International Journal of Science and Mathematics Education (2025) 23:97–120
https://doi.org/10.1007/s10763-024-10463-5

Reality-Based Tasks with Complex-Situations: Identifying Sociodemographic and Cognitive Factors for Solution

Alina Knabbe[1] · **Dominik Leiss**[1] · **Timo Ehmke**[2]

Received: 6 September 2023 / Accepted: 27 March 2024 / Published online: 19 April 2024

Abstract

Acquiring mathematical literacy requires students to apply mathematics in various real-world contexts. However, mathematics classes often provide brief, content-focused descriptions of reality-based tasks and tasks that describe the situation as more complex, closer to reality, are still lacking. Students with different sociodemographic characteristics and cognitive factors have difficulties in solving reality-based tasks in mathematics lessons. The relationship between sociodemographic characteristics and cognitive factors (language and mathematical competence) concerning complex situation descriptions has not yet been investigated. To identify disadvantaged students in integrating such complex-situation tasks in mathematics lessons, this study aims to investigate which sociodemographic characteristics predict the solving of complex-situation tasks and whether cognitive factors mediate the relationship. Experts created 30 complex situations with different mathematical questions. A total of 519 9th- and 10th-grade students participated in a paper–pencil test. Path analysis revealed that the competence to solve complex-situation tasks is directly linked to gender and social background, with mathematical content-related skills and language competence mediating this relationship.

Keywords Complex-situation tasks · Mathematical content-related skills · Language competence · Students' characteristics

In everyday life, individuals encounter a wide spectrum of questions in various contexts. Mathematics is a valuable tool for finding answers, making informed decisions, and solving problems. The goal of mathematics education is to enhance mathematical literacy, enabling students to become responsible citizens who utilize their

✉ Alina Knabbe
alina.knabbe@leuphana.de

[1] Institute of Mathematics and Its Didactics, Leuphana University of Lueneburg, Universitaetsallee 1, 21335 Lueneburg, Germany

[2] Institute of Educational Science, Leuphana University of Lueneburg, Lueneburg, Germany

mathematical knowledge to make informed decisions and take independent actions (Organization for Economic Co-operation and Development [OECD], 1999, 2023a).

In math lessons, solving problems in a variety of contexts is enabled by working on reality-based tasks. Reality-based tasks represent a real-world situation and require a translation process from the real-world situation into math to solve the task (Blum & Leiss, 2007; Niss et al., 2007; Verschaffel et al., 2000). Various cognitive factors, such as linguistic and mathematical skills, are required to solve such tasks. Reality-based tasks can incorporate different aspects of reality (e.g. lack of information, multiple solutions) and enable different degrees of reality references that students may potentially encounter in their present and future (e.g. Greefrath & Vorhölter, 2016; Verschaffel et al., 2020). In math lessons, however, tasks without any real-world reference are often worked on (Heinle et al., 2022; Wijaya et al., 2015). Most reality-based tasks in large-scale assessments describe real-world situations in less than five lines (e.g. OECD, 2023a). To develop mathematical literacy and take real-world situations seriously, reality-based tasks that describe the real-world situation in its complexity (complex-situation tasks) must be worked on in math lessons, so that students understand and cope with written challenge situations in everyday life.

However, such reality-based tasks with complex situation descriptions would probably reinforce the existing educational inequality regarding reality-based tasks, as students have more difficulties with complex-situation reality-based tasks (Plath, 2020). Girls and socially disadvantaged students perform worse in solving (shorter) reality-based tasks (e.g. OECD, 2023b). So far, it is unknown how sociodemographic characteristics influence the solving of complex-situation tasks. However, this is important because mathematics education aims to train all students for potentially relevant situations, regardless of their sociodemographic characteristics. To prevent unconsciously reinforcing educational inequity by integrating such situation-complex reality-based tasks in mathematics lessons, this study aimed to investigate which students had difficulties in solving reality-based tasks with complex situation descriptions. Furthermore, we explored whether language competence and mathematical content-related skills mediated the relationship to formulate possible support options.

Theoretical Background

Mathematical Literacy as an Educational Goal

For over twenty years, mathematical literacy has been demanded internationally so that students learn to recognise mathematics in the world and make reasoned mathematical judgements by applying mathematics meaningfully (OECD, 1999, 2023a). The aim is to train students to become active and responsible citizens who use math to solve real-world problems in a self-determined way and make informed decisions. To ensure this, mathematical action in different situations and contexts is emphasised.

In math lessons, flexible mathematical behaviour is developed by solving reality-based tasks in a wide range of situations. To solve such tasks, students understand the real-world situation presented and translate it into a mathematical model, solve it using mathematics, and then check the mathematical result against the real-world situation and interpret it (OECD, 2014a). A translation occurs between the real-world situation and mathematics. The solution steps described here are sub-processes of modelling competence (Niss et al., 2007). According to Blomhøj and Højgaard (2007), mathematical modelling competence can be divided into three dimensions: the degree of coverage, the technical level, and the radius of action. The degree of coverage measures how many different sub-processes are run through when solving and at what level the students can reflect on these. The technical level dimension describes which math the students use and how flexibly they can apply it. The radius of action indicates whether and in how many different situations and contexts students can activate their modelling skills.

Continuum of Reality-Based Tasks

In the scientific community, different terms and definitions are used to describe types of reality-based tasks. Verschaffel et al. (2020) argue that these tasks can be placed along a continuum of world problems and modelling tasks. The continuum describes the degree of reality or authenticity to categorised tasks according to their reality-based characteristics.

World problems focus on practising mathematical content-related skills within a linguistically represented (real) problem situation (Niss et al., 2007; Verschaffel et al., 2000). The real-world situations play a lesser role. Word problem questions are usually closed, contain all relevant information for solving the task, and commonly have one correct solution to the mathematical question (Verschaffel et al., 2020). The degree of reality-based characteristics is, therefore, low.

Modelling tasks focus on understanding and assessing the real-world situation using mathematics. Mathematics is a tool used to solve a real-world situation. Modelling tasks are classified according to different characteristics (see for an overview Greefrath & Vorhölter, 2016). Modelling tasks have (open) questions about authentic contexts potentially relevant to the students' reality. They can differ in terms of the information given in the task (Maaß, 2010). Some modelling tasks contain different data, asking students themselves to decide which are relevant. Furthermore, in some modelling tasks, students have to make assumptions and find the necessary data (Kaiser, 2017). Another characteristic is the level of modelling activity, i.e. which processes are run through when modelling the task (Maaß, 2010).

In this study, reality-based tasks are understood as tasks on the continuum of word problems and modelling tasks that involve various aspects of solving real-world problems in reality and whose solution (to varying degrees) requires a translation from real-world situations into mathematics.

Reality-based tasks that contain descriptions of complex real-world situations are called complex-situation tasks. Complex-situation tasks contain features of word problems and modelling tasks and can be classified on the continuum. Like word problems, the tasks have a closed question for which no assumptions need to be made to answer with one correct answer. To solve complex-situation tasks, students must be able to recognise irrelevant and relevant information, as in modelling tasks,

Varroa Mites

Most bee colonies suffer from a type of mite known as Varroa, which is a parasite. It multiplies extremely fast in a hive and damages the larvae of the bees. Even if there are on average only 50 female mites in the hive at the beginning of spring, by late summer there are already more than 2,500 mites. The mites have a special reproduction cycle that lasts only 12 days. At the beginning of spring, the first cycle of reproduction begins when the 50 female mites attach themselves to nurse bees. These are special bees that are responsible for feeding the bee larvae in their cells. As soon as a nurse bee feeds a larva in a cell, the mite secretly goes with it into the bee larva's cell. The mite remains there even when the bee with a wax cap closes the cell. The female mite now begins to lay eggs in the cell, from which initially only male mites hatch after about 70 hours. More eggs hatch 30 hours later, all of which are female. Next, 150 hours after hatching, these female mites are sexually mature, whereupon the male mites mate with them in the cell. When the nurse bees open the cell lids two days later, the cleaning bees remove all the male and most female mites. However, a total of 3 mated female mites survive per infested cell, completing the first reproductive cycle. Now, these 3 female mites can be transported by nurse bees to the next cells, where the reproduction cycle starts all over again.

