International Electronic Journal of Mathematics Education
2023, 18(3), em0744
e-ISSN: 1306-3030
https://www.iejme.com

MODESTUM

Research Article **OPEN ACCESS**

Taking notes as a strategy for solving reality-based tasks in mathematics

Lisa-Marie Wienecke [1*], Dominik Leiss [1], Timo Ehmke [1]

[1] Leuphana University Lueneburg, Lueneburg, GERMANY
*Corresponding Author: lisa-marie.wienecke@leuphana.de

Citation: Wienecke, L.-M., Leiss, D., & Ehmke, T. (2023). Taking notes as a strategy for solving reality-based tasks in mathematics. *International Electronic Journal of Mathematics Education, 18*(3), em0744. https://doi.org/10.29333/iejme/13312

ARTICLE INFO

Received: 03 Mar. 2023

Accepted: 17 May 2023

ABSTRACT

This study investigates the extent to which student and task-related characteristics are associated with different types of note-taking and analyzes how task success depends on these elements. For this purpose, a sample of n=866 students (age: mean=13.99) completing two reality-based tasks as part of a paper and pencil test are considered. The results demonstrate that the note-taking type differs significantly between the two parallel constructed tasks. For example, language skills (r=.26), interest in mathematics (r=.13), and the socio-economic statuses (r=.12) are observed to be significantly correlated to greater note-taking frequency. Based on linear regression (dependent variable: successful task solution), 34% of the variance is attributed to note-taking and other student characteristics. The most relevant predictor for a successful task solution (β=.36) is notes containing an elaboration of the given task information.

Keywords: note-taking, learning strategy, word problems, mathematical modelling, elaboration, reading competence

INTRODUCTION

Long before mathematical maturity was popularized (Boaler, 2022), efforts were made to reform mathematical teaching and task culture. Word problems (mathematical problems presented in words) have always been central to mathematical study in schools everywhere (Verschaffel et al., 2020). Verschaffel et al. (2020) reported that word problems were found in ancient Chinese and Indian manuscripts as well as Egyptian papyri. In more recent times, real-world functional applications have become the primary goal of mathematical education (Niss, 2016). Application-oriented mathematical instruction involves a broad spectrum of word problems and has been researched extensively in mathematical didactics for several decades (e.g., Blum et al., 2007; Brown, 2019; Frejd, 2013; Pollak, 1979). Reality-based tasks generally motivate students to understand and structure real-life situations using mathematical tools to solve underlying problems (Blum & Leiss, 2007). Numerous empirical studies have demonstrated that the associated translation process between reality and mathematics is associated with cognitive barriers (Galbraith & Stillman, 2006; Wijaya et al., 2014) and that understanding real problem situations is a source of concern for students (Leiss et al., 2010). Accordingly, to equip students with mathematical application competencies, they must be trained on reading (Schnotz & Dutke, 2004) and related learning strategies. One promising strategy is note-taking to understand problem situations provided in textual form. According to Lonka et al. (1994) and Slotte et al. (2001), note-taking can be classified into several types in this context, corresponding to elaboration strategies used during different stages of the problem-solving process, viz., dealing with the problem situation intensively, making the presented information explicit, and restricting it in a goal-oriented manner (Staub, 2006). However, differential analyses on the design of notes and the specific influence of note-taking on the solution process while working on reality-based mathematical tasks represent a research desideratum. This is the motivation for the present study, which investigates the extent to which personal and task-related characteristics are associated with different types of note-taking. Furthermore, it analyzes the dependence of task solution success on note-taking type.

STATE OF RESEARCH

Incorporating reality-based tasks and debating whether embedding problems in real-world contexts is beneficial or detrimental to mathematical learning are longstanding questions in mathematical education (Brown, 2019). They have been researched extensively as part of mathematical didactics ever since educational systems began to prioritize word problems to train students in real-world application (Reinhold et al., 2020). In the scientific community, different terms and definitions are used for different types of reality-based tasks (Verschaffel et al., 2020). They comprise a continuum between "dressed-up" word problems and modelling tasks (Leiss et al., 2019; Verschaffel et al., 2020). Herein, based on the study by Strohmaier (2020), we consider word problems as tasks that present information primarily in textual form and represent functional and realistic contexts besides the mathematical one. In contrast, more complex tasks containing, for example, non-essential information or different representations are described as mathematical modelling tasks (Strohmaier, 2020).

Reality-Based Mathematics Teaching

For over three decades, application-oriented mathematics teaching has been a central topic in mathematical didactics (Blum et al., 2007; Kaiser & Schukajlow, 2022). Reality-based tasks are a particularly relevant type of mathematical task. In addition, the act of presenting personalized and interesting tasks has been reported to enhance perception of the relevance of mathematics (Gijsbers et al., 2020) and increase the solution rates of tasks (Davis-Dorsey et al., 1991; Walkington, 2013). In these reality-based tasks, students have to filter and structure relevant aspects from the real-life situation described to solve the task. Essentially, students must interpret and understand a real situation to arrive at a solution. This requirement promotes logical thinking, constructive learning (van den Heuvel-Panhuizen, 2003) creativity, and critical questioning (UNESCO, 2017) through reflection and communication. In addition, these real-world contexts should be used as a didactic tool to support mathematical learning (Niss et al., 2007; Paredes et al., 2020; Wijaya et al., 2015). However, reality-based tasks should be clearly distinguished from intra-mathematical tasks, which have no relation to reality and only require mathematical procedures to solve them (Niss et al., 2007; Palm, 2002; Verschaffel et al., 2020). Different mathematical tasks represented in words can be classified on a spectrum of authenticity on which traditional word problems, as described in Cummins et al. (1988) and Verschaffel et al. (2000), represent the negative pole and modelling tasks, as described in Kaiser (2017), represent the positive pole (Verschaffel et al., 2020).

Moreover, Daroczy et al. (2015) and Reinhold et al. (2020) reported that reality-based tasks require special linguistic skills. In particular, they established that the complexity of a situation and its linguistic formulation is proportional to the difficulty of its comprehension and solution (Leiss et al., 2019). However, even when all information is acquired correctly, inadequate cognitive function of language comprehension can obstruct further processing. Dryvold et al. (2015) theorized that, in addition to general reading competence, subject-specific text comprehension is fundamental to effectively translate mathematical concepts into real solutions.

Depaepe et al. (2010) proposed shifting a part of the focus during the solution of reality-based tasks from intra-mathematical aspects to contextual and linguistic elements. Nevertheless, intra-mathematical skills are equally important, which has been demonstrated repeatedly (Daroczy et al., 2015; Fuchs et al., 2006; Leiss et al., 2010; Pongsakdi et al., 2020; Reinhold et al., 2020; Schnotz & Dutke, 2004). Leiss et al. (2010) related the term "reality-based" to the complexity of a task without its question and question complexity, revealing a correlation between the complexity of a situation (situation value) and the corresponding reference to reality. The study reports, however, that creating problems with high situation value and high empirical difficulty tends to be challenging (Leiss et al., 2010). Consequently, working on reality-based tasks requires a variety of special competencies, characterized, among other factors, by the translation processes between the real world (extra-mathematical context) and mathematics (intra-mathematical content) (Blomhøj & Jensen, 2007; Blum & Leiss, 2007; Borromeo Ferri, 2006; Matos & Carreira, 1995; Schukajlow et al., 2018).

Thus, reality-based tasks require a significant degree of maneuvering during task creation and processing, which can introduce difficulties for teachers and students. Various empirical studies have demonstrated that the translation process can lead to the introduction of cognitive barriers in the different phases (Clarkson, 1991; Galbraith & Stillman, 2006; Kaiser et al., 2011; Leiss et al., 2010; Mayer & Hegarty, 1996; Stillman et al., 2010; Wijaya et al., 2014). Therefore, the phase of comprehension must be the focus of research (Blum & Leiss, 2007). Errors in comprehension can affect subsequent decisions severely and induce further errors in subsequent phases (Clarkson, 1991; Leiss et al., 2010; Mayer & Hegarty, 1996; Wijaya et al., 2014).

Moreover, according to Kintsch and van Dijk (1978), the critical linguistic steps involved in solving word problems are (1) constructing a mental representation (situation model) and (2) transforming it into a mathematical model. This observation was complemented by Schukajlow et al. (2012) and Leiss et al. (2019), who demonstrated that constructing a situation model is critical for correctly solving these tasks and that it generally requires approximately 40% of the total solution time. Similarly, solution performance also relies significantly on the linguistic complexity of the task. These situation models rely on task features, such as semantic structure and thematic context, as well as on intra-personal characteristics, such as the reading competence of students and their prior contextual knowledge, which form the basis for subsequent task processing (Leiss et al., 2010; Reusser, 1988).

