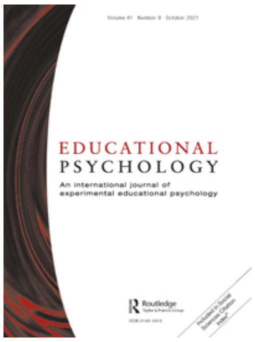
What makes mathematics difficult for adults? The role of reading components in solving mathematics items

Cruz Neri, Nadine; Wagner, Jenny; Retelsdorf, Jan

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article



Routledge
Taylor & Francis Group

Educational Psychology

An International Journal of Experimental Educational Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cedp20

What makes mathematics difficult for adults? The role of reading components in solving mathematics items

Nadine Cruz Neri, Jenny Wagner & Jan Retelsdorf

To cite this article: Nadine Cruz Neri, Jenny Wagner & Jan Retelsdorf (2021) What makes mathematics difficult for adults? The role of reading components in solving mathematics items, Educational Psychology, 41:9, 1199-1219, DOI: 10.1080/01443410.2021.1964438

To link to this article: https://doi.org/10.1080/01443410.2021.1964438

EDUCATIONAL PSYCHOLOGY
2021, VOL. 41, NO. 9, 1199–1219
https://doi.org/10.1080/01443410.2021.1964438

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

What makes mathematics difficult for adults? The role of reading components in solving mathematics items

Nadine Cruz Neri[a], Jenny Wagner[b] and Jan Retelsdorf[a]

[a]Faculty of Education, University of Hamburg, Hamburg, Germany; [b]Institute of Psychology, University of Hamburg, Hamburg, Germany

ABSTRACT

Although it is widely known that reading plays a significant role in mathematics performance, it remains unclear how specific reading component skills and item characteristics are associated with adults' mathematics performance. The aim of this study was to investigate reading component skills (printed vocabulary, sentence processing, passage comprehension), characteristics of mathematics items (picture/table, complex verb forms, number of prepositions, lexical density), and the possible interaction effects thereof on mathematics performance. The sample consisted of 368 German adults (age: $M = 50.45$; 59% female). Our results showed main effects on performance from adults' reading component skills and item characteristics. Furthermore, interaction effects between passage comprehension and two item characteristics on mathematics performance emerge. Implications include enhancing adults' reading component skills.

ARTICLE HISTORY
Received 13 May 2020
Accepted 2 August 2021

KEYWORDS
Reading components; item characteristics; mathematics; adult education

The role of reading in mathematics education has gained much attention in the recent decades (Adams, 2003; Pimm, 1987). Mathematical literacy not only enables performing mathematical deductions, but also fully participating in society and engaging in adult life (Burton & Morgan, 2000). While mathematics becomes more important in modern societies, all countries participating in the Programme for International Assessment of Adult Competencies study (PIAAC; OECD, 2016) have a considerable number of adults with mathematical deficits (Grotlüschen et al., 2016; Parsons & Bynner, 2005).

One requirement for succeeding in mathematical contexts are reading component skills (RCS), due to the verbal presentation of mathematics items (Schleppegrell, 2007). Furthermore, item characteristics may also affect mathematics performance (Abedi & Lord, 2001) and initial studies have shown that RCS and item characteristics may interact (Kulesz et al., 2016). Therefore, investigating item characteristics and their

CONTACT Nadine Cruz Neri nadine.cruz.neri@uni-hamburg.de
Supplemental data for this article can be accessed here.

interaction effects with RCS seems promising as a means to understanding how specific item characteristics become obstacles to learning mathematics.

School-based research has illustrated the association between reading and mathematics (Wu, 2010). Considering recent data on the sizeable number of adults without proficient RCS and mathematical skills (Grotlüschen et al., 2016), further research is needed to understand the relationship between RCS and mathematics performance beyond the school age. Given the lack of research in adulthood, the theoretical and empirical deductions for our research questions are based on the research on students. One could assume that the findings for students and adults will not differ significantly, since the processes in solving mathematics items should be similar. However, this assumption is not unequivocally clear, as the relevance of mathematics items is generally made much more explicit in school, compared to the daily life of adults. Therefore, students may approach mathematics items differently in school, compared to what adults would do in real-life contexts. Furthermore, we aim to go beyond previous research by investigating the interaction effects of person and item characteristics in more detail, given that studies on this matter have been sparse.

To address these outlined lacunas, we aimed to investigate the relation between adults' mathematics performance, their RCS, and the characteristics of mathematics items. We investigated the main effects of both RCS and item characteristics, and the interaction effects between these two variable types on mathematics performance, while controlling for gender and migratory background. We hypothesised that especially those participants with low RCS may face difficulties, resulting in lower mathematics performance, when confronting specific item characteristics.

The importance of RCS for mathematics performance

Bergqvist et al. (2018) nicely illustrate in their simple theoretical model that reading and mathematical skills partially overlap when solving mathematics items. Mathematics items are often verbal descriptions of mathematical problems, where a question demands deducing mathematical operations from the numbers provided in the item (Verschaffel et al., 2000). Solving such items involves many (mathematical) processes, called modelling processes (for an overview, see Borromeo Ferri, 2006). These processes include building a mental representation of a mathematical problem out of reading and comprehending written text (Hayes & Simon, 1974); these reading processes are the focal point of our study. Both reading and mathematical skills are essential for solving mathematics items (Bergqvist et al., 2018). Importantly, reading occurs at different reading levels. Based on the construction-integration model, reading (comprehension) involves the identification of words, the decoding of the semantics, and the connection of text elements into a mental representation of the written item (Kintsch, 1988; van Dijk & Kintsch, 1983). These different reading levels are, however, not addressed in the operationalizations of modelling processes: Mostly, it is only mentioned that reading is a crucial component in building an adequate mental representation of a mathematics item and the stated problem (Borromeo Ferri, 2006).

Following these assumptions of different reading levels, empirical research distinguishes RCS at three levels: vocabulary at the word level, sentence processing (SP) at the sentence level, and passage comprehension (PC) at the text level (O'Reilly & McNamara, 2007). Regarding vocabulary, studies have supported the importance of students' general receptive vocabulary. It has been found that it predicts mathematical vocabulary (Powell et al. 2017), which in turn affects mathematics performance, because knowledge of mathematical terms is needed to understand and solve mathematics items (Paetsch et al., 2015; Peng & Lin, 2019). However, little research has been conducted to investigate direct effects of general receptive vocabulary on mathematics performance, especially for adults.

On the basis of the construction-integration model (van Dijk & Kintsch, 1983), it has been argued that general vocabulary (GV) forms the foundation of RCS at the text level (Cain et al., 2004; Nalom et al., 2015). Less skilled readers usually have smaller vocabularies and tend to show many difficulties (Davis, 1968; Thorndike, 1973). Although a large vocabulary is argued to facilitate reading comprehension at higher reading levels (National Reading Panel, 2000), vocabulary alone is not sufficient for effective reading comprehension (Cain et al., 2004) and for building an appropriate mental representation of an item. In line with the construction-integration model (Kintsch, 1988), readers must integrate vocabulary into a coherent sentence and interpret it in the light of the text to ultimately get the central idea of an item (Perfetti, 1985).