Which of the following terms describes the number of female mites after x multiplication cycles in the hive?

- ☐1 $50 \cdot x^3$
- ☐2 $3 \cdot x + 50$
- ☐3 $50 \cdot 3 \cdot x$
- ☐4 $50 \cdot 3^x$
- ☐5 $3 \cdot 50^x$

Fig. 1 Complex-Situation Task: Varroa Mites

and go through different sub-processes of modelling competence (depending on the difficulty level of the mathematical question). Figure 1 shows a complex-situation task from this study, which is about Varroa mites threatening a beehive.

In the following, study results from modelling tasks and word problems are summarised under the term reality-based tasks, even though working on word problems is easier than mathematical modelling (Pongsakdi et al., 2020). The derivation of the results of word problems and modelling tasks as reality-based tasks is possible because complex-situation tasks, investigated in this study have task features that characterise both word problems and modelling tasks.

Solving Reality-Based Tasks

The solution process of reality-based tasks is described by various ideal–typical modelling cycles, which divide the individual steps into sub-processes (see for an overview Borromeo Ferri, 2006). The modelling cycle is suitable as a theoretical model for solving reality-based tasks in a variety of situations as a translation process occurs between the real-world situation and mathematics. According to Blum and Leiss (2007), we used the modelling cycle, shown in Fig. 2, as it represents the situation model's formation and the solution's verification embedded in the real-world situation.

In the beginning, students engage in reading and (1) understanding the written task (Blum & Leiss, 2007), which necessitates language competence, such as skills in reading

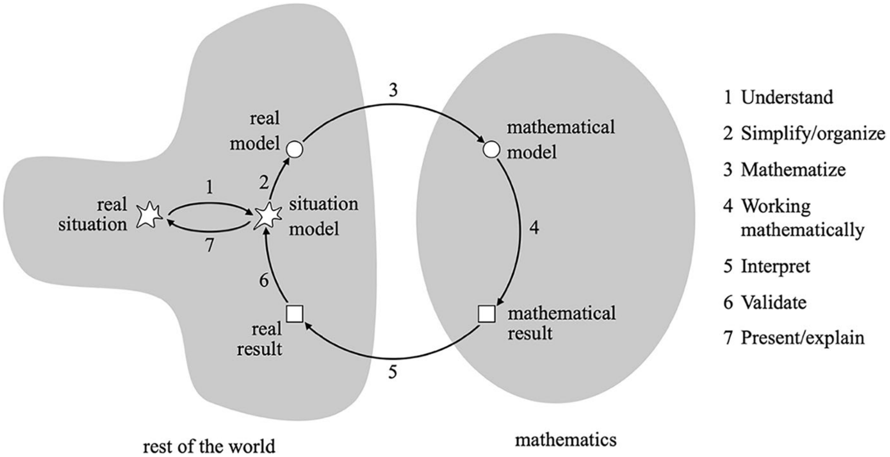


Fig. 2 Modelling Cycle after Blum and Leiss (2007)

and text comprehension. During this process, students (1) construct a situation model by comprehending and interpreting the task (Kintsch & Greeno, 1985; Leiss et al., 2010; Niss et al., 2007). The situation model is then transferred into a mathematical model through (2) structuring and (3) mathematizing (Blum & Leiss, 2007). These steps rely on the students' mathematical skills and language skills for cognitive processes. The result is (5) interpreted, (6) checked, and (7) written down or communicated to the real-world situation. Here, language competence is critical for effectively communicating or writing down the results.

The theory-driven modelling cycle is supported by empirical results, which show various cognitive factors necessary for solving reality-based tasks.

Necessary Cognitive Skills Various cognitive factors play a role in working on reality-based tasks. Cognitive skills such as spatial skills and verbal, numerical, and general reasoning are critical for solving reality-based tasks (e.g. Reinhold et al., 2020a, 2020b). Metacognitive knowledge, i.e. the knowledge of how to apply strategies, is also crucial (e.g. Vorhölter, 2023).

To "work mathematically", students need mathematical content-related skills appropriate to the mathematical content of the task (Blum & Leiss, 2007). These skills are summarised under mathematical content-related skills. The content-related skills refer to the competence of "Working with mathematical objects" in the German educational standards (Kultusministerkonferenz, 2022), including the factual knowledge of using mathematical procedures to process mathematical tasks efficiently. The connection between mathematical content-related skills and solving reality-based tasks has been demonstrated in many studies (Fuchs et al., 2006; Leiss et al., 2010; Reinhold et al., 2020a, 2020b; Strohmaier et al., 2021). Pongsakdi et al. (2020) examined the linguistic and numerical characteristics of arithmetical reality-based tasks and differences in text comprehension and arithmetic skills in relation to solving reality-based tasks. The results showed a significant correlation between content-related (arithmetical) skills and solving reality-based tasks across all difficulty levels $r = 0.59$, $p < 0.01$. Leiss et al. (2010) found a positive correlation of $r = 0.302$, $p < 0.01$ between content-related skills and the solution rate of reality-based tasks.

Kintsch and Greeno (1985) demonstrated the importance of general language competence in solving reality-based tasks by integrating the text comprehension model of van Dijk and Kintsch (1983) into their own. Numerous studies have highlighted the significance of language skills (see for an overview Daroczy et al., 2015), especially language competence such as text (e.g. Björn et al., 2016; Pongsakdi et al., 2020) and reading comprehension (e.g. Fitzpatrick et al., 2020; Leiss et al., 2019; Vilenius-Tuohimaa et al., 2008), in understanding the real-world situation in reality-based tasks. In a meta-analysis, Peng et al. (2020) identified a significant correlation of $r = 0.45$ between solving reality-based tasks and general language competence. In our study, language competence is understood as a general measure that combines reading and text comprehension, as students are required to read and understand a variety of reality-based tasks in complex situation descriptions.

Although mathematical content-related skills are not the only relevant mathematical skills in solving reality-based tasks, this study focused on mathematical content-related skills and language competence as cognitive factors because they influence not only the overall solution process of reality-based tasks but also the comprehension process of the real-world situation (Krawitz et al., 2022; Leiss et al., 2019). The comprehension process of the real-world situation (forming the situation model and real model) is an essential element in the solution process of reality-based tasks and is positively related to the solution rate of the task (Krawitz et al., 2022; Leiss et al., 2019). Especially at the beginning of a reality-based task, many students have difficulties in the translation process between the real-world situation and mathematics (Alsina & Salgado, 2022; Blum, 2015; Jankvist & Niss, 2020; Wijaya et al., 2014). Plath (2020) examined how students mathematically solved complex descriptions of real-world situations in the form of a newspaper article in comparison to short reality-based tasks focusing on mathematical content. Considering the duration of the comprehension processes, the comprehension process for complex situation descriptions (in the form of newspaper articles) accounted for 62% (including the first reading) and 48% (after the first reading) (Plath, 2020). In comparison, students needed 38% of the total processing time for the word problem comprehension process and 27% after the first reading.

Plath demonstrated that students solved reality-based tasks with complex descriptions of real-world situations in the form of a newspaper article more poorly than short reality-based tasks focused on mathematical content and required more time for the comprehension process. The results are not surprising considering the results on the use of reality-based tasks in math lessons in general.

Current Status: Use of Reality-Based Tasks in Math Lessons

In math lessons, students tend to work mainly on tasks not related to real-world situations (Heinle et al., 2022; Wijaya et al., 2015). Furthermore, these tasks typically contain short situation descriptions that focus on the current subject matter in large-scale assessments (OECD, 2023a). Even in a typical modelling task such as the "Filling up" task (Blum, 2011), which is used to represent various sub-processes of modelling, the real-world situation is described with a strong focus on the mathematical model. If we take the training of mathematical literacy seriously and want students to be able to apply mathematics in different situations, then in addition to the existing reality-based tasks, other types of reality-based tasks, which involve a complex description of real-world situations, need to be worked on to get closer to solving potentially relevant real-world problems in reality.