Thus, authentic tasks in mathematical education require specific language skills (Daroczy et al., 2015), particularly text comprehension (Fuchs et al., 2018; Pongsakdi et al., 2020) and reading comprehension (Lee et al., 2004; Leiss et al., 2019). For example, students experience difficulty in determining and organizing relevant objects or identifying appropriate mathematical objects and solution methods (Rellensmann et al., 2020). To overcome the aforementioned difficulties, Vorhölter (2018) and Schukajlow et al. (2021) recommended metacognitive strategies. Staub (2006) explained the special utilization of note-taking as a metacognitive elaboration strategy.

Note-Taking as a Solution Method

Writing is a student-initiated strategy that serves as a basis for note-taking. In schools, teachers often employ writing to support content learning (Gillespie et al., 2014; Ray et al., 2016).

Students who take notes are known to perform better in reading comprehension than those who do not (Chang & Ku, 2015; Graham et al., 2020). Graham et al. (2020) published a meta-analysis, revealing that asking students to write down content in mathematics classrooms improves learning. Writing improves students' understanding and application of content knowledge and facilitates the learning of material that is not the direct focus of instruction (Graham et al., 2020). Chang and Ku (2015) attributed this effect to the fact that, as explained in the previous section, note-taking helps students retain what they read and make connections between individual pieces of textual information. However, the nature of note-taking and its characteristics merit further investigation.

For this purpose, a product of note-taking is studied: notes. According to Staub (2006), various forms of notes can be used at diverse points of the solution process to express represented information explicitly. They could also serve as external storage for information that can be referred to during later stages of the solution process (Staub, 2006). The following two aspects are central to the function of notes: extracting information from the text of a task described by encoding or explicating the textually represented data that are essential to the solution and elaborating and expanding the given information and data. According to Staub (2006), the second facet is more interesting.

Herein, we focus on the recording of notes, called *note-taking,* as a problem-solving strategy as described in Rogiers et al. (2020), among others. Using (metacognitive) strategies, one can identify different student profiles. Out of all the profiles, the information organizers and integrated strategy users are conspicuous. They frequently use the strategies of summarizing, paraphrasing, elaborating, and highlighting (Rogiers et al., 2020). Rovers et al. (2019) found that performance approach goals do not predict strategies, while mastery goals predict strategies associated with organization and elaboration. Among other things, marking is related to understanding situation models (Rovers et al., 2019). These strategies can be divided into overt cognitive strategies, such as marking texts, and covert cognitive strategies, such as paraphrasing or elaborating (Rogiers et al., 2020). For example, deep strategies such as elaboration and organization are reportedly more effective than more superficial strategies such as rehearsal (Pintrich et al., 1993); even within these categories, certain strategies are likely to be more effective than others (Rovers et al., 2019). For example, organizational strategies lend themselves to capturing connections between constructs. However, to capture global theories and systems, elaboration strategies are more appropriate. Moreover, different strategies may be more effective and/or better aligned at different stages of learning (Greene & Azevedo, 2007). In particular, note-taking and summary writing (organization) are linked to comprehension of expository texts (Samuelstuen & Bråten, 2007).

Similarly, Weinstein and Mayer (1986) categorized learning strategies into memorization, organization, elaboration, and monitoring. Memorization involves using notes as an external memory. Organization refers to relating, grouping, or ordering information and ideas from the learning material (e.g., summarizing, outlining, and diagramming textual information). Elaboration serves to make the material more meaningful using connections between the information contained in the material and that from other sources (e.g., by linking it to relevant prior knowledge). Finally, in monitoring, individuals assess or regulate their learning (e.g., confirming comprehension, identifying the problem, and arriving at a solution) (Bråten & Samuelstuen, 2007). For learning from texts, research (National Reading Panel, 2000; Trabasso & Bouchard, 2002) indicates that organization, elaboration, and monitoring, are particularly important for successful performance (Bråten & Samuelstuen, 2007). Bråten and Samuelstuen (2007) focused on marking, note-taking, and summary writing in their study, because the literature (e.g., Chang & Ku, 2015; Hadwin & Winne, 1996; Jamieson-Noel & Winne, 2003; King, 1992; Kiewra, 1989; Kiewra et al., 1991; Kobayashi, 2005; Lahtinen et al., 1997) credits these strategies with promoting comprehension. Another reason was the uniqueness of the use of the chosen strategies, which was directly investigated in their study.

For mathematics and applied tasks described using a factual text, several studies have demonstrated that both the process and product of note-taking are beneficial to learning (Chang & Ku, 2015; Kiewra, 1989; Kiewra et al., 1991; Kobayashi, 2005). In particular, self-generated informal drawings support the translation process (between the real world and mathematics) by enabling students to include their prior domain knowledge more naturally than in a formal and abstract model (Leenaars et al., 2013). Furthermore, conclusive evidence exists to support the positive influence of note-taking on reading comprehension, memory, and retention (Laidlaw et al., 1993; Leiss et al., 2019; Kiewra et al., 1995; Spires & Stone, 1989). Leenaars et al. (2013) and Piolat et al. (2005) also demonstrated that note-taking promotes better comprehension of information, and consequently, of the situation. Thus, an adequate situation model can be designed based on note-taking, which has a positive effect on the solution process (Leenaars et al., 2013). Schukajlow et al. (2012) also identified note-taking as an essential tool in the solution process.

In his study of notes, Staub (2006) isolated features such as rendering the situation in one's own words, summaries, graphs, tables, conceptual networks, and matrix representations to describe the qualitative processing of notes. Two overarching qualitative features in any form of note-taking are the completeness and the generative transformation of the notes. These are fundamental to contextualizing individual statements and developing an adequate model of each situation (van Dijk & Kintsch, 1983). Lonka et al. (1994) and Slotte et al. (2001) corroborate the hypothesis that any strategy with the aim of learning a small detail reinforces its learning, whereas central ideas are learned independent of strategies. Similarly, they demonstrated that underlining is correlated to success in tasks requiring deep comprehension of the text. However, concept mapping, that is, structuring representations, is a successful strategy for solving tasks that require a critical review of the learned material. Notes can also be classified based on the aspects of content, design, and performance.

According to Chang and Ku (2014), note-taking can be classified as active learning behavior. Passive learning behavior involves merely reading a text, whereas active learning involves underlining or highlighting aspects of the text or taking *verbatim* notes. Constructive learning behavior involves active note-taking using one's own words and associated self-explanation attempts. For example, Cox (1999) reported a self-explanation effect that is more pronounced in higher performing students who take notes in their own words and enhance them with connections. Meanwhile weaker students tend to repeat textual contents word-for-word in external representations (Cox, 1999).

Summarizing and concept mapping are superior to *verbatim* note-taking, marking, and plain reading according to various studies (Lahtinen et al., 1997). However, most learners tend to take notes in a superficial, *verbatim*, and linear structure (Boch & Piolat, 2005; Lahtinen et al., 1997). Other possible structures include matrix forms (Atkinson et al., 1999) in which, for example, information is tabulated as juxtaposed notes. Buzan (1983) developed mind-mapping as a special note-taking technique that includes topic- or problem-related information and ideas, presented in a structured manner. Generally, concept maps are special diagrams that depict topic-related concepts and their interrelationships within a hierarchical structure (Brinkmann, 2005). Numerous studies have demonstrated that using concept and mind maps is conducive to mathematical reasoning (Brinkmann, 2003, 2005; Entrekin, 1992; Malone & Dekkers, 1984; Novak, 1990, 1996). Thus, the structuring of information is considered in the present study.

The different strategies discussed here can be classified using the following gradations comparable to elaboration scores (see Methods section) (adapted from Lahtinen et al., 1997; Lonka et al., 1994; Slotte et al., 2001):

(0) no recognizable notations,

(1) marking of parts of the text,

(2) active note-taking in the form of verbatim quotations,

(3) constructive note-taking in the form of summarizing paraphrases.

Note-Taking While Solving Reality-Based Tasks

The primary function of notes in this case is to quickly organize, structure, and elaborate information. However, note-taking affects different reality-based mathematical tasks based on their type. Therefore, several task types are considered in this study to provide a broader picture of the state of research. These task types differ in some categories from the ones presented here. Nevertheless, they all have the application and the textual form of presentation in common.