Contrary to the research regarding general receptive vocabulary and RCS at the sentence level, research has extensively explored the link between RCS at the text level and mathematics performance. Associations between these skills have been found repeatedly (Aiken, 1972) with correlations varying between .57 and .95 within the school context (Wu, 2010) and between .48 and .86 in adulthood (Parsons & Bynner, 2005). The associations remain, even after controlling for age, IQ (Pitts, 1952), or vocabulary and grammatical skills (Paetsch et al., 2015). While a few studies have further found that mathematical skills predict RCS at the text level (LeFevre et al., 2010), most studies focus on the prediction of mathematical skills by RCS at the text level. Cross-sectional studies revealed that the latter directly predicts mathematics performance (Fuchs et al., 2018), grade (Ivanoff et al., 1965) and self-concept (Beal et al., 2010). In fact, it has been noted that initial RCS predict subsequent increase in mathematics performance (Björn et al., 2016). In addition to such direct effects of RCS at the text level on mathematics performance, some studies have also investigated indirect effects. For instance, the effects seem to be further mediated by mathematics-specific language skills (Ufer & Bochnik, 2020) and relational processing (Boonen et al., 2013). All these outlined findings corroborate the basic assumption that reading and mathematical skills partially overlap (simple theoretical framework; Bergqvist et al., 2018). In accordance with multiple modelling processes (Borromeo Ferri, 2006), Leiss et al. (2019) illustrated that RCS at the text level were linked to building an adequate situation model. The authors further found that this link was affected by the linguistic characteristics of mathematics items. Thus, the solving process is affected by RCS at the text level and the mathematics item characteristics.

Drawing on these theoretical perspectives and empirical results, studies showing that non-native speakers perform worse in mathematics than do native speakers (Banks et al., 2016), may be explained by their lower RCS (Paetsch et al., 2015). For instance, low vocabulary may impede processes at the sentence and text level, which in turn may impede mathematics performance. At the same time, there are no indications as to why a certain gender should outperform the other(s). However, research has repeatedly illustrated that females outperform males in reading, while males outperform females in mathematics (OECD, 2019); this may partly be traced back to gender stereotypes (Muntoni et al., 2020) and various cognitive abilities (Zhu, 2007).

In sum, there is strong evidence that reading and mathematical skills overlap partially, aligning with the basic assumption of Bergqvist et al.'s theoretical framework (2018). Furthermore, research indicates that particularly RCS at the text level may be the key process of building a mental representation of a given mathematics item (Leiss et al., 2019). Unfortunately, in many studies, assessing RCS at different reading levels, researchers did not differentiate the effects of RCS, but rather combined their measures into one overall measure (Fuchs et al., 2015; Ufer & Bochnik, 2020). To the best of our knowledge, no previous study has explicitly investigated the effects of RCS at different reading levels on mathematics performance. Knowing that reading occurs at different reading levels (O'Reilly & McNamara, 2007) and that difficulties in mathematics may arise at all reading levels (Gürsoy et al., 2013), research should differentiate more clearly between different RCS. Looking further into RCS at the word, sentence and text levels will increase our knowledge of weak readers (OECD, 2012) and allow us to support them regarding their reading comprehension (in mathematics).

The importance of item characteristics for mathematics performance

The mathematical register, that is, a constellation of linguistic characteristics, characterises language in mathematics (Chan, 2015). The linguistic characteristics of the mathematical register differ significantly from everyday language in several respects, such as the use of technical vocabulary or the higher occurrence of prepositional phrases (Schleppegrell, 2004, 2007). Drawing on the Karlsruhe comprehensibility concept (Göpferich, 2009), the linguistic characteristics of the mathematical register may be part of different comprehensibility dimensions. For instance, the dimension simplicity refers to specific linguistic characteristics at different reading levels.

To build a mental representation of a mathematics item (Borromeo Ferri, 2006), one must decode these item characteristics of the mathematical register by reading to understand the problem stated in the item and to be able to solve the item correctly. School-based studies have repeatedly emphasised the relationship between mathematics performance and these item characteristics, showing that the (verbal) presentation of items may affect students' performance (Abedi & Lord, 2001). For our study, we focus on four item characteristics.

First, research suggests that presenting a picture with an item facilitates reading comprehension through the multimedia representation of a problem (Ainsworth, 2006). Elia and Philippou (2004) proposed, on the basis of the work of Carney and Levin (2002), a classification of pictures in mathematics items. First, decorative pictures serve as decoration, and do not provide additional information to solve the mathematics item. Second, representational pictures illustrate the problem portrayed in the item in part or as a whole. Third, organisational pictures support the solution procedure by providing directions. Finally, informational pictures provide crucial information for solving the item correctly. Studies showed that decorative pictures have no effect on students' performance, while the results for representational pictures are ambiguous, both organisational and informational pictures have positive effects on mathematics performance (Dewolf et al., 2014; Elia & Philippou, 2004).

Second, researchers suggest avoiding complex verb forms (CVF), and using only the present tense and active voice (Göpferich, 2009; Karlsruhe comprehensibility concept). CVF are assumed to be linked to decreased comprehension and performance due to difficulties arising when processing CVF (Berndt et al., 2004). Partly corroborating these assumptions, Shaftel et al. (2006) revealed that CVF (defined as consisting of more than three words in their study) were significantly linked to decreased mathematics performance in fourth graders. However, tenth graders scored significantly higher on items with CVF, which is a surprising finding the authors did not discuss. Regarding passive constructions, Shaftel et al. (2006) did not reveal any effects on students' performance. Banks et al. (2016) illustrated that although passive constructions lowered the chances of solving the mathematics items correctly, passive constructions were not to pose significant challenges for non-native speakers.

Third, researchers highlight the relevance of prepositions and prepositional phrases in mathematics given their prevalence in the field (Prediger et al., 2018). Despite a significant relevance to mathematics, a high number of prepositions in mathematics items often lead to difficulties and to decreased mathematics performance in students (Haag et al., 2013; Shaftel et al., 2006), especially for non-native speakers (Martiniello, 2008).

Finally, lexical density as a measure of information embedded in an item (Johansson, 2009), may also affect mathematics performance: The higher the lexical density in an item, the more information is embedded in it. According to a functional linguistics perspective (Schleppegrell, 2004, 2007), the mathematical register is characterised by higher lexical density compared to everyday language. This may lead to difficulties in reading because these more densely structured items need to be parsed precisely. Similar measures have been investigated in mathematics (lexical diversity: Gürsoy et al., 2013), but lexical density has not. Since high lexical density is considered to be a key characteristic of the mathematical register (Schleppegrell, 2004, 2007), it may be insightful to look further into its effect on mathematics performance.