When working on complex-situation tasks in mathematics lessons; however, the existing educational inequality with regard to mathematical literacy (OECD, 2023b) and reality-based tasks is likely to be exacerbated because students have even more difficulties with complex-situation tasks than with short reality-based ones (Plath, 2020). As there has been no research to date on the relationship between sociodemographic characteristics and complex-situation tasks, the following section presents study results investigating the general relationship between sociodemographic characteristics and solving reality-based tasks.

Sociodemographic Characteristics Influencing the Solving of Reality-Based Tasks

Theoretical models by Eccles and Wigfield (2020) were used to determine how sociodemographic characteristics are related to academic and mathematical performance. Empirical results on the direct relationship between sociodemographic characteristics and solving reality-based tasks are then reported. Finally, studies are presented that not only investigate the direct influence of sociodemographic characteristics and reality-based task solving but also take into account language competence and mathematical content-related skills in connection with solving reality-based tasks.

Direct Effect: Sociodemographic Characteristics and Solving Reality-Based Tasks

The PISA results reveal that students' sociodemographic characteristics, including gender, migration background, parental education, and cultural resources (measured by books in the home) are related to their solving reality-based tasks (OECD, 2013, 2014a). This relationship can be explained by the (situated) expectancy-value theory and model of parents' socialisation of motivation according to Eccles and Wigfield (2020). Students are raised in a social and cultural environment where they encounter the beliefs and expectations of those around them. These beliefs and expectations, especially from their parents, are influenced by their level of education and cultural background, such as migration background. Beliefs and expectations of parents and the social environment include, for example, gender-role stereotypes and efficacy beliefs. Depending on these expectations and beliefs, parents provide learning opportunities, such as books in the home as a cultural resource, which can cover a wide range of topics. These attitudes, beliefs, and expectations affect the student's (mathematical) performance.

The achievement differences between boys and girls in mathematics remain controversial (e.g. Lindberg et al., 2010; Reilly et al., 2015). However, in solving reality-based tasks boys have continually outperformed girls (Lindberg et al., 2010; OECD, 2019; Plath & Leiss, 2018; Reinhold et al., 2020a, 2020b). Even when considering cultural resources (measured by books in the home: $\beta = 0.104$, $p < 0.01$), Plath and Leiss' (2018) study shows that gender, favouring boys, remains a crucial predictor of solving reality-based tasks ($\beta = -0.177$, $p < 0.01$). The most pronounced differences in the performance of reality-based tasks in large-scale assessments are observed between socially advantaged and disadvantaged students (OECD, 2013). Students with a migration background tend to fall into the socially disadvantaged category (OECD, 2019). Students with a migration background perform worse than students without it (OECD, 2013). Although controlling for socioeconomic status reduces the performance gap, students with a migration background still perform worse than students without it. A significant difference in solving reality-based tasks has also been found between students whose parents are highly educated and those with less educated parents (OECD, 2013; Piel & Schuchart, 2014; Vilenius-Tuohimaa et al., 2008). Students with greater cultural resources demonstrate improved solving reality-based tasks (OECD, 2013; Plath & Leiss, 2018).

Interplay of Sociodemographic Characteristics and Cognitive Factors in Solving Reality-Based Tasks

It is assumed that sociodemographic characteristics influence solving reality-based tasks as they affect the development of the necessary skills required to solve the tasks, such as language competence and content-related skills (Peng et al., 2020; Plath & Leiss, 2018). This aligns with the model of Eccles and Wigfield (2020), as the social environment plays a crucial role in overall academic performance. For example, girls and socially advantaged students perform better in language skills (OECD, 2019; Vilenius-Tuohimaa et al., 2008), whereas boys and socially advantaged students perform better in different mathematical skills (Björn et al., 2016; Piel & Schuchart, 2014; Reinhold et al., 2020a, 2020b).

Vilenius-Tuohimaa et al. (2008) conducted path analysis to investigate the relationship between mother's education, gender, language skills, and performance in reality-based tasks. Controlling for gender and language competence, the results showed that higher parental education level was associated with better performance in reality-based tasks ($\beta = 0.30$, $p < 0.001$) and language competence ($\beta = 0.32$, $p < 0.001$). Gender exhibited no effect in this model. The overall between solving reality-based tasks and language competence was found to be $r = 0.38$, $p < 0.001$.

Plath and Leiss (2018) examined the relationship between personal characteristics (migration background, gender, and cultural resources) and solving reality-based tasks. They identified a correlative relationship of $r = —0.130$, $p < 0.001$ between migration background and solving reality-based tasks. They also calculated two regression analyses. According to the results of the first regression analysis, after controlling for German and math grades, gender and cultural resources exhibited no significant effect between migration background and solving reality-based tasks. The results showed that boys outperformed girls significantly in solving the reality-based tasks ($\beta = -0.177$, $p < 0.01$).

Similarly, greater cultural resources had a significant positive effect on solving reality-based tasks ($\beta = 0.104$, $p < 0.01$). The first model accounted for 8% of the variance in the solution rate. In the second model, when language competence was considered, gender influenced solving reality-based tasks favouring boys ($\beta = -0.209$, $p < 0.001$). Language competence had a significant influence on solving reality-based tasks ($\beta = 0.282$, $p < 0.001$). The effect between cultural resources and solving reality-based tasks was no longer significant when language competence was included, suggesting a possible mediation. Overall, the model explains 14% of the solution rate.

Björn et al. (2016) conducted a longitudinal study to explore the mediation of parental education via language competence and mathematical content-related skills on the development of solving reality-based tasks from primary to secondary school. The first model, focusing on language competence, showed that parental education had a partial mediation effect through language competence ($\beta = 0.31$, $p < 0.001$). In addition, a direct effect was found between parental education and solving reality-based tasks ($\beta = 0.12$, $p < 0.05$). In the second model, content-related skills were included. The results revealed that the influence of parental education on solving reality-based tasks was completely mediated by language competence and content-related skills.

Reinhold et al. (2020a, 2020b) examined the relationship between cognitive factors, such as content-related skills, language competence, spatial skills, and solving reality-based tasks and the role of gender. In the first model, where gender was the independent variable and modelling performance was the dependent variable, it was found that male students performed better at modelling tasks than female students ($\beta = 0.554$, $p < 0.001$). In the second model, the cognitive factors were included. The results showed that all cognitive factors explained solving reality-based tasks. Language competence emerged as the strongest predictor, with $\beta = 0.342$, $p < 0.001$, whereas content-related skills were the smallest predictor ($\beta = 0.164$, $p < 0.001$). Controlling for language competence, content-related skills, spatial skills, and general reasoning skills, boys performed on average 13.6% more correctly than girls. The effect of gender, however, decreased to $\beta = 0.285$, $p < 0.001$, suggesting a mediation of cognitive factors.

The results on the relationship between sociodemographic characteristics and solving reality-based tasks with less detailed real-world situation descriptions demonstrate that girls, students with a migration background, and students with less parental education and fewer cultural resources are disadvantaged. Even if initial results show that working on complex-situation leads to greater difficulties for students (Plath, 2020) and, thus, tends to increase educational inequality, working on complex-situation tasks is also necessary for math lessons to train students to become active, self-determined citizens in the sense of mathematical literacy who can make decisions in a variety of contexts using mathematics.

Present Study

This study aimed to investigate which students had particular difficulties in solving complex-situation tasks due to their sociodemographic characteristics to identify individuals who potentially need support in integrating such tasks into mathematics lessons.