For example, Hegarty and Kozhevnikov (1999) found that the use of schematic representations is related to the solution success of reality-based tasks, whereas pictorial representations are negatively correlated with success; Boonen et al. (2014) corroborated this result. Based on these two studies, we can conclude that a more detailed differentiation of notes is required to draw more meaningful conclusions. A finer classification by Verschaffel et al. (2020) revealed that correct schematic representations have a positive influence on the solution success of reality-based tasks. Further, Pape (2004) reported examples, where implementing note-taking as a reading strategy leads to a better understanding of the situation while solving reality-based tasks. A literature review also revealed three studies in particular devoted to note-taking in reality-based tasks (Leiss et al., 2019; Plath, 2017, 2020). All of these studies described note-taking as a strategy for comprehension and structuring. Plath (2020) demonstrated that note-taking is used in approximately 40% of all word problems and in more than 60% of newspaper-text tasks. They concluded that the use of comprehension strategies and the correctness of the solutions are unrelated. Leiss et al. (2019) reported that the comprehension process accounts for 40% of the entire solution process. Regarding the use of comprehension strategies, all three phases of the comprehension strategies were identified; in contrast to all other strategies, note-taking exhibited significant value as a comprehension strategy ($r=.263$) when correlated with the construction of a suitable situation model, which in turn was correlated with the successful solution of the task ($r=.116$). Consequently, the research conclusions on the correlation between note-taking and the successful solution of reality-based tasks are ambiguous. These studies demonstrated that learners also spontaneously use certain familiar reading strategies, such as marking, taking notes and writing down details, repeated reading, or pictorial imagining, while solving mathematical problems (Plath & Leiss, 2018). However, the tasks' solution rates remained unaffected by these strategies (Leiss et al., 2019; Plath, 2020). Moreover, Plath (2017, 2020) found no statistically significant correlations between successful solution and comprehension strategies, whereas Leiss et al. (2019) demonstrated a significant positive relationship.

Note-taking while solving reality-based tasks is supported in German schools by the strategy instrument "question, calculation, response." It is implemented to enable students to reflect actively on the aim of the task, possible methods of solution, and the actual solution itself. Thus, relevant actions include discerning the information provided in the text, identifying the information to be found, and identifying a method to obtain it. Using this methodology, students learn to use the varying outlines to structure their notes on a task (Buschmeier, 2017; Rinkens & Dingemans, 2014). In summary, the students do not merely write down the calculation of the solution but also note down all relevant information in advance.

In addition to the growing body of research on reality-based mathematical tasks, in general, and on sketches used in mathematical education, in particular (Rellensmann et al., 2020), numerous studies have been conducted on the general use of note-taking to promote long-term learning and note-taking while listening (Chen, 2021; Hagen et al., 2014; Witherby & Tauber, 2019). However, few studies have investigated how note-taking helps students to complete mathematical tasks in performance-oriented situations such as the classroom.

Furthermore, certain individual characteristics could influence the solution success rate of mathematical modelling tasks.

For example, the effect of gender differences on to mathematical education has already been investigated (for an overview: Reilly et al., 2015) in a few school achievement studies (Ludwig & Reit, 2013). Over the past two decades, meta-analyses have revealed small but stable mean gender differences in favor of men. However, the aforementioned studies (Leiss et al., 2019) have generally failed to establish a gender disparity in terms of reality-based task solving and note-taking as a tool for comprehension. Ludwig and Reit (2013) were unable to identify significant gender differences affecting the solution process of reality-based tasks.

Supplementary to these numerous studies (Fuchs et al., 2008; OECD, 2013, 2019; Plath & Leiss, 2018; Walkington et al., 2018), related research has indicated that the social background of students (e.g., parental education, migration background) is correlated to their success in solving mathematical problems with real-world contexts.

Research Questions and Hypotheses

Based on the above-mentioned research background, following four research questions were selected as the focus of this study:

(1) Which quantitative and qualitative characteristics of note-taking are exhibited by students during the solution of reality-based mathematical tasks?

(2) What is the extent of the relationship between taking different types of notes and solving reality-based mathematical tasks correctly?

(3) What is the extent of the dependence of (a) note-taking during the solution of reality-based mathematical tasks and (b) the degree of elaboration of those notes on specific student characteristics?

(4) In addition to personal characteristics, to what extent does taking (elaborated) notes explain additional variance with regard to the solution of reality-based mathematical tasks?

Regarding the first question, we assume that, irrespective of the students' personal preferences, specific task characteristics, such as the mathematical quantities encountered, the types of mathematical units used, or the situational complexity, exhibit varying influences on the type of notes used. In addition, we wish to note that the first question seeks clarification regarding the differences in quantitative and qualitative characteristics of note-taking between two reality-based mathematical tasks.

The second question addresses the effect of different types of notes on the successful processing of reality-based mathematical tasks. We assume that the use of notes tends to have a positive effect on the solution of mathematical tasks. However, to be precise, noting irrelevant information is associated with lower solution rates, whereas independent data processing or elaboration is associated with higher solution rates.

In the case of the third question, we assume that, in addition to specific task characteristics, personal characteristics, attitudes, and competencies also influence the process and final product of note-taking. In other words, the style of note-taking as well as the characteristics of the notes are influenced by the person taking notes. We assume that better school grades in German and mathematics and individuals belonging to a school type with a tendency toward a more performance-oriented milieu are correlated to increased note-taking (because higher performing students take notes as paraphrasing summaries). Furthermore, we assume that better language skills (in relation to the special linguistic competencies required) as well as higher interest in the task context is correlated to increased note-taking and increased note quality (because higher interest increases the likelihood of a correct solution). Concerning gender-specific effects on the use of note-taking as a learning strategy, the current conclusions are ambiguous. Thus, no justified hypotheses are assumed.

In the case of the fourth question, we assume that, in addition to empirically established influencing factors (e.g., gender, prior knowledge, and linguistic competence) even note-taking as a mere repetition strategy influences solution success positively. In addition, notes used as an elaboration strategy to present data in a goal-oriented structure exert an even stronger positive effect on the solution of mathematical tasks.

METHODS

To answer the aforementioned research questions, we used a quantitative design to analyze student solutions corresponding to two different reality-based mathematical tasks. Both reality-based problem situations were included in a broader achievement test. The entire paper-and-pencil achievement test was administered as a 90-min individual written test in a classroom and included various subject tasks (mathematics, physics, physical education, music, and German), related questions of interest, a questionnaire on the background characteristics of students, and a language test. A sample of 1,343 students in grades 7 and 8 from 17 schools participated in the test of which 34% reported a migration background (that is, neither parent was born in Germany). Each participant was assigned one of the two tasks.

Of the 1,343 students, only 866 started to work on the task using strategies such as marking content in the task text, which resulted in the following distribution: 407 students (♀: 51.8%, ♂: 48.2%, age: mean [M]=14.00 (standard deviation [SD]=0.78), grade[1] math: M=3.35 (SD=2.06)) attempted to solve the reality-based task, *Oven*, and 459 students (♀: 46.3%, ♂: 53.7%, age: M=13.98 (SD=0.79), grade math: M=3.20 (SD=2.02)) attempted to solve the reality-based task, *Petrol Station*. The 477 students who did not complete the tasks because of time constraints or other reasons were removed from the evaluation. In each case, one of the two reality-based tasks was contained in the test booklets randomly distributed to the students; each test booklet comprised three parts (see **Figure 1**).

Part 1 contained a C-test as a language test; part 2 comprised 13 subject tasks and related questions of interest, and part 3 contained the questionnaire with questions on gender, age, country of birth, school grades in mathematics, German, sport, music, and socio-economic status (SES), among others.

To provide a clear timeline of the study, refer to the flowchart in **Figure 2**; the students received a detailed introduction and instructions from trained test administrators to maintain consistency in possible influences. Subsequently, the students worked on the parts that were divided into time sections, as described earlier.

In the following section, the test portions relevant to the present study as well as the corresponding codings are described.

---

[1] In German school system, grades run from 1 to 6, with 1 being the best grade awarded and 6 being the worst and 4 is the passing grade.