Although previous research has highlighted strong associations between item characteristics and mathematics performance, further research is needed, since there are close to no studies investigating how item characteristics interact with individuals' RCS. Initial studies showed that particularly non-native speakers tend to show lower RCS and mathematics performance, indicating an interaction effect (Haag et al., 2015; Martiniello, 2008). However, the interaction effects between RCS and item characteristics are still unclear, since most interaction effects have been investigated either at the text level or with individuals of special subgroups, such as non-native speakers (Banks et al., 2016). Furthermore, studies conducted so far have focussed mostly on students. Adults may face similar difficulties, since one could assume that the processes in solving mathematics would be similar. However, this has not been investigated as yet, and for at least two reasons it could be assumed that solving processes differ in individuals of different ages. First, adults might be better able to engage with mathematics items and with the item characteristics that generate difficulties for students. Second, the relevance and explicitness of mathematics in real-life contexts differ from mathematics in school contexts. Therefore, it may be fair to assume that adults will approach mathematics items differently, compared to students.

The present investigation

Building on previous research and recent findings on the worrisome number of adults with low RCS and mathematical skills (Grotlüschen et al., 2016), we aim to get a more differentiated view of interaction effects between these variables and mathematics performance. In our study, we focus on RCS at different reading levels and on how specific characteristics of the mathematical register affect performance. Thereby, our study goes beyond previous research in at least two regards. First, we include several RCS and item characteristics in one study that have previously for the most part been researched in separate studies. Second, we focussed on an adult sample to investigate difficulties in mathematics beyond school, since the relevance and explicitness of mathematics may differ significantly from school to real-life contexts.

Specifically, our study addressed three research questions. First, we examined whether RCS have a favourable effect on adults' mathematics performance. For this, we examined three RCS at different reading levels. Drawing on the construction-instruction model (van Dijk & Kintsch, 1983), we assumed that at all levels, RCS are necessary to build an adequate mental representation of the mathematics items, and hypothesised that higher RCS are associated with higher performance.

Second, we investigated the association between four item characteristics and mathematics performance. Drawing on theoretical frameworks regarding learning and performing with multiple representations (Ainsworth, 2006), we hypothesised that presenting mathematics items with an informational picture or an assistance-providing table is linked to higher performance. Drawing on the Karlsruhe comprehensibility concept (Göpferich, 2009), we hypothesised that the use of CVF is associated with lower performance. In line with empirical results (Shaftel et al., 2006), we hypothesised that participants show lower performance on items with a high number of prepositions. Drawing on the functional linguistic perspective (Schleppegrell, 2004), we assumed that participants score lower on items with high lexical density, since these items are more densely structured, relative to common everyday language.

Third, we examined how interaction effects between RCS and item characteristics are related to mathematics performance. Drawing on the simple theoretical framework (Bergqvist et al., 2018), we assume that participants with lower RCS may be affected differently by the presentation of varying item characteristics. We hypothesised that participants score better on items with informational pictures or tables; this effect being stronger for participants with low RCS. Such pictures may help with processing the item better, as they can be used to make more sense of the problem, which they might have difficulties understanding (Ainsworth, 2006). Furthermore, we hypothesised that participants score lower on items using CVF; this effect being stronger for participants with low RCS. Finally, we assumed that the effects of the number of prepositional phrases and lexical density would be stronger for participants with low RCS. Participants with high RCS should be able to master items with a high number of prepositions and high lexical density, better than participants with low RCS, due to their proficient skills.

Finally, we controlled for gender and migratory background. As noted above, girls outperform boys in reading measures, while boys perform slightly better in mathematics (OECD, 2019). While the gender gap regarding reading seems to diminish in adulthood, men still outperform women in mathematics (OECD, 2013a). Furthermore, we controlled for migratory background, since research has repeatedly illustrated that native speakers outperform non-native speakers in RCS and mathematics assessments (OECD, 2013a).

Methods

Sample

The data stem from PIAAC, conducted by the Organisation for Economic Co-operation and Development (OECD; Rammstedt et al., 2016). More precisely, we used the data from German participants in round one; the data were collected between August 2011 and March 2012 (for a detailed description see OECD, 2016). Participation in the PIAAC was voluntary. In total 5,465 Germans took part in it. However, assessment specificities reduced the number of participants in our study. That is, participants could choose whether they wanted to complete the assessment on computers or on paper. Since RCS were only included in the paper-pencil test, we could only draw on data from the paper-based assessments. Due to the test rotation design of PIAAC (for a detailed workflow, see Zabal et al., 2014), only a small portion of the German sample completed both, the RCS and the mathematics items, resulting in our final sample of $N = 368$ German adults (age: $M = 50.45$, $SD = 11$; 59% female). In the final sample, there were no missing data on the relevant measures for our statistical models.

The majority of these participants were native speakers (81.52%): The participants were considered native speakers in cases where the first language learned was the same as the language the assessment was conducted in (OECD, 2013b). More than half of the participants had a high school qualification (56.3%), while one fifth finished school before qualifying for a high school qualification (21.5%) and one fifth had a degree above high school level (19.6%), respectively. The remaining participants did not provide further details about their education (2.7%).

Measures

We focussed on two measures: reading components and mathematics items. There was no time limit, and using working materials (e.g., calculators) was allowed. For further details about the test administration in PIAAC, see Zabal et al. (2014). For exemplary items of any of the following measures, see OECD (2012) or OECD (2016).

The participants' answers were coded by administrators of PIAAC. Participants received one point for each correct answer and no points for incorrect or missing answers.

Reading components

Three RCS were assessed. For the statistical analyses, we summarised the correct answers for each reading component for every participant.

General vocabulary (GV). This measure comprised 34 items assessing general receptive vocabulary at the word level, and consisting of a picture together with multiple-choice answers. Participants had to choose the word that matched the picture; these pictures depicted familiar and concrete nouns. The assessment of GV was called printed vocabulary in PIAAC, and referred to adults' everyday vocabulary, rather than to academic terms (Grotlüschen et al., 2016). To avoid confusion with mathematical vocabulary, we decided to refer to printed vocabulary as GV. Cronbach's alpha was .92, indicating excellent internal consistency. Participants needed 2.55 minutes ($SD = 1.47$) to complete this measure.

Sentence processing (SP). This assessment consisted of 22 items estimating skills at the sentence level. Participants read a series of sentences, ranging from three to 26 words in length (Grotlüschen et al., 2016), which 'require[d] the respondent to assess whether a sentence [made] sense in terms of the properties of the real world or the internal logic of the sentence' (OECD, 2012, p. 30). Cronbach's alpha was .86, indicating good internal consistency. Participants spent 2.68 minutes ($SD = 1.57$) on this measure.

Passage comprehension (PC). This measure assessed skills at the text level with 44 items. It consisted of four passages of prose texts ranging from 159 to 278 words in length, in two to three paragraphs (Grotlüschen et al., 2016). Participants were 'asked to read a passage and when they [came] to the underlined alternatives, [to circle] the word that [made] sense' (OECD, 2012, p. 30). Hence, these items used a forced-choice close answer format. Cronbach's alpha was .98, indicating excellent internal consistency. Participants needed 7.22 minutes ($SD = 3.87$) to complete this measurement.