Plath (2020) showed that the comprehension process took longer and complex-situation tasks were difficult to solve. It is precisely during the comprehension process that errors often occur (Alsina & Salgado, 2022; Blum, 2015; Jankvist & Niss, 2020; Wijaya et al., 2014). As language skills and mathematical content-related skills are critical for the comprehension process (Krawitz et al., 2022; Leiss et al., 2019), we also investigated whether language competence and mathematical content-related skills mediated the relationship between sociodemographic characteristics and the competence to solve complex-situation tasks to formulate support options for the students concerned. This study aimed to answer the following research questions:

1. How much variance in the competence to solve complex-situation tasks can be explained by sociodemographic characteristics of students (gender, migration background, parental education, and cultural resources)?
2. To what extent do (a) mathematical content-related skills and (b) language competence play a mediating role between sociodemographic characteristics and the competence to solve complex-situation tasks?

Based on the results on the relationship between sociodemographic characteristics and solving reality-based tasks, we assumed that gender in favour of boys, migration background, parental education and cultural resources would directly affect the competence to solve complex-situation tasks. We also assumed that language competence and mathematical content-related skills would mediate the connection, as they are crucial for the comprehension process.

Figure 3 presents the theoretical model for solving these tasks, incorporating sociodemographic characteristics and two cognitive factors. The sociodemographic characteristics serve as distal predictors, whereas the cognitive factors act as presumed mediators in the middle. The theoretical model presents direct and indirect pathways.

Previous studies have either examined some of the sociodemographic characteristics using reality-based tasks with other task characteristics or looked at situational complexity as a task characteristic (Plath, 2020), but have not considered the sociodemographic characteristics of the students. Thus, our study is the first to investigate the relationship between sociodemographic characteristics, cognitive factors, and solving complex-situation tasks.

Methods

Sample

To address the research questions, 519 students (female: 49.9%, male: 50.1%) at two schools (German: Gesamtschulen) in grades ninth and tenth were surveyed in the fall of 2020 in Germany. The average age of the students in the sample was 15.2 years ($SD = 0.913$). Among the participants, 84.6% had no migration background,

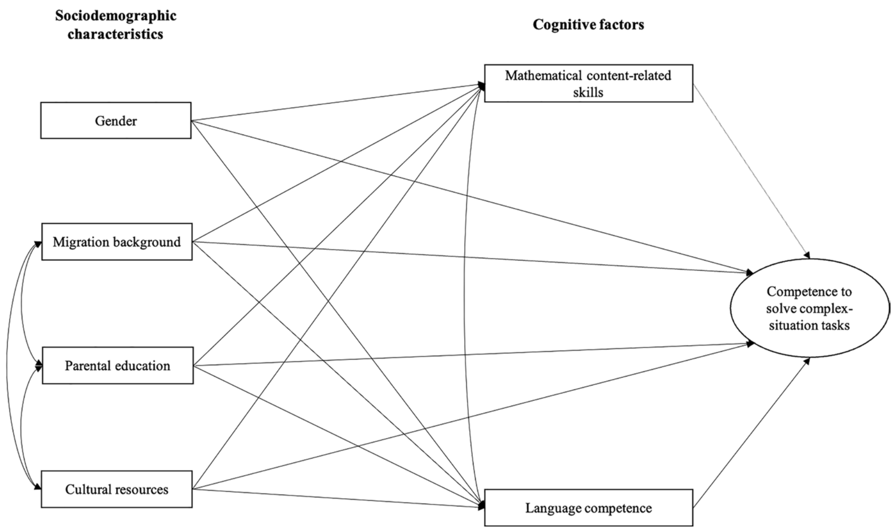


Fig. 3 Theoretical Model of Predictors for the Competence to Solve Complex-Situation Tasks

whereas 15.4% had a migration background. The average highest level of parental education was 12.7 years (*SD* = 2.790), which corresponded to either holding a higher school certificate or having completed vocational training in Germany. Most students (23.2%) reported that they had 26–100 books at home.

Instruments

The students participated in an 85-min paper–pencil test, administered by trained test administrators who had received prior training. The questionnaire comprised four sections: solving complex-situation tasks, a cloze test to assess language competence, a mathematical content-related test, and a sociodemographic questionnaire. Table 1 presents the means, standard deviations, and reliability of the variables. A rotating design was used to test as many complex-situation tasks as possible with various real-world situations. This design involved arranging the complex-situation tasks in 18 test booklets (see Fig. 4).

Complex-Situation Tasks

The complex-situation tasks consisted of two sections: the complex description of the real-world situation (complex situation) and the mathematical question with possible solutions (item). An interdisciplinary team of mathematics didacticans, educational scientists, and linguists developed 30 complex situations. These situations were based on real-world problems from various subject areas such as digital media, environmental protection, alternative energy sources, health, transportation, technology, nature, and finance. The complex situation, illustrated in Fig. 1 about the threat of varroa mites to the hive is from the subject area nature. The idea was to develop

Table 1 Psychometric Properties for Scales

Scale	*M*	*SD*	Range	Number of items	Reliability
Competence to solve complex-situation tasks [a]	0.00	0.99	-4.06–3.87	15	0.73
Language competence [b]	19.25	4.81	0–25	25	0.88
Content-related skills [b]	3.07	2.11	-3.53–5.30	11	0.74

[a] WLE Person Separation Reliability. [b] Cronbach's α

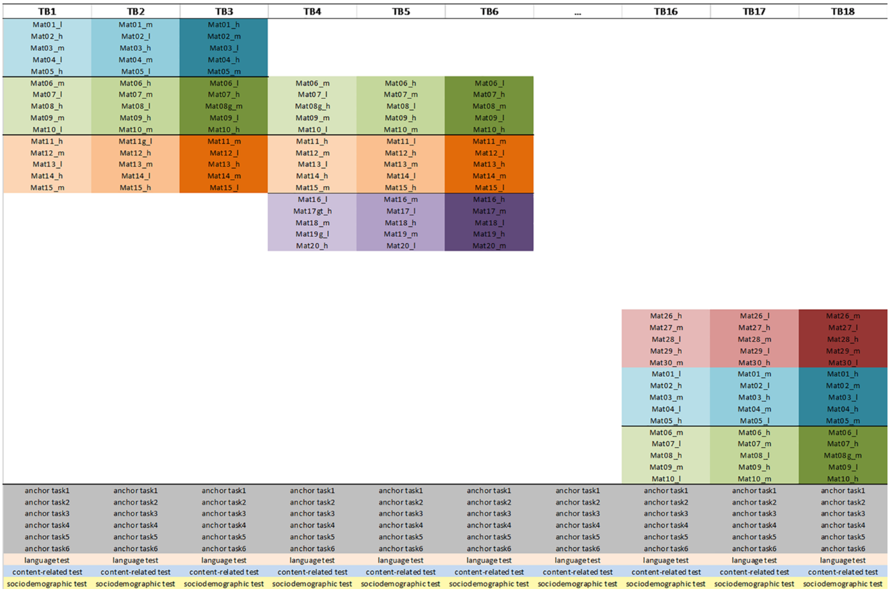


Fig. 4 Structure of the Study

diverse complex situations potentially relevant to students in their future and present to ensure a close resemblance to the complexity of reality. Unlike reality-based tasks in textbooks, these complex situations included longer descriptions (number of words: $M = 207$, $SD = 29$) with real numerical and factual information. On average, the complex situation contained 8,02 ($SD = 2.01$) numerical data and an average of 3.18 numerical data (SD = 1.31) from the complex situation is required to solve the tasks. The linguistic level of all complex situations was intermediate and measured with the readability index LIX (Anderson, 1981).