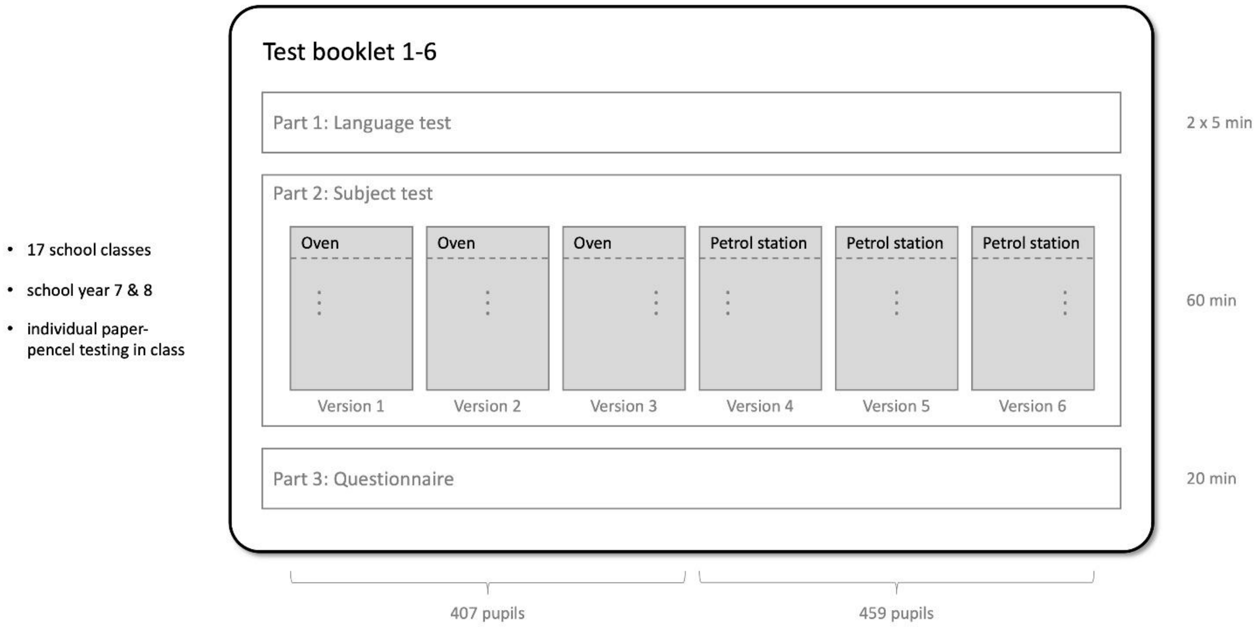


Figure 1. Test design (Source: Authors' own elaboration)

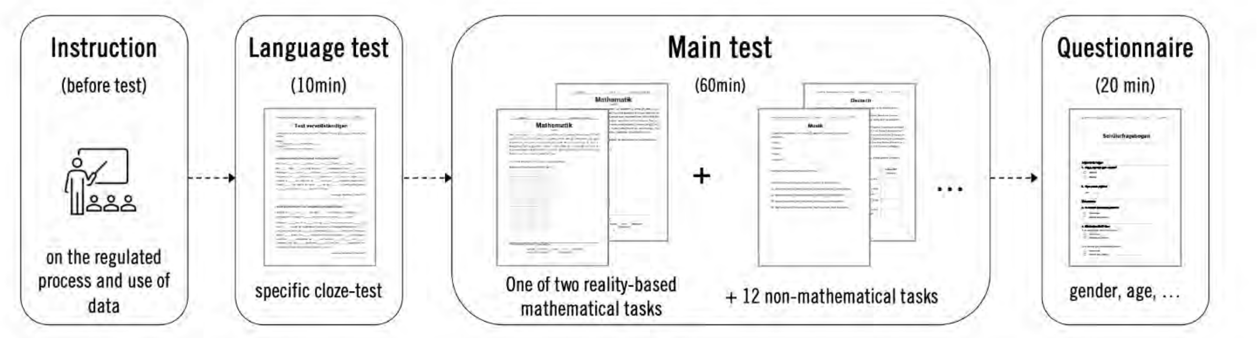


Figure 2. Timeline of the study (Source: Authors' own elaboration)

Part 1–Language Test

A modified C-test[2] was used to assess the written language skills of students (Baghaei & Grotjahn, 2014). For this purpose, the pupils were presented with two short texts in German with word counts of 92 words each. The topics of the two texts were "The capital, Lisbon" and "After-work activities." In both texts, the second halves of every second word in each sentence except for the first and the last one were systematically deleted.

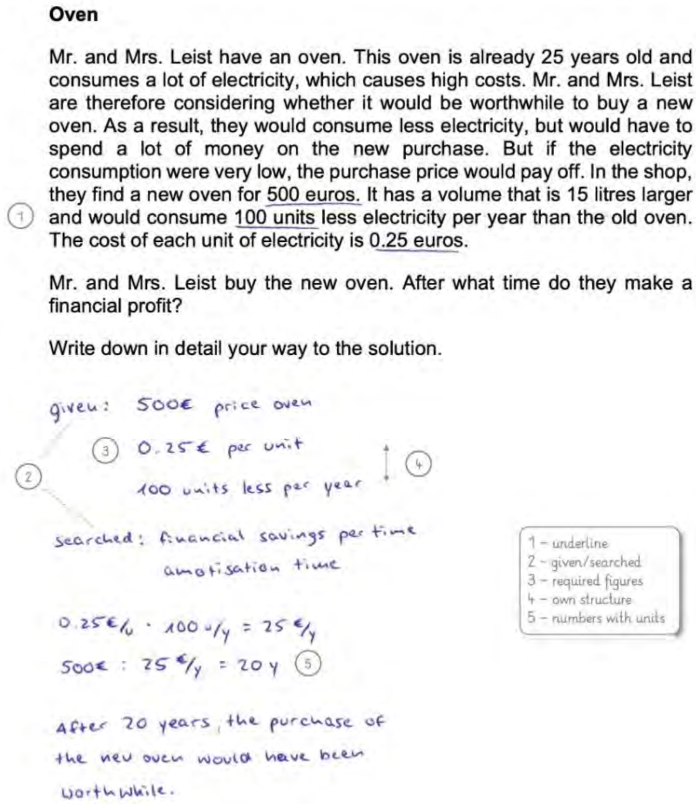
The following is an excerpt from "The Capital, Lisbon" (first 15 gaps):

> … Ein Groß teil $_{1}$ ihres Cha rmes $_{2}$ rührt v on $_{3}$ ihrer maler ischen $_{4}$ Lage h er $_{5}$. So li egt $_{6}$ die südeurop äische $_{7}$ Metropole z um $_{8}$ einen a n $_{9}$ einer Bu cht $_{10}$ des Flu sses $_{11}$ Tejo, andere rseits $_{12}$ ist d as $_{13}$ Stadtbild du rch $_{14}$ sanfte Hü gel $_{15}$ geprägt ... (German original–gaps underlined).
>
> ... Much of its charm stems from its picturesque location. On the one hand, the southern European metropolis lies on a bay of the River Tagus, on the other hand, the cityscape is characterized by gentle hills ... (English translation without gaps).

Pupils were required to fill in these 2×30 gaps within 2×5 min; the correct completion of the word element in terms of the content, and not its orthographic correctness, was assessed. Because the individual gaps can be scaled as largely independent test items (Harsch & Hartig, 2010), only two such texts from a tried-and-tested C-test from the Hamburg Institute for educational monitoring and quality development (IfbQ) were used for test-economic. An AEAP/PV test reliability of 0.90 was noted, which reveals the high internal consistency of the measurement scale.

---

[2] A C-test is a specific version of a close test format, well-established as a reliable and valid measure of general language proficiency (Grotjahn & Drackert, 2020; Grotjahn et al., 2002). The score of a C-test is based on the proficiency with which students correctly identify the searched word individually for each gap. Based on the case, the orthographic and grammatical correctness of the response is evaluated.

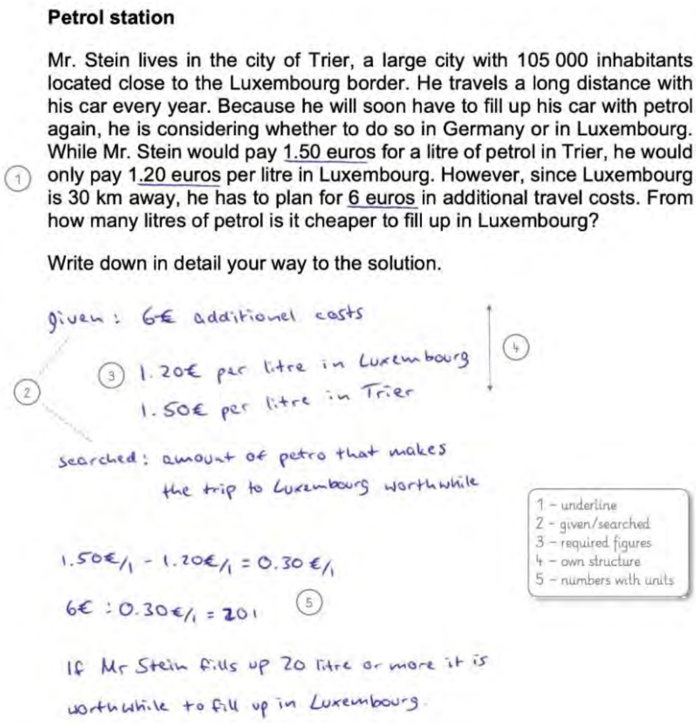

Oven

Mr. and Mrs. Leist have an oven. This oven is already 25 years old and consumes a lot of electricity, which causes high costs. Mr. and Mrs. Leist are therefore considering whether it would be worthwhile to buy a new oven. As a result, they would consume less electricity, but would have to spend a lot of money on the new purchase. But if the electricity consumption were very low, the purchase price would pay off. In the shop, they find a new oven for 500 euros. It has a volume that is 15 litres larger and would consume 100 units less electricity per year than the old oven. The cost of each unit of electricity is 0.25 euros.

