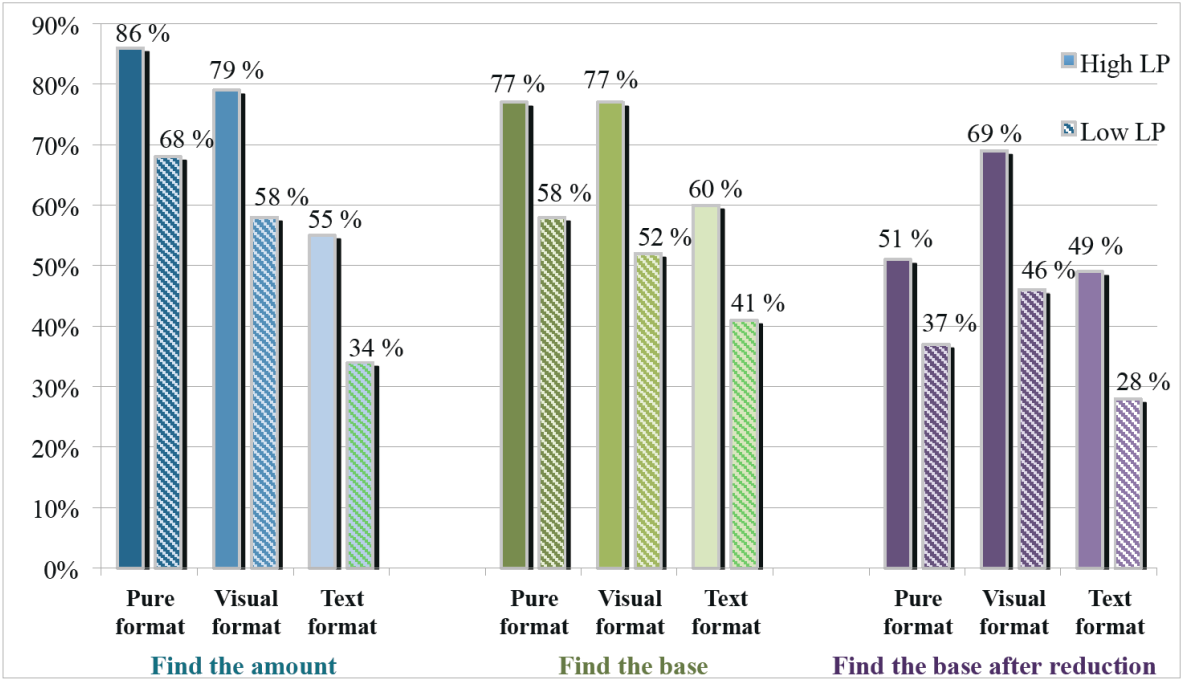


Figure 2: Frequencies of correct mathematization in subsamples with different language proficiency

In contrast, the Swiss subsample, which was more acquainted with the problem type, could solve the pure format more frequently than the text format.

Hypothesis (H2) on the role of the visual format is confirmed for all problem types for the text format: In both subsamples, problems with the visual bar model were solved significantly more frequently than in text format ($p < 0.0009$). In contrast, the visual format is easier than the pure format only for the less known problem type "Find the base after reduction" in both subsamples ($p < 0.021$). Also the interviews show that the visual bar model can enhance students' access to the problem when the key concepts do not suggest routine solutions.

Except for one deviance (text format for less known problem type "Find base after reduction"), the general pattern of difficulties between problem formats seem to be parallel in the German and Swiss subsamples. Unexpectedly, the mentioned differences are even significant in the separate subsamples (except for (H2) in the German subsample).

As Figure 2 shows, the described tendencies seem to apply for students with low as well as with high language proficiency in similar ways: the difference between pure format and text format is very similar for "Find the amount" (31 percent points for high LP and 34 percent points for low LP; of course with a higher rate) and "Find the base" (17 percent points for both samples).