Mathematics items

The paper-pencil version comprised 20 mathematics items with open-answer formats. The items are supposed to be 'derived from real-life stimuli and pertain to all types of contexts or situations […] that can be expected to be of importance or relevance' (OECD, 2012, p. 39). Two independent raters coded the item characteristics for each mathematics item. For additional information regarding the contents and coding of the mathematics items, see Text OS1 in the Supplemental Material.

Picture/table

All items were presented with pictures and/or tables. We distinguished between items only presented with a decorative picture ($n = 6$) and items presented with at least one informational picture or one assistance-providing table ($n = 14$). We dummy-coded the variable (0 = decorative picture/no table, 1 = informational picture/assistance-providing table). Hence, the number of pictures and tables was not considered in the statistical analyses. We did not differentiate between informational pictures and tables, since tables provide similar support as do organisational pictures: The readers are presented with a 'useful structural framework' (Carney & Levin, 2002, p. 7), and a lot of information can be put in a table while simultaneously reducing word count. Cohen's kappa was 1.00, indicating a perfect degree of agreement among two raters (Cohen, 1960).

Complex verb forms (CVF)

We differentiated between items using only the present tense ($n = 8$) and items with CVF, such as passive constructions or a shift in tense ($n = 12$). Again, we dummy-coded this variable (0 = present tense, 1 = CVF). Cohen's Kappa was 1.00.

Number of prepositions

The number of prepositions was counted. Cohen's Kappa was .84, indicating an excellent degree of agreement (Cohen, 1960).

Lexical density

We coded the lexical density by dividing the number of content words (e.g., nouns) by the number of total words (Johansson, 2009; Ure, 1971). Cohen's Kappa was .78, indicating a good degree of agreement (Cohen, 1960).

Statistical analyses

We tested our hypotheses by applying multilevel logistic regression models, using the software M*plus* 8.3 (Muthén & Muthén, 1998-2017), due to the structure and hierarchy of the given data. We applied multilevel modelling, since we considered the mathematics items with varying item characteristics as being nested in participants. We applied logistic regression, due to the binary nature of our dependent variable mathematics performance at item level (0 = false answer, 1 = correct answer). It may be important to note that multilevel logistic regression models are equivalent to item response theory models (Kamata, 2001) and may even have advantages over traditional techniques (Reise, 2000).

In all models, the item characteristics were treated as within-level variables (on Level 1). All RCS were handled as a between-level variable (on Level 2), respectively. Furthermore, cross-level interactions between the RCS and the item characteristics were included in the models, to test interaction effects between these factors on mathematics performance. We included two covariates in our models: gender (0 = male, 1 = female) and migratory background (0 = native speaker, 1 = non-native speaker). Finally, all non-binary variables were standardised ($M = 0$, $SD = 1$) for an easier interpretation of the results, and odds ratios were estimated.

Results

Descriptive statistics and correlations can be found in the Supplemental Material (see Text OS2 and Table OS1). In this paper, we focus on the results for the models with covariates presented in Table 1. The results for the models without covariates are presented in Text OS3, Table OS2, and Figure OS1 in the Supplemental Material.

Table 1. Results of the multilevel analyses with RCS and covariates predicting mathematics performance (Unstandardized Parameters).

	GV				SP				PC			
	B	*SE*	*p*	*OR*	*B*	*SE*	*p*	*OR*	*B*	*SE*	*p*	*OR*
Intercept	0.209	.066	<.001		0.243	.063	<.001		0.213	.067	.001	
Level 1 (item characteristics)												
Picture/table	1.535	.054	<.001	4.641	1.501	.044	<.001	4.486	1.510	.060	<.001	4.527
CVF	−1.112	.054	<.001	0.329	−1.112	.043	<.001	0.329	−1.119	.050	<.001	0.327
Number of prepositions	−0.102	.020	<.001	0.903	−0.093	.021	<.001	0.911	−0.091	.020	<.001	0.913
Lexical density	−0.333	.025	<.001	0.717	−0.333	.019	<.001	0.717	−0.334	.024	<.001	0.716
Level 2 (person characteristics)												
RCS	0.094	.047	.047	1.100	0.272	.052	<.001	1.313	0.242	.048	<.001	1.274
Gender	−0.165	.077	.014	0.848	−0.146	.071	.014	0.8640	−0.168	.074	.012	0.845
Migratory background	−0.329	.101	.001	0.720	−0.149	.095	.055	0.862	−0.249	.095	.004	0.780
Cross-level interactions												
RCS × picture/table	0.022	.049	.349	1.022	0.004	.081	.487	1.004	−0.065	.057	.089	0.937
RCS × CVF	0.015	.043	.336	1.015	0.050	.075	.286	1.051	0.026	.050	.251	1.026
RCS × number of prepositions	−0.011	.017	.235	0.989	−0.047	.028	.065	0.954	−0.036	.023	.037	0.965
RCS × lexical density	0.025	.017	.089	1.025	0.048	.035	.100	1.049	0.056	.024	.009	1.058
Variance components												
Mathematics performance	.399	.042	<.001		.332	.036	<.001		.369	.038	<.001	
Picture/table	.008	.005	<.001		.007	.009	<.001		.005	.009	<.001	
CVF	.006	.008	<.001		.009	.010	<.001		.007	.010	<.001	
Number of prepositions	.004	.003	<.001		.002	.003	<.001		.002	.003	<.001	
Lexical density	.002	.003	<.001		.002	.003	<.001		.002	.003	<.001	

CVF: complex verb forms; GV: general vocabulary; RCS: reading component skills; SP: sentence processing; PC: passage comprehension. The variables picture/table (0 = decorative picture/no table, 1 = informational picture/assistance-providing table), CVF (0 = present tense, 1 = CVF), gender (0 = male, 1 = female), and migratory background (0 = native speaker, 1 = non-native speaker) were dummy-coded.

Main effects (level 1 and level 2)

In accordance with our first hypothesis, the analyses showed that all RCS were positively linked to mathematics performance (Table 1; Level 2). Furthermore, the main effects of the item characteristics were significant in all models (Table 1; Level 1), corroborating our hypothesis. While participants showed higher performance on mathematics items presented with an informational picture or an assistance-providing table, the use of CVF was linked to lower performance. Lower mathematics performance was associated with an increasing number of prepositions and lexical density in the items.

Cross-Level interaction effects

The analyses detected no interaction effects with GV nor with SP. Interaction effects between PC and the number of prepositions as well as lexical density were significant. Participants with high PC scored lower on items with an increased number of prepositions compared to participants with low PC. Furthermore, participants with high PC scored higher on items with increasing lexical density compared to participants with low PC.

To illustrate, the significant interaction effects between PC and item characteristics on adults' mathematics performance are presented in Figure 1, which also include the 95% confidence interval (grey line). This type of presentation helps to understand at which levels the RCS is significant (region of significance; cf. Johnson & Neyman, 1936). In every plot, the y-axis depicts the range of values for the conditional slope of mathematics performance on the item characteristic; the x-axis depicts participants' range of values for their PC from -2*SD* to +2*SD*. For instance, the increasing function can be interpreted as follows: With increasing PC, the negative effect of lexical density on performance decreased, resulting in higher mathematics performance.