Mr. and Mrs. Leist buy the new oven. After what time do they make a financial profit?

Write down in detail your way to the solution.

Figure 3. Example for note-taking & corresponding solution: Task 1–*Oven* (Source: Authors' own elaboration)


Petrol station

Mr. Stein lives in the city of Trier, a large city with 105 000 inhabitants located close to the Luxembourg border. He travels a long distance with his car every year. Because he will soon have to fill up his car with petrol again, he is considering whether to do so in Germany or in Luxembourg. While Mr. Stein would pay 1.50 euros for a litre of petrol in Trier, he would only pay 1.20 euros per litre in Luxembourg. However, since Luxembourg is 30 km away, he has to plan for 6 euros in additional travel costs. From how many litres of petrol is it cheaper to fill up in Luxembourg?

Write down in detail your way to the solution.

Figure 4. Example for note-taking & corresponding solution: Task 2–*Petrol Station* (Source: Authors' own elaboration)

Part 2–Subject Test

The students were assigned 60 min to complete the second part of the test. Based on a rotating framework of six different versions of part 2, 13 subject tasks and related questions of interest had to be completed within this time. However, only the two tasks are of interest to the current study. Three versions of the question booklet contained the reality-based task, *Oven*, and three versions contained the reality-based task, *Petrol Station*. The two tasks are depicted with corresponding sample solutions in **Figure 3** and **Figure 4**; ideally, all potentially performance-enhancing features of notes are integrated.

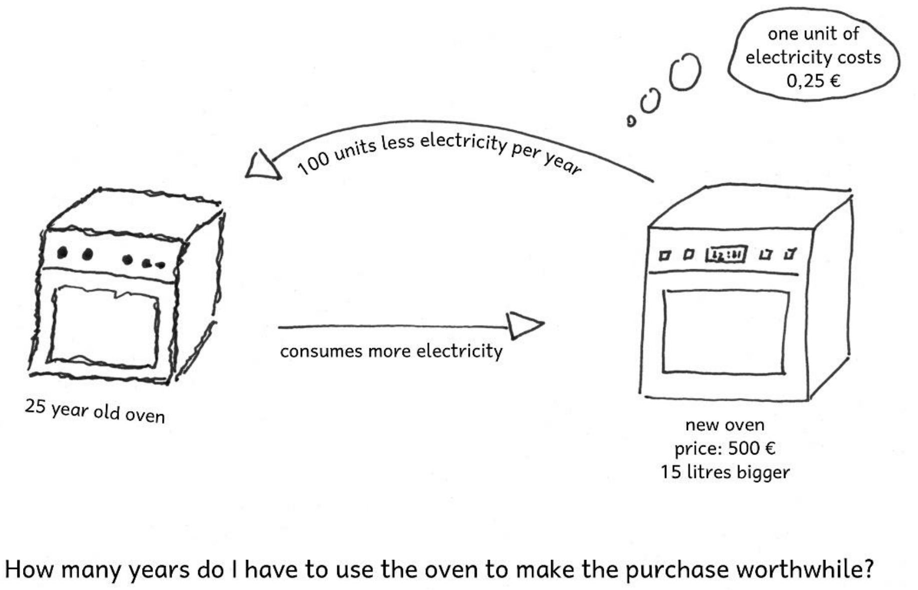


Figure 5. Possible situation model: Task 1–*Oven* (Source: Authors' own elaboration)

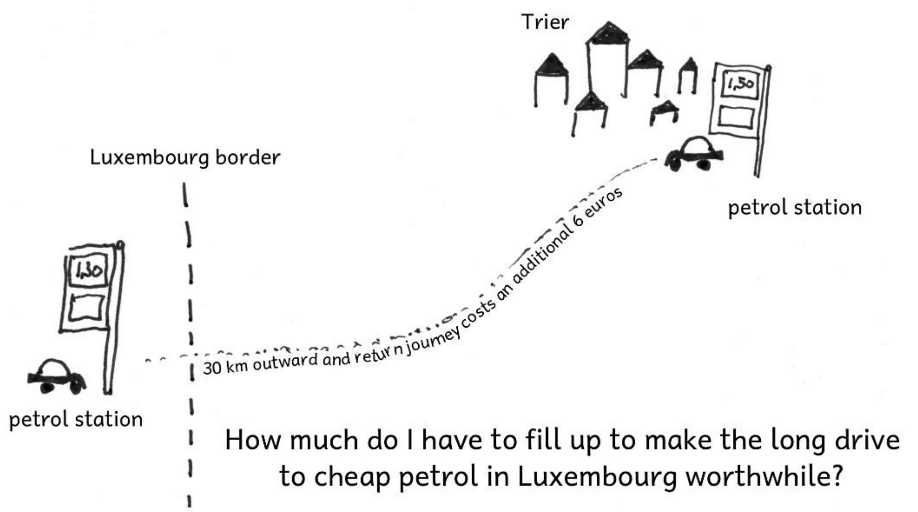


Figure 6. Possible situation model: Task 2–*Petrol Station* (Source: Authors' own elaboration)

These two reality-based tasks are similar to the word problems introduced by Verschaffel et al. (2020); they are primarily mathematical problems based on a real-world context of financial decision-making that have been posed in linguistic form as mentioned in the theory section. Moreover, they contain both essential as well as irrelevant numerical data. To solve the problem, students have to put themselves in the presented situation, extract information from the text, and use it to find a solution. Mechanical manipulation of the numbers extracted from the text will not lead students to the solution; understanding of the connotations of each figure is essential.

The primary aim during the construction of these two tasks was the construction of parallel requirements with respect to the central mathematical idea (functional relationship), the required mathematical competences (modelling & technical work), the number of relevant (#3) and irrelevant (#2) figures, the complexity of the required quantities and units (monetary amounts as decimal numbers and euros per unit or euros per liter), the empirical degree of difficulty (40%-50%), as well as the basic content-related problem (financial decision-making task). The tasks were intended to be sufficiently complex to motivate a majority of the students to use notes as support strategies. Although this parallelization succeeded in terms of both the subject didactic categories and the level of difficulty (38.33% for *Oven* and 38.56% for *Petrol Station*), we expected significant differences in terms of the two central aspects.

Furthermore, the contexts of the two tasks are different. For example, pupils are likely to have limited experience of how the purchase of an oven can be profitable based on savings in electricity costs. The context of the other task, that of a car that can be refueled at different petrol stations at different rates, is more commonplace, at least passively. From this perspective, the task *Oven* could be more challenging than task *Petrol Station* (**Figure 5** and **Figure 6**) in terms of the construction of a situation model.

Table 1. Significant codes & their derivation

Codename & general description	Examples for a score 1
Structuring elements: Even in German primary school, the curriculum includes a structured approach to mathematical word problems. However, such approaches are not retained by the pupils in lower secondary school. Accordingly, the student solutions were dichotomously coded in terms of the extent to which such structuring elements, e.g., given/searched or question-calculation-answer, were part of the notes.	Example for *Petrol Station* task in **Figure 3** (2)
Mathematical units: A partial-credit coding was used to record the extent to which units were noted for all (score 1), some (score 0.5), or none (score 0) of the quantities.*	For *Oven: 500* ***€****, 100* ***units****, & 0.25* ***€*** For *Petrol Station: 105 K inhabitants & 30 km*
Marking relevant data: To record the marking-reading strategy (including highlighting and underlining), partial-credit coding was implemented to determine the extent to which none (score 0), one (score 0.33), two (score 0.66), or three (score 1) of the required quantities were marked.*	Example for *Petrol Station*: **Figure 3** (1)
Marking irrelevant data: To investigate the extent to which focusing on irrelevant data negatively influences the solution process, partial-credit coding was used to determine the extent to which none (score 0), one (score 0.5), or both (score 1) of the irrelevant numerical data were marked.*	For *Oven: 25 years & 15 L* For *Petrol Station: 105 K inhabitants & 30 km*
Noting relevant data: As in the case of the marking strategy, the data copied by the students based on the task text were analyzed. In terms of content, the extent to which the students wrote down all the numerical data required to answer the mathematical problem was evaluated. Exception 1: If an irrelevant number was noted, code 1 was assigned (score 2). Exception 2: If the profit was calculated in units (the 100 units less electricity per year are not needed for this), a score of 3 was assigned. If one relevant number was noted, a score of 4 was assigned. If two relevant numbers were noted, a score of 5 was assigned. If one or more irrelevant numbers were noted in addition to one or more relevant numbers, a score of 6 was assigned. If only one or more irrelevant numbers were noted, a score of 7 was assigned. If the task was not processed, a score of 9 was assigned. Only the notes were considered; the calculation pathways were ignored.*	For *Oven: 500 €, 100 units, & 0.25 €* Example for *Petrol Station*: **Figure 3** (3)
Noting irrelevant data: This strategy was also used to analyze the extent to which the students did not note down any (score 0), one (score 0.5), or both (score 1) of the numerical data for the tasks that were not relevant for the successful completion of the question.*	For *Oven: 25 years & 15 L* For *Petrol Station: 105 K inhabitants & 30 km*
Relationships: This code recognizes relationships within the notation before mathematization during students' notes. Score 0 corresponds to case in which the relationship was not recorded; score 1 corresponds to the case in which the relationship was captured completely correctly; score 2 corresponds to case in which the relationship was captured but included mathematical/content errors.**	For *Oven: 1 unit costs 0,25 €* For *Petrol Station: 1 L "arrow" 1,20 €.*
Order of notation: The score 0 corresponds to the case in which the order given in the task text was retained in the notes. This case did not contain any intentional restructuring of information. The score 1 corresponds to the case in which the order given in the task text is not preserved in the notes; instead, the information is presented in an order suitable for the computational steps.	Example for *Petrol Station*: **Figure 3** (4)
Position: The score 0 corresponds to the case in which no textual notes are included before mathematization, the score 1 corresponds to the case in which textual notes are observed before mathematization (**Figure 3** [5]).**	Example for *Petrol Station*: **Figure 3** (1-4)