After the development of 30 complex situations, at least three mathematical questions (items) were formulated for each complex situation. All mathematical question tasks focused on the mathematical content area of functional relationships, ensuring that all students, regardless of their level, had covered the content in class. These questions had three different difficulty levels, light, medium, and heavy, and targeted

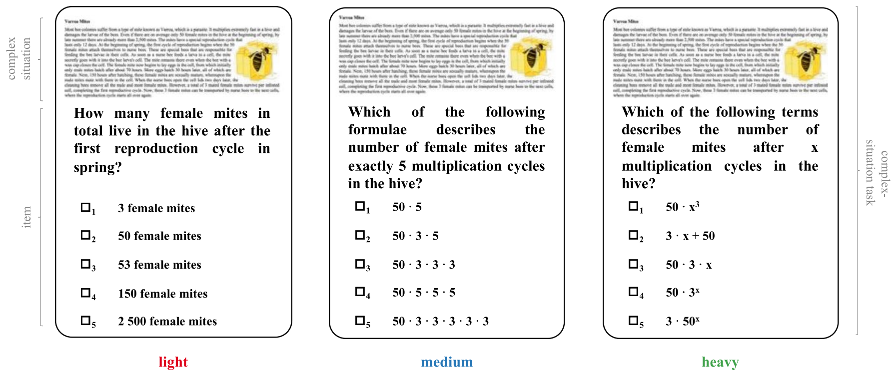


Fig. 5 Example of Three Difficulty Levels in the Same Complex Situation

different sub-processes of the solution process. Figure 5 shows the complex situation of Varroa Mites with three different mathematical questions of varying difficulty levels. Students responded by selecting one of the five response options.

In total, each student completed 15 complex-situation tasks, consisting of a complex situation (pool of 30) and an item with a difficulty level of light, medium, or heavy. Therefore, each student had five complex-situation tasks, at each difficulty level.

Competence to Solve Complex-Situation Tasks The competence to mathematically solve such complex described real-world situations is called “competence to solve complex-situation tasks" in this study. The term is based on the dimensions of modelling competence by Blomhøj and Højgaard (2007) and focuses mainly on the third competence, the radius of action. More precisely, it measures how students solve different complex situations mathematically. Although the three different levels of difficulty activate different sub-processes during the solving process (degree of coverage), these are not analysed individually. The dimension of the technical level is limited to the mathematical content area of functional relationships.

To determine student ability estimates for competence to solve complex-situation tasks, one-dimensional IRT models were applied using ConQuest (Wu et al., 2007) due to the rotating design. The solutions and difficulty level of the tasks were considered in the Rasch scaling of competence to solve complex-situation tasks. Six additional reality-based tasks in each test booklet served as anchor items for the Rasch scaling process in which all students solved the tasks (Fischer et al., 2021).

Cognitive Factors

The test instrument used to record mathematical content-related skills was focused on the same content area as the reality-based tasks, specifically the functional relationship. The test instrument consisted of a total of 11 items, with eight items taken from a standardised test that required solving equations, e.g. $5 + x = 3$, $x = ?$ (Schmidt et al., 2012). To capture all the necessary content-related skills, three further tasks were self-developed, which required knowledge of technical terms for their solution, e.g. “Given are the two following functions: $y_1 = 2 \cdot x + 1$ and $y_2 = 2{,}5 \cdot x$. What is the intersection of the two functions?”. To select the test items, two experts in mathematics didactics first independently analysed the necessary content-related skills for solving all complex-situation tasks on three difficulty levels and selected them on this basis. As not all content-related skills were covered by the standardised test, each of the experts developed additional tasks to cover all content-related skills. In the final step of item selection, the experts compared their results and agreed on the 11 items. The reliability of the mathematical content-related skills test ($M = 3.07$, $SD = 2.11$) can be rated as good, with Cronbach's $\alpha = 0.74$ (Table 1). The students' content-related skills were assessed by calculating the sum score of correctly solved tasks.

Students' language competence was assessed using a brief cloze test (DCLL + 3).[1] The cloze format has been proven to be a fairly robust, reliable, and valid measure of general language competence (Tremblay, 2011). The text consisted of an excerpt from a German newspaper focused on the topic of people going to work despite being sick. This is a classic information text describing a real-world problem, like complex-situation tasks. The text comprised 96 words. During the test, students were given five minutes to complete the second part of 25 words. The total score of the student's language competence is the number of correctly spelt words.

Sociodemographic Characteristics

Sociodemographic characteristics were determined using a sociodemographic questionnaire, which included items on gender, age, migration background, parental education, and cultural resources. Gender was recorded as binary: male = 1 and female = 2. Parental education was measured as the highest educational attainment of the parents in years. The PARED index was used to calculate parental education (for more details see: OECD, 2014b). Cultural resources were assessed by asking students to indicate the number of books they had at home. This was measured on a six-point scale ranging from “0–10 books” to “more than 500 books” (for more details see: OECD, 2014b). Each item was dummy coded: “0–10 books” = 1, and “more than 500 books” = 6. The book variable has been used for years in large-scale assessments as an indicator of socio-economic status (OECD, 2014b; Yin & Fishbein, 2020). Even though e-books have been increasingly bought, the book question is a valid instrument for measuring cultural resources (Schwippert, 2019). Migration backgrounds were determined by students indicating the country in which they and their parents were born. The response options were “Germany” or “other country”. Migration background was divided into two categories: students without a migration background (i.e. both parents were born in Germany, coded with 1) and students with a migration background (i.e. both parents were not born in Germany, coded

[1] The DCLL + 3 is a standardised test and is available from the Test Development and Diagnostics Unit of the Hamburg Institute for Educational Monitoring and Quality Assurance.

Table 2 Correlation of Study Variables

Variable	1	2	3	4	5	6	7
1. Competence to solve complex-situation tasks	–						
2. Language competence	0.43**	–					
3. Content-related skills	0.47**	0.26**	–				
4. Gender	-0.11*	0.16**	-0.02	–			
5. Migration background [b]	-0.18**	-0.35**	-0.13**	0.03	–		
6. Parental education	0.19**	0.26**	0.12**	-0.06	-0.24**	–	
7. Cultural resources	0.25**	0.26**	0.20**	-0.03	0.23**	0.28**	–

[a] *1 = male and 2 = female.* [b] *1 = without migration background and 2 = with migration background*
$^{*}p<0.05$. $^{**}p<0.01$

with 2). Table 2 shows the correlation among the seven measures in this study. None of the correlation coefficients exceeded $r=0.47$, indicating that there were no spurious inter-correlations between the measures expected for the path model.

Data Analysis

To ensure that the planned path analysis had sufficient statistical power, we conducted two apriori G-power analyses for research questions 1 and 2. The result of the G-power analysis revealed that a minimum sample size of 191 (RQ1) and 215 participants (RQ2) was required. The sample of this study is considered sufficient for accurately calculating the effects.

We conducted path analyses using Mplus (Muthén & Muthén, 1998-2017) to examine the association between gender, sociodemographic characteristics (parental education, cultural resources, and migration background), language competence, and mathematical content-related skills while solving complex-situation tasks. Due to its dichotomous coding, gender was standardised with the STDY option. The other coefficients were standardised with the STDYX option. For each research question, we calculated two models: (RQ1: Model 1, RQ2a-b: Model 2). Their model fits were Model 1: $\chi^2(3)=1.750$, $p=0.623$, CFI = 1.000, TLI = 1.000, RMSEA = 0.000, SRMR = 0.017 and Model 2: $\chi^2(3)=1.841$, $p=0.606$, CFI = 1.000, TLI = 1.000, RMSEA = 0.000, SRMR = 0.013.

To address the first research question, we calculated Model 1, which included gender, migration background, parental education, and cultural resources as independent variables, and competence to solve complex-situation tasks as dependent one. In addition, inter-correlations between the sociodemographic characteristics were estimated. For the second research question, we utilised Model 2. This model incorporated cognitive factors as possible mediators. Two paths were estimated, each representing the relationship between sociodemographic characteristics as independent variables and content-related skills (RQ2a) and language competence (RQ2b) as dependent variables. In the third path of the second model, the sociodemographic characteristics (gender, migration background, parental education, and cultural resources) and the two cognitive factors (content-related skills and language competence) were treated as independent variables and competence to solve complex-situation tasks as dependent ones. We also considered the inter-correlations between content-related skills and language competence and between sociodemographic characteristics.

Results

The presentation of results focuses on the relationship between sociodemographic characteristics and competence to solve complex-situation tasks and whether the relationship is mediated by cognitive factors (content-related skills and language competence). Figure 6 shows the results of the path analysis.