Discussion

The aim of this study was to investigate the effects of RCS and item characteristics on adults' mathematics performance. We investigated the main effects of RCS and mathematics item characteristics, and how interaction effects between these two variables were related to performance. Due to restrictions in the PIAAC data, the following results and implications could conceivably apply only to the preselected sample with moderately lower reading and mathematical skills (see Text OS2 in the Supplemental Material).

The main effects of RCS and item characteristics

Supporting our hypothesis, our results showed significant main effects of RCS on adults' mathematics performance. While other studies mainly report one measure for students' RCS at one reading level, we analysed the effects of adults' RCS at three reading levels and their importance for solving mathematics items. In accordance with school-based research, we found that adults' GV at the word level (Paetsch et al., 2015) and adults' PC at the text level (Fuchs et al., 2015; Wu, 2010) were positively

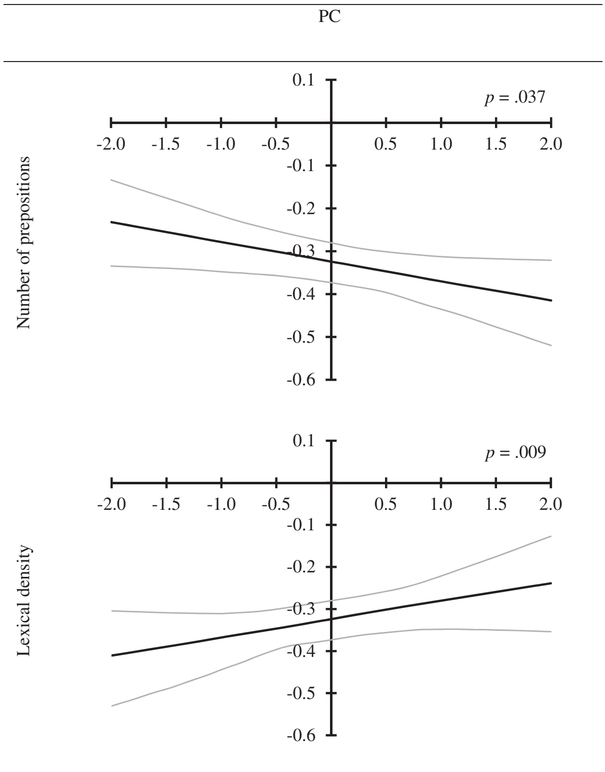


Figure 1. Plots of significant interaction effects.
Note. PC: passage comprehension. The y-axis depicts a range of values for the conditional slope of mathematics performance on the item characteristics, while the x-axis depicts participants' range of values for PC from $-2SD$ to $+2SD$. The plots include the 95% confidence interval.

associated with their mathematics performance. We further analysed adults' SP at the sentence level, finding positive main effects on mathematics performance. Therefore, it seems that high(er) RCS at all reading levels are linked to better mathematics performance in adults.

These results align with the construction-integration model (Kintsch, 1988), which assumes that reading occurs at different levels, all of which need to be processed effectively to comprehend an item. Furthermore, these main effects confirm the theoretical assumption that processing written text is an essential part of building an adequate mental representation of a mathematics item (Borromeo Ferri, 2006), and that reading and mathematical skills overlap (simple theoretical framework; Bergqvist et al., 2018). These results enable us to enrich the research to date in two respects. First, school-based research may be transferred onto adults, as it provides us with indications that the links between adults' RCS and mathematics performance may be similar to those in the student population. Second, considering two common theoretical perspectives regarding the solving of mathematics items (modelling processes: Borromeo Ferri, 2006) and the overlap of reading and mathematical skills (simple theoretical framework: Bergqvist et al., 2018), both only address RCS at the text level. In our study, we were able to indicate that not only RCS at the text level are important for solving mathematics items, but also RCS at the word and sentence levels.

Our hypotheses regarding the item characteristics were also corroborated. Prior research generally confirms that informational pictures presented with mathematics items have a positive effect on students' performance (Elia & Philippou, 2004). According to the conceptual framework for learning with multiple representations (Ainsworth, 2006), presenting an item with a picture tends to facilitate comprehension of a problem presented, because it might help to capture its essence. Therefore, using pictures in mathematics may work as a tool to support individuals in their understanding and solving of mathematics items. Our results support these findings for adults, and reinforce them further by extending results onto the use of assistance-providing tables. That is, our findings emphasise that tables reduce word count substantively and that readers are able to gather information that is relevant to solving the mathematics item, with less text.

Furthermore, the results illustrated that adults experience difficulties with CVF, aligning with the Karlsruhe comprehensibility concept (Göpferich, 2009). School-based research to date has yielded mixed results regarding CVF (Shaftel et al., 2006). In our study, adults seem to suffer from CVF and perform substantially better on items that exclusively use the present tense in an active voice. Moreover, our hypothesis that a high number of prepositions is associated with lower mathematics performance in adults was supported; which is in accordance with school-based research (Prediger et al., 2018; Shaftel et al., 2006). Finally, we were the first to illustrate that mathematics items with high lexical density are linked to lower mathematics performance in adults. In line with a functional linguistics perspective, this could be due to two reasons. First, a mathematics item with high lexical density may be more difficult to solve, since it requires more cognitive resources and processes to extract all the information provided in the item (Fang et al., 2016). Second, the mathematical register differs from everyday registers; lexical density is one of the characteristics that differs significantly between those registers (e.g., Schleppegrell, 2004). Hence, individuals who are less familiar with items of high lexical density may struggle more with the uncommon structure of an item with high lexical density.

In sum, all the investigated item characteristics of the mathematical register in this study are significantly linked to adults' mathematics performance. That is, the items' presentation significantly affects adults' performance. For the most part the effects seem to parallel findings in the school context. This aligns with the basic assumption of the Karlsruhe comprehensibility concept (Göpferich, 2009) that the verbal presentation affects the readability of an item. Interestingly, all these characteristics seem to further explain variance in mathematics performance.

The interaction effects of RCS and item characteristics

No significant interaction effects were detected between adults' GV/SP and item characteristics, indicating that adults' GV and SP do not cohere with adults' skills to deal with linguistic demands in mathematics. However, significant interaction effects were found between adults' PC, the number of prepositions and lexical density. Surprisingly, the negative effect of the number of prepositions on mathematics performance was stronger for participants with high PC. In this case, our hypothesis was not confirmed. Our findings are contrary to our expectations and are difficult to interpret. We did not find any previous research providing clues as to why especially adults with high PC seem to suffer from a high number of prepositions. Past studies in which interaction effects were examined have illustrated that particularly non-native speakers (Martiniello, 2008) and younger students (Shaftel et al., 2006), might suffer from the number of prepositions in items, since their RCS are usually less proficient or still developing. Nor does the functional linguistics perspective provide any plausible explanation for this result (e.g., Schleppegrell, 2004, 2007).