Note. *For codes *mathematical units*, *marking (ir)relevant data*, & *noting relevant data,* we decided to introduce a *partially correct* code. Underlying idea was to capture tendencies & not have dichotomous coding for these categories. This approach enabled us to capture tendencies in results & draw conclusions for subsequent research & **Score 8 corresponds to the case involving no mathematization, score 9 corresponds to the case in which no notation was observed prior to mathematical text or the case in which the task was not attempted.

However, *Oven* task requires the design and solution of a single linear equation (see Eq. 1), whereas the *Petrol Station* task requires the design and comparison of two linear equations (see Eq. 1 and Eq. 2). Thus, the formation of a correct mathematical model and its mathematical processing is more demanding in the case of the second task.

$$0.25\ €/u\ *\ 100\ u/y\ *\ xy\ =\ 500\ € \tag{1}$$

€-Euro / u-unit / y-year

$$1.50\ €/1\ *\ x/>\ 1.20\ €/|\ *x/+6\ € \tag{2}$$

€-Euro / l-liters

Both tasks were assessed to have been correctly solved if the students wrote down 20 years and 20, as the solutions, respectively, and either included an explanatory sentence or explained the calculation method.

After completing the respective tasks, the question "How interesting did you find the task?" was posed with responses recorded on a five-point Likert scale (1: not at all interesting, 2: rather uninteresting, 3: neutral, 4: rather interesting, and 5: interesting) to assess the interest in the task.

To answer the aforementioned research questions, the students' solutions were not only evaluated in terms of correctness. Based on the theoretical explanations of the possible influences of different note elements during the solution of reality-based tasks, all student solutions were coded in terms of the occurrence of notes and their specific elements. In **Table 1**, we describe the codes and their derivation.

In addition, two meta-codes were used. As mentioned in the theory section, elaboration as a type of strategy significantly impacts learning and comprehension. This code was used to evaluate students who did not simply write the (relevant) numerical data of the task text (that is, their notes served as a mere relief for working memory [score 0]) but included a specific elaboration of the information with the notation (score 1). The relevant subcodes of elaboration were relationships within the notation (before and within mathematization), the order of the notes, and the position of use of the notes. If one of these codes was coded as true (score 1), the meta-code elaboration was also coded as 1.

The second meta-code summarizes dichotomously the occurrence of any note-taking. The two categories "notes are present" and "notes are not present" correspond to the presence and absence of the variants or possibly learning-promoting aspects of notes described in the theory section. Accordingly, excerpts of relevant (numerical) information, marking, literal quotations, summarizing paraphrases, or structuring representations were coded positively. The calculation methods or mere notation of the (correct) result were not evaluated as notes. For a deeper understanding of how notes were coded in this study, the meta-analysis discussed by Graham et al. (2020) is briefly presented. They did not consider the act of writing down figures to solve a mathematical problem, which is a classical arithmetic method, as a written learning activity. However, examples of written learning activities included writing down an explanation or additions to the task. We also decided to code only notes adjacent to the calculation. Thus, if only the calculation to obtain the correct solution was included, it was not considered as notes (score 0). Distinctions were made between various forms of notes, such as notations, sketches, excerpts, tables. However, because the analyses did not reveal any relevant differences, they were merged into a common code, capturing presence or absence of any note, irrespective of form.

For each aforementioned code, 10% of the student responses were independently assessed by two coders to ensure reliability. The Cohen's Kappa average inter-rater reliabilities of 0.87 for *Oven* task and 0.91 for the *Petrol Station* task indicate that the aspects of student responses were reliably assessed. In addition, the absence of systematic deviations at the level of the individual codes corroborate this conclusion.

Part 3–Questionnaire (Student Characteristics)

After the subject test, the students were assigned 20 min to complete a questionnaire containing various questions on different variables. The following four data points are relevant to the current study, in addition to the age of the respondents:

Gender

Until recently, the use of binary options was common in lower grades–male or female. Approximately 2% of the respondents did not respond to this question, which is slightly less than the proportion of people with non-binary gender identity in Germany.

School grade mathematics

To record the prior mathematical knowledge, the mathematics grade obtained on the last school report was requested; the grade scale comprises six options, from 1 (very good) to 6 (insufficient).

School grade German

While the C-test specifically assesses written linguistic competence, German grade on the last school report was also requested to assess general linguistic/literary competence. The grading scale comprises six options, from 1 (very good) to 6 (insufficient).

Socio-economic Status

To assess the SES, which has been shown in several studies to be a relevant factor influencing academic performance, the tried-and-tested question about the number of books at home ("How many books do you have at home?") was used. In addition, the students were given the hint "You can fit about 40 books on a bookshelf of one meter in length. Please do not count magazines, newspapers, or your schoolbooks". The answer categories were 0-10 books, 11-25 books, 26-100 books, 101-200 books, 201-500 books, and more than 500 books.

The data were analyzed using the statistical software SPSS version 28.0.0.0 (190). For research question 1, we calculated descriptive statistics (frequencies) regarding the characteristics of notes in both tasks as well as related effect sizes of the differences. To answer research question 2, we observed task-specific correlations between various note-taking characteristics and the solution success rate. For research questions 3 and 4, regression models were calculated by considering personal and note-taking characteristics as independent variables and the note-taking score and test score (referring to the correct task solution) as dependent variables. For all analyses, a significance level of $\alpha=0.05$ was used.

RESULTS

The results corresponding to each task are presented separately, because, contrary to our expectations, they exhibited limited generalizability. Despite this limitation, two particular similarities were observed between the two tasks.

The two tasks corresponded to identical levels of difficulty (38.33% for *Oven* and 38.56% for *Petrol Station*). Furthermore, the results revealed that both tasks (interest for *Oven* M=2.46 [SD=1.17] and interest for *Petrol Station* M=2.57 [SD=1.15]) were rated by the students as neither particularly interesting nor particularly uninteresting.

Characteristics of Note-Taking

The aim of the first research question was to identify quantitative and qualitative characteristics of note-taking in students while solving the two reality-based mathematical tasks and analyze any differences. **Table 2** presents the observed note-taking characteristics of different students corresponding to the two tasks.

Table 2. Percentages of different note-taking characteristics

Characteristics of note-taking	Task 1: *Oven*	Task 2: *Petrol Station*	t-test ANOVA significance	Effect size
	%	%	p	d*
Notes (structuring elements)	2.21	1.74	ns	0.034
Notes (some mathematical units)	36.1	39.70	<.05	-0.352
Notes (all mathematical units)	6.40	16.60		
Notes (marking some relevant data)	4.20	3.90	<.05	0.160
Notes (marking all relevant data)	15.70	10.20		
Notes (marking some irrelevant data)	9.80	7.20	<.05	0.198
Notes (marking all irrelevant data)	5.70	2.40		
Notes (noting some relevant data)	17.70	10.70	ns	-0.091
Notes (noting all relevant data)	6.40	13.30		
Notes (noting some irrelevant data)	3.90	6.10	ns	-0.024
Notes (noting all irrelevant data)	1.00	0.20		
Note-taking (anything)	65.60	71.68	<0.05	-0.131
Elaboration of data	48.89	59.48	<0.05	-0.212

Note. *Cohen's effect size d is a measure of the standardized mean difference between two groups: small effect corresponds to d=0.20, medium effect corresponds to d=0.50, & strong effect corresponds to d=0.80 (Cohen, 1969)

Overall, a high percentage of the students (task 1: 66%, task 2: 72%) took notes of some type. Most of them (task 1: 49%, task 2: 60%) took notes that served as an elaboration of relevant data. Results corresponding to both tasks revealed that many students used the strategy of note-taking while solving reality-based tasks. At least a third of the students wrote down mathematical units. Moreover, the strategy of marking data was employed quite often. Several students not only marked the relevant data but also the irrelevant data presented within the tasks (task 1: 16% (all relevant) vs. 6% (all irrelevant); task 2: 10% (all relevant) vs. 2% (all irrelevant)). Several students also explicitly noted down data during the process of solution. In contrast to the marking strategy, explicit noting of data was mostly done for relevant data than for irrelevant data (task 1: 18% (some relevant) vs. 4% (some irrelevant); task 2: 11% (some relevant) vs. 6% (some irrelevant)). Approximately two percent of the students structured the problem-solving process by writing down the structuring elements (both tasks).