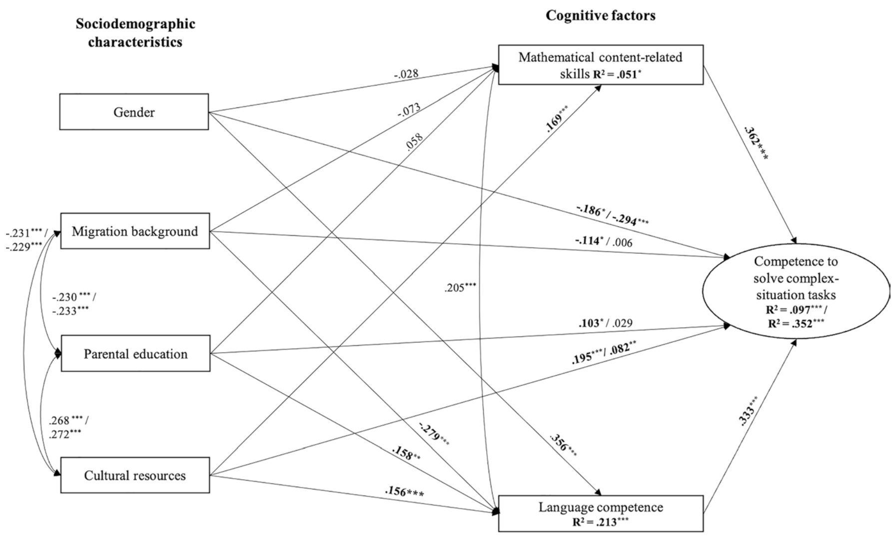
Prediction of Sociodemographic Characteristics for Competence to Solve Complex-Situation Tasks (RQ1)

The first research question investigated the relationship between sociodemographic characteristics (gender, migration background, parental education, and cultural resources) and competence to solve complex-situation tasks. The sociodemographic characteristics accounted for 9.7% ($p < 0.001$) of the competence to solve complex-situation tasks. The analysis showed that students with more cultural resources ($\beta = 0.195$, $p < 0.001$) and those without a migration background ($\beta = -0.114$, $p < 0.05$) had significantly higher competence to solve complex-situation tasks. Higher parental education ($\beta = 0.103$, $p < 0.05$) had a significant prediction effect on solving complex-situation tasks. Gender had a prediction effect in favour of boys ($\beta = -0.186$, $p < 0.\ 05$) when compared to girls.

Role of Cognitive Factors in the Relationship between Sociodemographic Characteristics and Competence to Solve Complex-Situation Tasks

Role of Content-related Skills Between Sociodemographic Characteristics and Competence to Solve Complex-Situation Tasks (RQ2a)

Concerning the second question (a) regarding whether content-related skills mediated the relationship between sociodemographic characteristics and competence to solve complex-situation tasks, the path analysis showed that content-related skills served as a mediator of the relationship between cultural resources and competence to solve complex-situation tasks. Among the sociodemographic characteristics, only cultural resources had a significant prediction of content-related skills ($\beta = 0.169$, $p < 0.001$).



Note. The standardized beta coefficients and significance levels are shown alongside the arrows. Each arrow representing the direct effects of sociodemographic characteristics on competence to solve complex-situation tasks is associated with two beta coefficients. The first coefficient corresponds to the beta coefficient in Model 1, which only includes sociodemographic characteristics and competence to solve complex-situation tasks. The second beta coefficient represents the direct effect despite the mediation effect in the analysis of Model 2 (for research questions 2a and 2b). $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Fig. 6 Path Analysis Prediction of Sociodemographic Characteristics for Competence to Solve Complex-Situation Tasks with Cognitive Factors as Mediators

Role of Language Competence Between Sociodemographic Characteristics and Competence to Solve Complex-Situation Tasks (RQ2b)

In terms of language competence, all sociodemographic characteristics were mediated by language competence concerning competence to solve complex-situation tasks. Gender in favour of girls had the strongest effect on language competence ($\beta = 0.356$, $p < 0.001$). Students without a migration background had significantly better language competence than those with one ($\beta = -0.279$, $p < 0.001$). Parental education ($\beta = 0.158$, $p < 0.01$) and cultural resources ($\beta = 0.156$, $p < 0.001$) had a significant effect on student language competence. Students whose parents had a higher level of education and more cultural resources at home developed better language competence.

Direct and Indirect Effects of Sociodemographic Characteristics on Competence to Solve Complex-Situation Tasks

Despite the significant mediation of gender on language competence, student gender directly affected competence to solve complex-situation tasks ($\beta = -0.294$, $p < 0.001$). This means that when language competence, content-related skills, and social background were equal, boys outperformed girls in solving complex-situation tasks. Compared to the first model, this effect increased with the inclusion of cognitive factors. Cultural resources also had a significant but negligible effect ($\beta = 0.082$, $p < 0.01$) on competence to solve complex-situation tasks, which was mediated through language competence ($\beta = 0.156$, $p < 0.001$). More specifically, cultural resources indirectly affected competence to solve complex-situation tasks through their impact on language competence. Therefore, for students with the same language competence, migration background and parental education did not have a direct influence on competence to solve complex-situation tasks.

In terms of cognitive factors, content-related skills ($\beta = 0.362$, $p < 0.001$) and language competence ($\beta = 0.333$, $p < 0.001$) emerged as the strongest predictors of competence to solve complex-situation tasks. Overall, Model 2 explained 35.2% ($p < 0.001$) of the variance in solving complex-situation tasks. Comparing the variance resolutions of Models 1 and 2, it is evident that the inclusion of cognitive factors in Model 2 explained 25.5% more variance in the competence to solve complex-situation tasks.

Discussion

In terms of mathematical literacy, students should be empowered to make decisions with mathematics. Accordingly, in addition to the existing reality-based tasks, mathematics education must also include tasks that take into account the complexity of the situation in reality. So far, we know that such complex-situation tasks are particularly difficult for students (Plath, 2020) and that students with certain sociodemographic characteristics are at a disadvantage when solving reality-based tasks in general (e.g. OECD, 2023b). To identify students who potentially need support in maths lessons when working on a variety of complex-situation tasks, this study aims to determine which students need support and whether certain cognitive factors mediate the relationship to show possibilities for support.

Empirical Insights and Theoretical Contributions

The results of previous studies (e.g. OECD, 2013, 2014a; Vilenius-Tuohimaa et al., 2008) support the findings in this study concerning the relationship between sociodemographic characteristics and solving complex-situation tasks (RQ1). This study demonstrated that factors such as gender (favouring boys), migration background, parental education and cultural resources (at the expense of socially advantaged students) explained a total of 9.7% of the variance in competence to solve complex-situation tasks. When considering cognitive factors as mediators (RQ2), boys and students with more cultural resources maintained a direct association with better competence to solve complex-situation tasks. The impact of students' migration background and parental education is fully mediated by language competence. Furthermore, gender and cultural resources are mediated by language competence, and cultural resources are mediated by content-related skills. Among cognitive factors, content-related skills ($\beta = 0.362$, $p < 0.001$) are the strongest predictor followed by language competence ($\beta = 0.327$, $p < 0.001$). In comparison, Reinhold et al., (2020a, 2020b) considered other cognitive factors and found that content-related skills ($\beta = 0.164$, p < 0.001) were the lowest predictor of solving reality-based tasks, whereas language competence ($\beta = 0.342$, p < 0.001) was the highest. Overall, the path analysis in this study accounts for 35.2% of the variance in solving complex-situation tasks which is considered a high-value (Cohen, 1988). The results of the study are in line with the theoretical models of Eccles and Wigfield (2020), as the investigated sociodemographic characteristics of the students (social and cultural environment) are directly and indirectly predictive of the competence to solve complex-situation tasks (performance). Concerning the modelling cycle, it can be stated that language competence as well as content-related skills affect solving. This study did not investigate the specific sub-processes of the modelling cycle in which cognitive factors play a role. Further research is necessary to determine this.