Finally, the results showed that mathematics performance decreased when the lexical density of items increased; this effect was stronger for participants with low PC. Hence, lexical density seems to be particularly difficult for participants with low PC. It seems that it is difficult for these participants to extract the information embedded in a mathematics item with high lexical density to solve it correctly. This may support and enhance the results of Gibson and Levin (1975), who found that students with high RCS are better able to abstract the substance of a written text.

In sum, our study indicates that mathematics item characteristics do not interact with adults' GV and SP, but rather only with their PC at the text level, which places more complex demands on the reader. Hence, for the first time, it has been shown that the assumed interaction effects of individuals' RCS and the item characteristics of written items, emerge only at the text level. It seems as that the interaction effects become more crucial, the higher and more complex the reading level and their associated demands get. Hence, the item characteristics only show interaction effects with adults' skills that affect the understanding and solving of the item as a whole. These results are particularly noteworthy when considering the construction-integration model (Kintsch, 1988): It seems that RCS at the text level are especially relevant for reading in mathematics, since the items' presentation affects individuals differently at different levels of PC.

Theoretical and practical implications

This study contributes to our understanding of the interplay between RCS and item characteristics in affecting adults' mathematics performance. Considering that adults are often confronted with verbal descriptions of mathematical problems in their daily and professional life (Grotlüschen et al., 2016), the effect on mathematics performance becomes relevant. Performance changed significantly depending on the item characteristics, and these effects partially changed according to adults' PC. Hence, mathematics performance is linked to person and item characteristics, which emphasises the interaction between person and item characteristics in solving mathematics items. Our study does not suggest general differences in these associations between adults and students. Considering that large adult populations show low reading and mathematical skills, associated with negative outcomes (Grotlüschen et al., 2016), it becomes obvious that adults need to be given the opportunity to enhance their skills.

We suggest two practical implications regarding the enhancement of individuals' RCS for different age groups. First, handling the demands of the mathematical register is an essential prerequisite to learning mathematics in school (Prediger et al., 2018; Shaftel et al., 2006). Hence, it is crucial to make teachers aware of the relevance of RCS (Boonen et al., 2013) and of item characteristics (Lager, 2006) for mathematics learning and performance. Considering the lasting effects that the quality of schooling has on adults' reading and mathematical skills (Gustafsson, 2016), it is necessary to start interventions in school. Teachers should be made aware of the difficulties that specific item characteristics generate to address these in classes and to guide students in learning to deal with them. Prediger and Wessel (2013), for instance, proposed a language- and mathematics-integrated intervention in which linguistic challenges are explicitly discussed in class. The results of their intervention were promising, as students' conceptual understanding was shown to increase significantly. Since then, more instructional approaches have been proposed to include reading proficiency, and dealing with the demands of the mathematical register (Prediger, 2019).

Second, it is essential to support adults in attaining appropriate reading and mathematical skills, and handling the demands of the mathematical register. For instance, most adults need reading and mathematical skills in both daily and professional life (Grotlüschen et al., 2016). These skills should be explicitly targeted in adult education training programs (Batalova & Fix, 2015), which around half of the adult population already take part in (participation varies between 20 and 66%; OECD, 2014). Regarding our results, it may be particularly effective to increase individuals' RCS at higher reading levels to support them with reading and dealing with mathematics items. However, it is important to point out that simply increasing participants' RCS is not in itself sufficient, since the mathematical register has many features that differ from everyday language and may pose additional challenges (Daroczy et al., 2015). Therefore, it is equally important to guide individuals through the mathematical register and support them in learning to deal with its linguistic characteristics in academic and everyday settings.

Limitations and future directions

As the study draws on secondary data, some limitations need to be considered when interpreting our results. First, the study was not conducted with an experimental design and consequently, causal conclusions cannot be made. Future research should systematically modify item characteristics in mathematics assessments to further investigate the role of RCS in adults' performance, enabling to draw causal conclusions. While there are experimental studies elaborating the role of item characteristics in mathematics items (Haag et al., 2015), these tend to disregard the interaction effects of item characteristics with participants' RCS.

Second, we were restricted in investigating the item characteristics. Hence, some item characteristics that are represented in previous research could not be considered due to the lack of variation. Nonetheless, this study focussed on four item characteristics and managed to investigate their role in adults' mathematics performance in respect of adults' RCS.

Third, due to the test rotation design of PIAAC (Zabal et al., 2014), only a small portion of the German sample completed both assessments (RCS and mathematics items). The preselection resulted in an analytical sample that scored significantly lower in reading and mathematical skills than did the overall German sample. Hence, our results apply to a sample with relatively low RCS and mathematical skills. In future research, it may be reasonable to investigate a larger sample with more variance in the RCS. Large(r) sample sizes also help to identify the small interaction effects that are endemic in the psychological field (Trautwein et al., 2012).

Fourth, we needed to conduct three different models for each RCS instead of one model including all RCS due to multicollinearity. Hence, it remains unclear whether all RCS independently contribute to the explanations of adults' mathematical skills. In future research, it should be examined whether GV and SP uniquely explain variance in adults' mathematical performance.

Finally, there may be ceiling effects for the three RCS. PIAAC initially included the measurement of RCS in the paper-based assessment to get a better understanding of adults with low reading proficiency. Hence, the assessment of RCS was initially specifically developed for adults with low proficiency (Grotlüschen et al., 2016) and was not meant to differentiate adults with average or high reading proficiency. However, most participants working on the reading components were participants who voluntarily opted for the paper-based assessment, rather than adults with low proficiency (Grotlüschen et al., 2016). Given that the assessment of RCS does not differentiate well between performances in the upper range of RCS, the interaction effects may be rather small due to restrictions in the variance of the RCS. A future study should focus on an adult sample with greater variance in RCS to get a better understanding of (interaction) effects for adults.

Conclusion

Our study supports previous findings regarding the relevance of RCS in mathematics. It highlights that both RCS and mathematics item characteristics are associated with adults' mathematics performance. Furthermore, interaction effects between person and item characteristics were detected at the text level. Our study is one of the few that investigates several aspects of RCS and item characteristics, showing where interaction effects emerge between these two factors.

Acknowledgement

This paper uses data from the German part of the first cycle of the Programme for International Assessment of Adult Competencies (PIAAC) conducted by the Organisation for Economic Co-operation and Development (OECD). We would like to thank the OECD for providing the data and German test booklets. The data used in this study are partially openly available online at http://www.oecd.org/skills/piaac/publicdataandanalysis/. Further citations of the data set are included in the manuscript.

The authors thank Stephen McLaren for his editorial support.

Disclosure statement

The authors have no known conflict of interest to disclose.

ORCID

Nadine Cruz Neri http://orcid.org/0000-0002-5414-3854
Jan Retelsdorf http://orcid.org/0000-0002-9892-9613

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219–234. https://doi.org/10.1207/S15324818AME1403_2

Adams, T. L. (2003). Reading mathematics: More than words can say. *The Reading Teacher*, *56*(8), 786–795.