These rather heuristic comparisons are confirmed by statistically more elaborate methods: The estimated item-difficulties in the Rasch-model had no differences for both subsamples, low LP and high LP. Significant differential item functioning was only found with respect to students' mathematics achievement, but *not* between the language proficiency subsamples. Hence, LP does not seem to determine specific difficulties. As a consequence, *hypothesis (H3)* can partly be falsified: Students with low language proficiency perform lower in all items, but their weaker achievement in the text formats is not per se an evidence for reading challenges being most dominant. In contrast, the differences between the formats of the problems proceeds in a related way for students with low LP as for students with high LP. The visual format which might be assumed to be relatively less difficult for students with low LP shows similar differences to the pure format for "Find the amount" (7 percent points for high LP and 10 percent points for low LP) and "Find the base" (0 percent points for high LP and 6 percent points for low LP). These comparisons of formats allow us to conclude that not the language alone, but the conceptual understanding (needed when the item format does not betray the problem type) is the highest difficulty for students with low LP.

This conclusion is strengthened by the analysis of the interviews, from which we show only one singular example, Tom's way of solving the Item Cross trainer (cf. Table 1). The student with low LP makes evident his conceptual understanding when subtracting the number indicating the percentage from the price (#2). He even validates and corrects his solution (#6 / 8), but still with no understanding that Euros and percent cannot be combined directly.

2	Tom:	Well, first, the client buys a crosstrainer in a sport shop. Yes, she pays 450 € for the equipment. That is the price which – how much it costs. She receives a discount of 10 % because she is a member. And now, we should calculate the normal price.... of the cross trainer. And I received: 440. And I have calculated 450 minus 10 %.
...	...	...
6	Tom:	But actually, the normal price should be [reads the text again] should be higher.
...	...	...
8	Tom:	Yes, now I have 460.

Like Tom, many students with low LP succeed in understanding the situation in the text but have too limited conceptual understanding to mathematize correctly. As a consequence, not only problems in text format are difficult for them, but all problems in which the mathematization is not pre-given by technical terms (like in pure format).

Reduce to versus reduce by: An example for linguistic challenges

Even if conceptual challenges are most crucial for students with low LP, there exist also linguistic challenges. We give an example from the deeper analysis of two similarly formulated items, both in text format:

Students' identifications of the problem types "Find the base" (Item Jeans) versus "Find the base after reduction" (Item Dress)	Frequency in *whole sample* (N=250)	Frequency in *subsample low LP* (n=60)	Frequency in *subsample high LP* (n=84)	Significance of differences low-high LP
recognizing a difference	40 %	36 %	49 %	p=0.049
(*both* problem types correctly identified)	14 %	7 %	21 %	p=0.004
(*one* problem type correctly identified)	6 %	7 %	4 %	n.s.
(*no* problem type correctly identified)	20 %	22 %	24 %	n.s.
recognizing no difference	59 %	65 %	52 %	p=0.049
(*one* problem type correctly identified)	39 %	35%	37 %	n.s.
(*no* problem type correctly identified)	6 %	12 %	5%	n.s.
(*both* items not treated)	14 %	18 %	10 %	n.s.

Table 4: Identifying the difference between Item Jeans ("reduced to") and Item Dress ("reduced by")

Item Jeans (Find the base): 40 € for jeans, were *reduced to* 80 %. What did they cost before?

Item Dress (Find base after red.): 30 € for dress, was *reduced by* 40 %. What did it cost before?

The prepositions in the expressions "reduce to" versus "reduce by" determine the problem type, and the typical challenge to recognize this difference (as discussed by Parker & Leinhardt, 1995, p. 439). Table 4 shows students' mathematizations for both items.

Only 40 % of the students have identified different problem types and have hence recognized a difference between the two items, 36 % of the students with high LP and 49 % of those with low LP (the difference being significant with p =0.049). In contrast, 59 % have not recognized the difference while treating the items (65 % with low LP and 52 % with high LP again a significant difference with p =0.049).

However, the fact that only 14 % identified both problem types correctly show that al-though there is a significant linguistic challenge that requires attention in classrooms, the conceptual challenges are still virulent. This can again be illustrated by Tom's way of solving the Item Jeans (cf. Table 1):

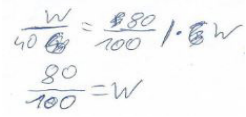


Figure 3

1 Tom: [after reading and calculating]: Yes okay, okay the amount is to be found, short W.

CONCLUSION AND OUTLOOK

The results of the presented test with N=250 students confirm that the same problem formats can provide different challenges: students have most difficulties in cracking percent problems in text format compared to those in pure or visual format. Whereas for elementary problem types, items in pure format are solved better than those in visual format, the visual model seems to enhance the accessibility for more complex, more unknown problem types. A methodological limitation of the study is that we did not account for the specific classroom curricula, which might influence students' varied abilities.