A comparison between the different characteristics of note-taking corresponding to the two tasks revealed that note-taking strategies (d=-0.13) and elaboration of data (d=-0.21) were used more frequently in task 2 than in task 1. In contrast, students underlined relevant (d=0.16) or irrelevant data (d=0.20) more often in task 1 than in task 2. The largest effect size was recorded for taking notes on units (d=0.352). Thus, in summary, although the two tasks were mathematically very similar, the corresponding characteristics of note-taking differed significantly.

Different Types of Notes and Students' Mathematical Performance

The second research question refers to the relationship between the frequency of taking different types of notes and correctly solving the mathematical problem. **Table 3** presents the correlation between different types of note-taking and the test scores for tasks 1 and 2. The results revealed that, for both reality-based tasks, the pattern and order of statistically significant correlation coefficients were identical. The most relevant correlations were observed corresponding to the "elaboration of data" (task 1: r=.40; task 2: r=.20), "note-taking" (anything) (task 1: r=.21; task 2: r=.13), and "noting mathematical units" (task 1: r=.35; task 2: r=.20). In the case of both tasks, the "marking of irrelevant data" (task 1: r=-.14; task 2: r=-.11) and "note-taking of irrelevant data" (task 1: r=-.10; task 2: r=-.17) were negatively correlated with the test scores.

Table 3. Correlations between different types of note-taking & test scores

Different types of note-taking	Task 1: *Oven*	Task 2: *Petrol Station*
	r	*R*
Notes (structuring elements)	-.015	-.003
Notes (mathematical units)	.351*	.198*
Notes (marking relevant data)	-.053	-.083
Notes (marking irrelevant data)	-.143*	-.105*
Notes (noting relevant data)	-.031	.049
Notes (noting irrelevant data)	-.096*	-.167*
Note-taking (anything)	.209*	.130*
Elaboration of data	.402*	.198*

Note. Statistically significant coefficients are characterized with an asterisk (* for p<.05)

Student Characteristics and Note-Taking

The third research question considers the extent to which student characteristics determine the practice of note-taking and the extent of elaboration of data during the solution of reality-based tasks. This correlation is presented in **Table 4**.

Table 4. Dependence of note-taking (anything) on student characteristics

Student characteristics	Task 1: *Oven*		Task 2: *Petrol Station*	
	Correlation	Regression	Correlation	Regression
	r	Beta	r	Beta
Gender (0=female, 1=male)	-.13*	-.12	-.09	.06
School grade mathematics	.02	.02	.04	.12*
School grade German	.48*	.12	-.09	.02
Linguistic competence (C-test)	.25*	.25*	.27*	.26*
Student's interest (task)	.13*	.16*	.08	.07
Socio-economic status (SES)	.12*	-.01	.11*	.03
Explained variance, R^2 [in %]		7.6		5.6

Note. Statistically significant coefficients are characterized with an asterisk (* for p<.05)

Some differences in the types of note-taking were observed between male and female students. For example, female students took significantly more notes than male students in task 1. Further, the school grade in mathematics was observed to be predictive only for the performance corresponding to task 2. Conversely, the German school grades were correlated with note-taking for task 1 but not for task 2. However, after controlling the other predictors, the German school grades were not predictive for note-taking corresponding to either task. Therefore, we concluded that the student characteristics, gender and school grades, were not relevant for any kind of note-taking. In contrast, both linguistic competence and interest regarding task context were observed to be predictive for note-taking. Thus, students with higher linguistic competence and higher interest in the task context were observed to take more notes corresponding to at least one of the tasks compared to students with lower linguistic competence and lower interest in the task context. Finally, the correlations between note-taking and SES are also evident from the table; higher SES is observed to be correlated to a higher probability of note-taking.

Concerning the elaboration of data, specific effects were observed corresponding to linguistic competence and students' interest in task contexts. However, with respect to gender and school grades in mathematics and German, the results were inconsistent. Gender was observed to be correlated to elaboration of data corresponding to task 2 but not task 1 (**Table 5**). Small but statistically significant, differences corresponding to gender were observed repeatedly, which will be further elucidated in the Discussion section. Significant correlations were observed corresponding to the German school grade for task 2 but not for task 1.

Table 5. Dependence of elaboration of data on student characteristics

Student characteristics	Task 1: *Oven*		Task 2: *Petrol Station*	
	Correlation	Regression	Correlation	Regression
	r	Beta	r	Beta
Gender (0=female, 1=male)	.01	-.02	-.10*	.03
School grade mathematics	.02	.04	-.05	.05
School grade German	.04	.10	-.14*	.01
Linguistic competence (C-test)	.22*	.24*	.26*	.24*
Student's interest (task)	.26*	.28*	.15*	.15*
Socio-economic status (SES)	.15*	.00	.16*	.06
Explained variance, R^2 [in %]		12.0		7.8

Note. Statistically significant coefficients are characterized with an asterisk (* for p<.05)

Concerning the third research question, stable effects were observed only corresponding to linguistic competence and interest in the task context.

Elaboration of Data as a Predictor for Solving Reality-Based Tasks

Finally, the fourth research question considers the extent to which note-taking and the elaboration of data determine the solution success rate in the two reality-based tasks (corresponding to controlled values of student characteristics). The results of a bivariate correlation and two ordinary least squares (OLS) regression analyses for each task are presented in **Table 6**.

Table 6. Predicted success rates for solving two reality-based tasks on student characteristics & note-taking prevalence

Student characteristics	Task 1: Oven			Task 2: Petrol Station		
	Correlation	RM 1	RM 2	Correlation	RM 1	RM 2
	r	Beta	Beta	r	Beta	Beta
Gender (0=female, 1=male)	.240*	.279*	.27*	-.001	.08	.08
School grade mathematics	-.156*	-.059	-.07	-.205*	-.10	-.10
School grade German	-.074	.016	-.01	-.161*	.05	.05
Linguistic competence (C-test)	.335*	.294*	.23*	.347*	.33*	.32*
Students interest (task)	.279*	.254*	.17*	.317*	.27*	.25*
Socio-economic status (SES)	.240*	.048	.05	.212*	.05	.05
Note-taking (anything)	.209*		-.11	.130*		-.07
Elaboration of data	.402*		.36*	.198*		.11
Explained variance, R^2 [in %]		26.8	34.2		20.9	20.9

Note. Statistically significant coefficients are characterized with an asterisk (* for p<.05) & RM: Regression model

In the case of task 1, note-taking and elaboration of data were observed to be strong predictors for correctly solving the task. Both variables exhibited significant correlation with the task score. A comparison of the explained variance (R^2) between regression models 1 and 2 corresponding to task 1 revealed that including the predictors note-taking and elaboration of data into the analysis increased the amount of explained variance significantly (R^2=26.8% vs. R^2=34.2%). Moreover, both regression coefficients provided specific and statistically significant predictions of correctly solving task 1 corresponding to controlled student characteristics. In the case of task 2, both note-taking and elaboration of data were observed to be positively correlated with the solution's success. However, the correlations were significantly lower than those corresponding to task 1. Furthermore, in this case, accounting for note-taking and elaboration of data did not increase the proportion of explainable variance between the regression models 1 and 2 (R^2=20.9% vs. R^2=20.9%). Thus, in this case, the effects of note-taking and elaboration of data were statistically insignificant corresponding to controlled linguistic competencies and interest in task context. Interestingly, gender was a significant predictor only corresponding to model 1 and task 1. In summary, students who took more notes and elaborated more data were observed to be more successful in solving task 1. Both parts of strategy of *note-taking* (general and elaboration of data) were important for solving the reality-based mathematical task.