Limitations and Suggestions for Further Research

In this study, complex-situation tasks were created to approximate complexity in potentially relevant real-world problems of students in their future and present. Importantly, we mainly focused on investigating the dimension radius of action to complex-situation tasks (Blomhøj & Højgaard, 2007). The tasks were designed with mathematical questions and one-choice solutions to allow for the exploration of multiple real-world situations. Further research should consider examining all dimensions and other task characteristics, such as open questions and multiple solutions in complex-situation tasks.

In this study, we focused on investigating two cognitive factors: general language competence and content-related skills as mediators. Other studies have highlighted the importance of cognitive factors such as subject-related language abilities (e.g. Ufer & Bochnik, 2020), metacognitive knowledge (e.g. Vorhölter, 2023), spatial abilities, and general reasoning (e.g. Reinhold et al., 2020a, 2020b). Therefore, future studies could consider including these additional cognitive factors as mediators and re-examine the relationship between sociodemographic characteristics and solving complex-situation tasks. It would also be interesting to see whether auxiliary material, as found by Guerrero-Ortiz et al. (2018) in primary school students, would also help solve complex-situation tasks. This would allow for the identification of additional support possibilities for students. Future research could investigate whether the different real-world situations influence the solving of complex-situation tasks, as Albarracín et al. (2021) found that different strategies were used to solve tasks in different real-world situations.

Conclusions and Implications for Practice

The results indicate that language competence and content-related skills are crucial for the development of competence to solve complex-situation tasks. Supporting the development of these cognitive factors is critical for socially disadvantaged students, as language competence and content-related skills are influenced by sociodemographic characteristics. However, the small direct effect of cultural resources and the indirect effect of migration background and parental education on cognitive factors showed that socially disadvantaged students need further support for solving complex-situation tasks in addition to cognitive factor support. Socially disadvantaged students are a particular focus, as both language and content-related skills are influenced by sociodemographic characteristics. However, female students perform worse in complex-situation tasks despite better language skills and no significant differences in content-related skills. Therefore, it is crucial to provide additional support and encouragement for female students in the field of complex-situation tasks.

The complex-situation tasks developed in this study are closed questions at three different levels of difficulty. Each student was only given one mathematical question (at one difficulty level) on one complex-situation. To prepare complex tasks for a heterogeneous group of students in class, the teacher could offer tasks of different levels, both open and closed. In the context of low ceiling high floor tasks (e.g. Boaler, 2022; Sullivan et al., 2009), students can work on the three tasks of a complex situation one after the other. They can then be encouraged to develop further, more challenging questions through problem-posing. This would involve engaging deeply with the complex situation and the mathematical content at various levels of difficulty. Importantly, even if certain groups may have advantages in the competence to solve complex-situation tasks, it is vital to incorporate more complex-situation tasks into mathematics lessons while considering the special individual students' needs. The goal of mathematics education is to provide all students with an opportunity to develop mathematical literacy to make informed decisions based on mathematical considerations.

Acknowledgements The VAMPS project (Variation Tasks Mathematics, Physics, and Language) is funded by the German Research Foundation (DFG-project no.: 417017613, project leader: Dominik Leiss, Knut Schwippert, Timo Ehmke, Lena Heine, and Dietmar Hoettecke).

Author Contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Alina Knabbe, Dominik Leiss and Timo Ehmke. The first draft of the manuscript was written by Alina Knabbe and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The data collection and analysis described here was conducted as part of an interdisciplinary project funded by the German Research Foundation (DFG-project no. 417017613). After the end of the project period, the data will be made available to researchers on a national education server for further analysis.

Declarations

Ethical Approval All relevant ethical guidelines and principles were carefully considered in the preparation of this scientific article. The conduct of the research, as well as data collection, analysis, and interpretation, was performed in strict adherence to ethical standards to ensure that potential impacts on humans and the environment were minimized.

A comprehensive ethical evaluation was conducted prior to the study; this weighed all potential risks and benefits of the research. Any interaction with human participants was voluntary and informed consent was obtained. Participant privacy and confidentiality were always respected, and appropriate measures were taken to maintain anonymity. All study participants provided informed consent.

Competing Interests We have no conflict of interest to disclose.

References

Albarracín, L., Ferrando, I., & Gorgorió, N. (2021). The Role of Context for Characterising Students' Strategies when Estimating Large Numbers of Elements on a Surface. *International Journal of Science and Mathematics Education, 19*(6), 1209–1227. https://doi.org/10.1007/s10763-020-10107-4

Alsina, Á., & Salgado, M. (2022). Understanding Early Mathematical Modelling: First Steps in the Process of Translation Between Real-world Contexts and Mathematics. *International Journal of Science and Mathematics Education, 20*(8), 1719–1742. https://doi.org/10.1007/s10763-021-10232-8

Anderson, J. (1981, August). *Analysing the Readability of English and Non-English Texts in the Classroom with LIX.* Paper presented at the Annual Meeting of the Australian Reading Association, Darwin, Australia.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, Blum, & S. Khan (Eds.), *Mathematical Modelling Education, Engineering and Economics - ICTMA 12* (pp. 222–231). Woodhead. https://doi.org/10.1533/9780857099419.5.221

Leiss, D., Schukajlow, S., Blum, W., Messner, R., & Pekrun, R. (2010). The role of the situation model in mathematical modelling - Task analyses, student competencies, and teacher interventions. *Journal for Didactics of Mathematics, 31*(1), 119–141. https://doi.org/10.1007/s13138-010-0006-y

Plath, J., & Leiss, D. (2018). The impact of linguistic complexity on the solution of mathematical modelling tasks. *ZDM, 50,* 159–171. https://doi.org/10.1007/s11858-017-0897-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

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

Blomhøj, M., & Højgaard, T. (2007). What's all the Fuss about Competencies? In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education*: *The 14th ICMI Study* (pp. 45–56). Springer US. https://doi.org/10.1007/978-0-387-29822-1_3

Blum, W. (2011). Can modelling be taught and learnt? Some answers from empirical research. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling: ICTMA14* (Vol. 1, pp. 15–30). Springer Netherlands. https://doi.org/10.1007/978-94-007-0910-2

Blum, W. (2015). Quality teaching of mathematical modelling: What do we know, what can we do? In S. J. Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education: Intellectual and attitudinal challenges* (pp. 73–96). Springer International Publishing. https://doi.org/10.1007/978-3-319-12688-3_9

Boaler, J. (2022). *Mathematical mindsets: Unleashing students' potential through creative mathematics*. Jossey-Bass.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Taylor & Francis Group. https://doi.org/10.4324/9780203771587

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,* Article 348. https://doi.org/10.3389/fpsyg.2015.00348

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Fischer, L., Rohm, T., Carstensen, C. H., & Gnambs, T. (2021). Linking of Rasch-Scaled Tests: Consequences of Limited Item Pools and Model Misfit. *Frontiers in Psychology*, *12*, Article 633896. https://doi.org/10.3389/fpsyg.2021.633896

Fitzpatrick, C. L., Hallett, D., Morrissey, K. R., Yıldız, N. R., Wynes, R., & Ayesu, F. (2020). The relation between academic abilities and performance in realistic word problems. *Learning and Individual Differences, 83–84*, 101942. https://doi.org/10.1016/j.lindif.2020.101942

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29

Greefrath, G., & Vorhölter, K. (2016). Teaching and learning mathematical modelling: Approaches and developments from German speaking countries. In G. Greefrath & K. Vorhölter (Eds.), *Teaching and learning mathematical modelling* (pp. 1–42). Springer Cham. https://doi.org/10.1007/978-3-319-45004-9_1

Guerrero-Ortiz, C., Mena-Lorca, J., & Soto, A. M. (2018). Fostering transit between real world and mathematical world: Some phases on the modelling cycle. *International Journal of Science and Mathematics Education, 16*(8), 1605–1628. https://doi.org/10.1007/s10763-017-9856-9