Aiken, L. R. (1972). Language factors in learning mathematics. *Review of Educational Research*, *42*(3), 359–385. https://doi.org/10.3102/00346543042003359

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Banks, K., Jeddeeni, A., & Walker, C. M. (2016). Assessing the effect of language demand in bundles of math word problems. *International Journal of Testing*, *16*(4), 269–219. https://doi.org/10.1080/15305058.2015.1113972

Batalova, J., & Fix, M. (2015). *Through an immigrant lens: PIAAC assessment of the competencies of adults in the United States*. Migration Policy Institute.

Beal, C. B., Adams, N. M., & Cohen, P. R. (2010). Reading proficiency and mathematics problem solving by high school English language learners. *Urban Education*, *45*(1), 58–74. https://doi.org/10.1177/0042085909352143

Bergqvist, E., Theens, F., & Österholm, M. (2018). The role of linguistic features when reading and solving mathematics tasks in different languages. *The Journal of Mathematical Behavior*, *51*, 41–55. https://doi.org/10.1016/j.jmathb.2018.06.009

Berndt, R. S., Mitchum, C. C., Burton, M. W., & Haendiges, A. N. (2004). Comprehension of reversible sentences in aphasia: The effects of verb meaning. *Cognitive Neuropsychology*, *21*(2), 229–244. https://doi.org/10.1080/02643290342000456

Björn, P. M., Aunola, K., & Nurmi, J.-E. (2016). Primary school text comprehension predicts mathematical word problem-solving in secondary school. *Educational Psychology*, *36*(2), 362–377. https://doi.org/10.1080/01443410.2014.992392

Boonen, A. J. H., van der Schoot, M., van Wesel, F., de Vries, M. H., & Jolles, J. (2013). What underlies successful word problem solving? A path analysis in sixth grade students. *Contemporary Educational Psychology*, *38*(3), 271–279. https://doi.org/10.1016/j.cedpsych.2013.05.001

Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *ZDM*, *38*(2), 86–95. https://doi.org/10.1007/BF02655883

Burton, L., & Morgan, C. (2000). Mathematicians writing. *Journal for Research in Mathematics Education*, *31*(4), 429–453. https://doi.org/10.2307/749652

Cain, K., Oakhill, J., & Bryant, P. E. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*(1), 5–26. https://doi.org/10.1023/A:1013176309260

Chan, S. (2015). Linguistic challenges in the mathematical register for EFL learners: Linguistic and multimodal strategies to help learners tackle mathematics word problems. *International*

Journal of Bilingual Education and Bilingualism, *18*(3), 306–318. https://doi.org/10.1080/13670050.2014.988114
Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104
Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H.-C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, *6*, 348. https://doi.org/10.3389/fpsyg.2015.00348
Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly*, *3*(4), 499–545. https://doi.org/10.2307/747153
Dewolf, T., van Dooren, W., Ev Cimen, E., & Verschaffel, L. (2014). The impact of illustrations and warnings on solving mathematical word problems realistically. *The Journal of Experimental Education*, *82*(1), 103–120. https://doi.org/10.1080/00220973.2012.745468
Elia, I., & Philippou, G. (2004). The functions of pictures in problem solving. In: M.J. Hoines, & A.B. Fuglestad (Eds.), *Proceedings of the 28th conference of the international group for the psychology of mathematics education.* (Vol. 2, pp. 327–334). University College.
Fang, Z., Schleppegrell, M. J., & Cox, B. E. (2016). Understanding the language demands of schooling. Nouns in academic registers. *Journal of Literacy Research*, *38*(3), 247–273. https://doi.org/10.1207/s15548430jlr3803_1
Fuchs, L. S., Fuchs, D., Compton, D. L., Hamlett, C. L., & Wang, A. Y. (2015). Is word-problem solving a form of text comprehension? *Scientific Studies of Reading*, *19*(3), 204–223. https://doi.org/10.1080/10888438.2015.1005745
Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., & Martin, B. N. (2018). Text comprehension and oral language as predictors of word-problem solving: Insights into word-problem solving as a form of text comprehension. *Scientific Studies of Reading: The Official Journal of the Society for the Scientific Study of Reading*, *22*(2), 152–166. https://doi.org/10.1080/10888438.2017.1398259
Gibson, E. J., & Levin, H. (1975). *The psychology of reading.* MIT Press.
Göpferich, S. (2009). Comprehensibility assessment using the Karlsruhe comprehensibility concept. *The Journal of Specialised Translation*, *11*, 31–53.
Grotlüschen, A., Mallows, D., Reder, S., Sabatini, J. (2016). *Adults with low proficiency in literacy or numeracy* (OECD Education Working Papers, No. 131). Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/docserver/5jm0v44bnmnx-en.pdf?expires=1587106017&id=id&accname=guest&checksum=C1C45CB15C247A20804F9986504BD255
Gürsoy, E., Benholz, C., Renk, N., Prediger, S., & Büchter, A. (2013). Erlös = Erlösung? – Sprachliche und konzeptuelle Hürden in Prüfungsaufgaben zur Mathematik [Language- and concept-induced obstacles in mathematics exams]. *Deutsch Als Zweitsprache*, *13*(1), 14–24.
Gustafsson, J. E. (2016). Lasting effects of quality of schooling: Evidence from PISA and PIAAC. *Intelligence*, *57*, 66–72. https://doi.org/10.1016/j.intell.2016.05.004
Haag, N., Heppt, B., Roppelt, A., & Stanat, P. (2015). Linguistic simplification of mathematics items: Effects for language minority students in Germany. *European Journal of Psychology of Education*, *30*(2), 145–167. https://doi.org/10.1007/s10212-014-0233-6
Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction*, *28*, 24–34. https://doi.org/10.1016/j.learninstruc.2013.04.001
Hayes, J. R., & Simon, H. (1974). Understanding written problem instructions. In: L. W. Gregg (Ed.), *Knowledge and cognition.* (pp. 167–200). Erlbaum.
Ivanoff, J. M., De Wane, E. T., & Praem, O. (1965). Use of discriminant analysis for selecting students for ninth-grade algebra or general mathematics. *The Mathematics Teacher*, *58*(5), 412–416. https://doi.org/10.5951/MT.58.5.0412
Johansson, V. (2009). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working Papers*, *53*(2008), 61–79.
Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their applications to some educational problems. *Statistical Research Memoirs*, *1*, 57–93.

Kamata, A. (2001). Item analysis by the hierarchical generalized linear model. *Journal of Educational Measurement*, *38*(1), 79–93. https://doi.org/10.1111/j.1745-3984.2001.tb01117.x

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295x.95.2.163

Kulesz, P. A., Francis, D. J., Barnes, M. A., & Fletcher, J. M. (2016). The influence of properties of the test and their interactions with reader characteristics on reading comprehension: An explanatory item response study. *Journal of Educational Psychology*, *108*(8), 1078–1097. https://doi.org/10.1037/edu0000126

Lager, C. A. (2006). Types of mathematics-language reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology*, *27*(2-3), 165–204. https://doi.org/10.1080/02702710600642475

LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767. https://doi.org/10.1111/j.1467-8624.2010.01508.x

Leiss, D., Plath, J., & Schwippert, K. (2019). Language and mathematics – key factors influencing the comprehension process in reality-based tasks. *Mathematical Thinking and Learning*, *21*(2), 131–153. https://doi.org/10.1080/10986065.2019.1570835

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, *78*(2), 333–368. https://doi.org/10.17763/haer.78.2.70783570r1111t32

Muntoni, F., Wagner, J., & Retelsdorf, J. (2020). Beware of stereotypes: Are classmates' stereotypes associated with students' reading outcomes? *Child Development*, *92*(1), 189–204. https://doi.org/10.1111/cdev.13359

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide. Eighth Edition*. Muthén & Muthén.