Although the language proficient students outperformed the students with low language proficiency in all items, the general pattern of differences in item difficulties is similar for both groups. This contradicts commonly held assumptions that disparities are bigger for text formats than for pure formats. It suggests that not exclusively the students' restricted reading proficiency is responsible for difficulties in tests, but also their lacking conceptual understanding in percentages (similar in grade 10, cf. Prediger et al. 2014).

In the future research, the findings will be extended by (1) an extended sample, (2) by deeper qualitative insights into students' difficulties by the interview study, and (3) by investigating whether these findings also apply to other problem formats.

Already the current state of the results is taken into account when designing a remediating course for enhancing students' conceptual understanding for percentages and dealing with word problems in this topic.

ACKNOWLEDGEMENT

We thank Esther Brunner for collecting data in Switzerland.

REFERENCES

Abedi, J. (2006). Language issues in item-development. In S. M. Downing & T. M. Haldyna (Eds.), *Handbook of test development* (pp. 377–398). Mahwah, NJ: Erlbaum.

Davison, A.C. (2003). Statistical Models. Cambridge, UK: Cambridge University Press.

Dole, S., Cooper, T.J., Baturo, A.R., & Conoplia, Z. (1997). Year 8, 9 and 10 students' understanding and access of percent knowledge. In A. Begg (Ed.), *People in mathematics education. Proceedings of 20th MERGA* (pp. 7–11). Rotorua, New Zealand: Merga.

Duarte, J., Gogolin, I., & Kaiser, G. (2011). Sprachlich bedingte Schwierigkeiten von mehrsprachigen Schülerinnen und Schülern bei Textaufgaben. In S. Prediger & E. Özdil (Eds.), *Mathematiklernen unter Bedingungen der Mehrsprachigkeit* (pp. 35–53). Münster, Germany: Waxmann.

Fischer, G.H., & Molenaar, I.W. (1995). *Rasch models: foundations, recent developments and applications*. New York, NY: Springer-Verlag.

Grotjahn, R., Klein-Braley, C., & Raatz, U. (2002). C-Test: an overview. In J. Coleman et al. (Eds.), *University Language Testing and the C-Test* (pp. 93–114). Bochum, Germany: AKS.

Hafner, T. (2012). *Proportionalität und Prozentrechnung*. Wiesbaden, Germany: Vieweg/Teubner.

Kouba, V., Brown, C., Carpenter, T., Lindquist, M., Silver, E., & Swafford, J. (1988). Results of 4th NAEP Assessment of Mathematics. *Arithmetic Teacher, 35*(8), 14–19.

Jitendra, A. & Star, J. (2012). An exploratory study contrasting high- and low-achieving students' percent word problem solving. *Learning & Individual Differences,* 22, 151–158.

OECD (2007). *PISA 2006*. Vol. 2: Data. Paris: OECD.

Parker, M., & Leinhardt, G. (1995). Percent: A Privileged Proportion. *Review of Educational Research,* 65(4), 421–481.

Prediger, S. (2013). Focusing structural relations in the bar board. In B. Ubuz et al. (Eds.), *Proceedings of CERME 8* (pp. 343–352). Ankara, Turkey: METU University.

Prediger, S., Renk, N., Büchter, A., Gürsoy, E., & Benholz, C. (2013). Family background or language disadvantages? Factors for underachievement in high stakes tests. In A. Lindmeier & A. Heinze (Eds.), *Proceedings of 37th PME* (vol. 4, pp. 49–59). Kiel, Germany: PME.

Prediger, S., Wilhelm, N., Büchter, A., Benholz, C., & Gürsoy, E. (2014, submitted). *Sprachkompetenz und Mathematikleistung*. Submitted manuscript.

Van den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics*, 54, 9–35.

Van den Heuvel-Panhuizen, M. (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics*, 25 (2), 2–9.

Walkington, C., Cooper, J., & Howell, E. (2013). Effects of visual representations and interest-based personalization on solving percent problems. In Martinez, M. & Castro Superfine, A. (Eds.), *Proceedings of 35th PME-NA* (pp. 533–536). Chicago, IL: University of Illinois.