Heinle, A., Schiepe-Tiska, A., Reinhold, F., Heine, J.-H., & Holzberger, D. (2022). Supporting student motivation in class: The motivational potential of tasks. *Zeitschrift Für Erziehungswissenschaft, 25*(2), 453–470. https://doi.org/10.1007/s11618-022-01090-3

Jankvist, U. T., & Niss, M. (2020). Upper secondary school students' difficulties with mathematical modelling. *International Journal of Mathematical Education in Science and Technology, 51*(4), 467–496. https://doi.org/10.1080/0020739X.2019.1587530

Kaiser, G. (2017). The teaching and learning of mathematical modelling. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 267–291). National Council of Teachers of Mathematics.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*(1), 109–129. https://doi.org/10.1037/0033-295X.92.1.109

Kultusministerkonferenz. (2022). *Bildungsstandards für das Fach Mathematik. Erster Schulabschluss (ESA) und Mittlerer Schulabschluss (MSA)* [Educational standards for mathematics. First school-leaving certificate (ESA) and intermediate school-leaving certificate (MSA)]. Retrieved from https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2022/2022_06_23-Bista-ESA-MSA-Mathe.pdf

Krawitz, J., Chang, Y.-P., Yang, K.-L., & Schukajlow, S. (2022). The role of reading comprehension in mathematical modelling: Improving the construction of a real-world model and interest in Germany and Taiwan. *Educational Studies in Mathematics, 109*(2), 337–359. https://doi.org/10.1007/s10649-021-10058-9

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Maaß, K. (2010). Classification scheme for modelling tasks. *Journal für Mathematik-Didaktik, 31*(2), 285–311. https://doi.org/10.1007/s13138-010-0010-2

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User Guide* (8th ed.). Muthén & Muthén. https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Niss, M., Blum, W., & Galbraith, P. L. (2007). Introduction. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI Study* (pp.3–32). Springer US. https://doi.org/10.1007/978-0-387-29822-1

Organization for Economic Co-operation and Development. (2013). *PISA 2012 Results: Excellence trough Equity*. Author. https://doi.org/10.1787/9789264201132-en

Organization for Economic Co-operation and Development. (2014a). *PISA 2012 Results: What students know and can do: Student performance in mathematics, reading and science* (Volume I). Author. https://doi.org/10.1787/9789264208780-en

Organization for Economic Co-operation and Development. (2019). *PISA 2018 results (Volume II): Where all students can succeed*. Author. https://doi.org/10.1787/b5fd1b8f-en
Organization for Economic Co-operation and Development. (2023a). *PISA 2022 assessment and analytical framework*. Author. https://doi.org/10.1787/dfe0bf9c-en
Organization for Economic Co-operation and Development. (2023b). *PISA 2022 results (Volume I): The state of learning and equity in education*. Author. https://doi.org/10.1787/53f23881-en
Organization for Economic Co-operation and Development. (1999). *Measuring student knowledge and skills: A new framework for assessment.* Author. Retrieved from https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33693997.pdf
Organization for Economic Co-operation and Development. (2014b). *PISA 2012 technical report*. Author. Retrieved from https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf
Peng, P., Lin, X., Ünal, Z. E., Lee, K., Namkung, J., Chow, J., & Sales, A. (2020). Examining the mutual relations between language and mathematics: A meta-analysis. *Psychological Bulletin, 146*(7), 595–634. https://doi.org/10.1037/bul0000231
Piel, S., & Schuchart, C. (2014). Social origin and success in answering mathematical word problems: The role of everyday knowledge. *International Journal of Educational Research, 66*, 22–34. https://doi.org/10.1016/j.ijer.2014.02.003
Plath, J. (2020). Verstehensprozesse bei der Bearbeitung realitätsbezogener Mathematikaufgaben: Klassische Textaufgaben vs. Zeitungstexte [Comprehension Processes During Solving Reality-based Mathematical Problems: Word Problems vs. Newspaper Texts]. *Journal für Mathematik-Didaktik, 41*(2), 237–266. https://doi.org/10.1007/s13138-019-00148-w
Pongsakdi, N., Kajamies, A., Veermans, K., Lertola, K., Vauras, M., & Lehtinen, E. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM Mathematics Education, 52*(1), 33–44. https://doi.org/10.1007/s11858-019-01118-9
Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. *Journal of Educational Psychology, 107*(3), 645–662. https://doi.org/10.1037/edu0000012
Reinhold, F., Hofer, S., Berkowitz, M., Strohmaier, A., Scheuerer, S., Loch, F., Vogel-Heuser, B., & Reiss, K. (2020a). The role of spatial, verbal, numerical, and general reasoning abilities in complex word problem solving for young female and male adults. *Mathematics Education Research Journal, 32*(2), 189–211. https://doi.org/10.1007/s13394-020-00331-0
Reinhold, F., Strohmaier, A., Hoch, S., Reiss, K., Böheim, R., & Seidel, T. (2020b). Process data from electronic textbooks indicate students' classroom engagement. *Learning and Individual Differences, 83–84*, 101934. https://doi.org/10.1016/j.lindif.2020.101934
Schmidt, S., Ennemoser, M., & Krajewski, K. (2012). *DEMAT 9*. Hogrefe.
Schwippert, K. (2019). Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft [What's about the books? Social background of students and educational opportunities from the perspective of international large-scale surveys]. *Journal for Educational Research Online, 11*(1). https://doi.org/10.25656/01:16789
Strohmaier, A. R., Reinhold, F., Hofer, S., Berkowitz, M., Vogel-Heuser, B., & Reiss, K. (2021). Different complex word problems require different combinations of cognitive skills. *Educational Studies in Mathematics, 109*(1), 89–114. https://doi.org/10.1007/s10649-021-10079-4
Sullivan, P., Mousley, J., & Jorgensen, R. (2009). Tasks and pedagogies that facilitate mathematical problem solving. In B. Kaur, B. H. Yeap, & M. Kapur (Eds.), *Mathematical problem solving* (pp. 17–42). World Scientific. https://doi.org/10.1142/9789814277228_0002
Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition, 33*(3), 339–372. https://doi.org/10.1017/S0272263111000015
Ufer, S., & Bochnik, K. (2020). The role of general and subject-specific language skills when learning mathematics in elementary school. *J Math Didakt, 41*(1), 81–117. https://doi.org/10.1007/s13138-020-00160-5
van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.
Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM Mathematics Education, 52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4
Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making Sense of World Problems*. Swets & Zeitlinger.

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology, 28*(4), 409–426. https://doi.org/10.1080/01443410701708228

Vorhölter, K. (2023). Metacognition in mathematical modeling: the connection between metacognitive individual strategies, metacognitive group strategies and modeling competencies. *Mathematical Thinking and Learning, 25*(3), 317–334. https://doi.org/10.1080/10986065.2021.2012740

Wijaya, A., van den Heuvel-Panhuizen, M., Doorman, M., & Robitzsch, A. (2014). Difficulties in solving context-based PISA mathematics tasks: An analysis of students' errors. *The Mathematics Enthusiast, 11*(3), 555–584. https://doi.org/10.54870/1551-3440.1317

Wijaya, A., van den Heuvel-Panhuizen, M., & Doorman, M. (2015). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics, 89*(1), 41–65. https://doi.org/10.1007/s10649-015-9595-1

Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACER ConQuest version 2.0: Generalised item response modelling software*. ACER Press.

Yin, L., & Fishbein, B. (2020). Creating and interpreting the TIMSS 2019 Context Questionnaire Scales. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and procedures: TIMSS 2019 Technical Report* (pp. 16.1–16.331). IEA. Retrieved from https://timssandpirls.bc.edu/timss2019/methods/chapter-16.html

Declaration of originality This manuscript has not been published or presented elsewhere in part or in entirety and is not under consideration by another journal. We have read and understood your journal's (ethical) policies, and we believe that neither the manuscript nor the study violates any of these.