Nalom, A. F. O., Soares, A. J. C., & Cárnio, M. S. (2015). The relevance of receptive vocabulary in reading comprehension. *CoDAS*, *27*(4), 331–338. https://doi.org/10.1590/2317-1782/20152015016

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4754). https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

O'Reilly, T., & McNamara, D. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "highstakes" measures of high school students' science achievement. *American Educational Research Journal*, *44*(1), 161–196. https://doi.org/10.3102/0002831206298171

OECD (2012). *Literacy, numeracy and problem solving in technology-rich environments: Framework for the OECD survey of adult skills*. https://www.oecd.org/skills/piaac/PIAAC%20Framework%202012–%20Revised%2028oct2013_ebook.pdf

OECD (2013a). *OECD skills outlook 2013: First results from the survey of adult skills*. https://www.oecd.org/skills/piaac/Skills%20volume%201%20(eng)–full%20v12–eBook%20(04%2011%202013).pdf

OECD (2013b). *Technical report of the survey of adult skills (PIAAC)*. https://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf

OECD (2014). *Education at a glance 2014: OECD indicators*. http://www.oecd.org/education/Education-at-a-Glance-2014.pdf

OECD (2016). *The survey of adult skills: Reader's companion, second edition, OECD skills studies*. https://www.oecd.org/publications/the-survey-of-adult-skills-9789264258075-en.htm

OECD (2019). *PISA 2018 results (Volume II): Where all students can succeed*. https://www.oecd-ilibrary.org/docserver/b5fd1b8f-en.pdf?expires=1602147835&id=id&accname=ocid49017929&checksum=62C15ECA001D7936D7E6993FB0BA9B37

Paetsch, J., Felbrich, A., & Stanat, P. (2015). Der Zusammenhang von sprachlichen und mathematischen Kompetenzen bei Kindern mit Deutsch als Zweitsprache [Second language proficiency and mathematical achievement: Their interrelations. *Zeitschrift Für Pädagogische Psychologie*, *29*(1), 19–29. https://doi.org/10.1024/1010-0652/a000142

Parsons, S., Bynner, J. (2005). *Does numeracy matter more?* National Research and Development Centre for Adult Literacy and Numeracy. https://discovery.ucl.ac.uk/id/eprint/1566245/1/parsons2006does.pdf
Peng, P., & Lin, X. (2019). The relation between mathematics vocabulary and mathematics performance among fourth graders. *Learning and Individual Differences*, *69*, 11–21. https://doi.org/10.1016/j.lindif.2018.11.006
Perfetti, C. (1985). *Reading ability.* Oxford University Press.
Pimm, D. (1987). *Speaking mathematically: Communication in mathematics classrooms*. Routledge & Kegan Paul.
Pitts, R. J. (1952). Relationship between functional competence in mathematics and reading grade levels, mental ability, and age. *Journal of Educational Psychology*, *43*(8), 486–492. https://doi.org/10.1037/h0055218
Powell, S. R., Driver, M. K., Roberts, G., & Fall, A.-M. (2017). An analysis of the mathematics vocabulary knowledge of third- and fifth-grade students: Connections to general vocabulary and mathematics computation. *Learning and Individual Differences*, *57*, 22–32. https://doi.org/10.1016/j.lindif.2017.05.011
Prediger, S. (2019). Investigating and promoting teachers' expertise for language-responsive mathematics teaching. *Mathematics Education Research Journal*, *31*(4), 367–392. https://doi.org/10.1007/s13394-019-00258-1
Prediger, S., & Wessel, L. (2013). Fostering German-language learners' constructions of meanings for fractions – design and effects of a language- and mathematics-integrated intervention. *Mathematics Education Research Journal*, *25*(3), 435–456. https://doi.org/10.1007/s13394-013-0079-2
Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E., & Benholz, C. (2018). Language proficiency and mathematics achievement: Empirical study of language-induced obstacles in a high stakes test, the central exam ZP10. *Journal Für Mathematik-Didaktik*, *39* (S1), 1–26. https://doi.org/10.1007/s13138-018-0126-3
Rammstedt, B., Martin, S., Zabal, A., Konradt, I., Maehler, D., Perry, A., Massing, N., Ackermann-Piek, D., & Helmschrott, S. (2016).Programme for the International Assessment of Adult Competencies (PIAAC), Germany – reduced version (ZA5845 Datefile Version 2.2.0) [Data set]. *GESIS – Leibniz-Institut Für Sozialwissenschaften*. https://doi.org/10.4232/1.12660
Reise, S. P. (2000). Using multilevel logistic regression to evaluate person-fit in IRT Models. *Multivariate Behavioral Research*, *35*(4), 543–568. https://doi.org/10.1207/S15327906MBR3504_06
Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Lawrence Erlbaum Associates.
Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading and Writing Quarterly*, *23*(2), 139–159. https://doi.org/10.1080/10573560601158461
Shaftel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment*, *11*(2), 105–126. https://doi.org/10.1207/s15326977ea1102_2
Thorndike, R. L. (1973). *Reading comprehension education in fifteen countries*. Wiley.
Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology*, *104*(3), 763–777. https://doi.org/10.1037/a0027470
Ufer, S., & Bochnik, K. (2020). The role of general and subject-specific language skills when learning mathematics in elementary school. *Journal Für Mathematik-Didaktik*, *41*(1), 81–117. https://doi.org/10.1007/s13138-020-00160-5
Ure, J. (1971). Lexical density and register differentiation. In G. Perren, & J. L. M. Trim (Eds.), *Applications of linguistics.* (pp. 443–452). Cambridge University Press.
Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* Academic Press.
Verschaffel, L., Greer, B., & de Corte, E. (2000). *Making sense of word problems.* Academic Press.

Wu, M. (2010). *Comparing the similarities and differences of PISA 2003 and TIMSS* (OECD Working Papers No. 32). Organisation for Economic Co-operation and Development. https://files.eric.ed.gov/fulltext/ED530738.pdf

Zabal, A., Martin, S., Massing, N., Ackermann, D., Helmschrott, S., Barkow, I., Rammstedt, B. (2014). *PIAAC Germany 2012: Technical report*. GESIS – Leibniz-Institut für Sozialwissenschaften. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-50410-5

Zhu, Z. (2007). Gender differences in mathematical problem solving patterns: A review of literature. *International Education Journal*, *8*(2), 187–203.