DISCUSSION

Scientific Significance of the Study

These analyses reveal, among other things, that higher language skills (r=.26), higher interest in the task (r=.13), and higher SES (r=.12) are significantly associated with greater note-taking frequency. In addition, taking notes in general (r=.17), taking notes on units (r=.27) and elaboration of data (r=.30) are significantly correlated to the solution success rate of the modeling problems, in contrast to merely writing down the (relevant) numerical data. Successful task completion can be explained by 34,2 % of the variance (for task 1), particularly by elaborating the information from the task text. Students who processed the data further, rearrange them in their own order, or include relationships are significantly more successful than those who merely write down the data given in the task or do not take notes. Despite the parallelly constructed tasks with different contexts, the predictors were found to differ significantly (e.g., elaboration of data). Although more notes were taken in task 2 than in task 1, the meta-code elaboration of data became significant only in task 1.

Based on the literature, the comprehension phase is known to be important to the solution of reality-based mathematical problems. Errors in this phase (e.g., in deciding computational pathways and interpretations) affect subsequent decisions significantly. Learners often experience difficulty in constructing an adequate situation model following the comprehension phase. This observation indicates that working on reality-based tasks requires special linguistic and intra-mathematical competencies because linguistically and situationally complex problems are difficult to understand and solve. We assume that notes could be a way to overcome these difficulties. In this study, the role of notes in effective situational comprehension is considered. Leiss et al. (2019) proved that the mere presence of notes exerts a direct influence on the creation of an adequate situation model. Similarly, based on the categorization of notes presented in the theory section, we can see that different forms of notes have different effects on comprehension and model creation. This self-constructed model for ranking notes on a scale of elaboration also reflects the hierarchy of notes in practice, as shown for one task (**Table 3** and **Table 6**). *Elaboration of data* was identified to be a salient predictor of note-taking. This code was obtained based on the representation of relationships between certain quantities, the introduction of a custom order within the notes, and the location of note usage. Accounting for this quantity increased the explained variance from 27% to 34% corresponding to one of the tasks. At this point, the difference between the two tasks is again interesting. While in task 1, data elaboration is highly correlated (r=.4) and contributes a significant portion to the explained variance (β=.36), this relationship changes in task 2 (for *Petrol Station*), where the elaboration is still correlated (r=.20) but drops out in the regression.

In addition, linguistic competence was identified to be another significant predictor, as mentioned in the theory section. Interest in the task was already known to be a significant predictor (Gijsbers et al., 2020), which is corroborated by this study (β=.21). Other important correlations are found in the cases of school grades, SES, overall note-taking, and gender (**Table 3** and **Table 6**).

However, most of these influencing factors have not been researched extensively. Nevertheless, such studies as relevant considering that the literature and the state of research indicate that notes may be an option to overcome existing difficulties. The effects of qualitative and quantitative characteristics and their predictors comprise relevant questions in this context. This study investigates the extent to which personal and task-related characteristics determine specific note-taking characteristics. It also analyzes the connection between these elements and successful task processing. It shows that note-taking exhibits several characteristics rather than just being a form of externalized memory. In addition, it suggests that mathematical education in schools should use note-taking as a subject-specific learning strategy and, therefore, integrate this topic as an explicit learning element into the curriculum. In particular, the elaboration and further processing of numerical data presented in the text is observed to be a special predictor.

Implications for Practical Application

The results indicate that schools should use note-taking specifically as a learning and elaboration strategy. Students should be trained to leverage the benefits highlighted in this study–particularly that of the elaboration of relevant information given in the text. In addition, instructing students to consider which numerical data are relevant to solving the task because a correlation is observed between successful solution and marking and writing down irrelevant numerical data. In the long term, the implementation of the strategies could support the creation of more mathematically mature citizens.

According to Graham et al. (2020), when using writing-based learning methods in the classroom, such as note-taking, teachers should consider the purpose of the approach. We recommend that teachers always focus on promoting content comprehension rather than mere writing skills. Similarly, the predictors reported in Graham et al. (2020) also hold true for the present study. For example, extremely different results were obtained for the predictor gender in different cases. It was observed to be correlated to note-taking corresponding to task 1 (r=-.13*; girls tend to make more notes); however, this correlation disappears when linguistic competence and interest in task context are also considered. In the case of the second task, no similar correlation was observed. The elaboration of text was observed to be negatively correlated to note-taking in the case of task 2 (r=-.10*; girls are more inclined to elaborate on their notes) but was not significant during regression. Gender was strongly correlated to solution success (r=.24) for task 1 and observed to be a strong predictor in both models. Looking only at the correlations, it is relevant to note that while girls take more elaborate notes for tasks 1 and 2 (see **Table 4** and **Table 5**), boys' mathematical performance is better in task 1 (see **Table 6**). Regarding mathematical performance in task 2, no clear influence of gender was observed. These findings suggest that further research is needed regarding gender and its influence on solving reality-based tasks.

Another important finding is that noting units is associated with better mathematical performance (r=.27) but only a quarter of the students actually write down the units (**Table 2**). In this regard, teachers should educate students on the importance of noting units as it relates to better performance in mathematics. However, as the directionality of this correlation remains uncertain, future studies can explore this relationship further. Likewise, students should be taught to mark more selectively, as marking irrelevant information correlates negatively with mathematical performance (r=-.124). Teachers can emphasize the importance of paying attention to the information needed to solve the problem and ignoring irrelevant data. Similarly, writing out irrelevant data is associated with lower mathematical performance (r=-.132).

Thus, various predictors exhibited varying effects and a certain number of predictors remain unidentified. In addition, in the performance scenario considered in the study, a normal classroom situation was only partially replicated. The different effects of notes are particularly apparent in the difference corresponding to two parallelly constructed tasks.

As shown in **Table 6**, elaboration seems to be a (strong) predictor for only one of the two tasks, although the frequency distribution (**Table 2**) suggests otherwise. The two tasks considered in this study were similar in terms of their central mathematical idea (functional relationships), the mathematical competencies required (modelling and technical work), and the basic problem content (a financial decision-making task). Furthermore, the amount of relevant and irrelevant numerical data and the complexity of the required units (monetary amounts presented as decimal numbers and in terms of euros per unit or euros per liter), as well as the degrees of difficulty and the interest in the tasks were similar. Differences were observed primarily in the contexts of the tasks. For example, students are likely to have limited experience regarding the extent to which an oven can help to reduce electricity costs. Conversely, the situation of refueling a car at various inexpensive petrol stations is more commonly understood, at least passively. From this perspective, task 1 could be considered to be more demanding than task 2. Nevertheless, more students were observed to take notes corresponding to task 2. They also noted down all units more frequently in this case. Similarly, highly significant differences are observed in variance resolution between the two tasks, as evidenced by the results in Table 6. This disparity suggests the presence of other relevant predictors that have not been considered. Hence, further research is required to identify these hidden predictors. Regarding non-contextual differences, the problem is stated concretely in task 1, whereas it is only stated implicitly in task 2. Moreover, task 1 contains an abstract unit ("unit"), whereas all actual units are specified in task 2. The distractors in the text (15 l/30 km) may also exert different influences.

CONCLUSIONS

The effectiveness of notes in the case of different tasks and their characteristic properties merit further investigation. In particular, an intensive and extended training study would be appropriate for this purpose.

The predictors discussed in the previous section should be maintained as constants in subsequent studies. Similarly, follow-up studies should be conducted to distinguish between deliberate prompts for note-taking, as present here, and voluntary or intuitive note-taking. An implied expectation that notes will be taken would help us to study actual responses in comparison to elicited ones.

The number of items should also be increased significantly. As this survey was conducted as part of an interdisciplinary study, only two distinct tasks could be considered. Furthermore, the study indicates that task characteristics are a predictor for note-taking. Thus, more tasks should be considered in subsequent studies.

In addition, the cases in which units are noted should be identified. The calculation pathways or errors therein could also be analyzed in greater detail based on more precise coding and in the context of recorded notes. Overall, the extent to which the calculation path is part of notes should be reflected in the study because tasks can also be solved using a calculator. The role of noting down irrelevant data in the derivation of incorrect solutions is another interesting observation.

If numerous other predictors are found and the variance is further elucidated, a training study with a comparison group could be performed to corroborate the conclusion that note-taking is conducive to learning, which has already been asserted multiple times in the literature (Graham et al., 2020; Kiewra, 1989; Spires & Stones, 1989). In general, Chang and Ku (2015) also demonstrated that using a note-taking strategy significantly improves students' note-taking and reading comprehension performance. Demonstrating this result for reality-based tasks can significantly help in minimizing the difficulties faced by students while solving them.

Despite parallelly constructed tasks, no consistent effect of note-taking could be demonstrated in this study. A more in-depth elaboration of the required information turned out to be particularly effective. The elaboration of the information needed to solve the task must be clearly distinguished from the mere writing out of the required numerical data. The meta-code of elaboration is shown to be a strong predictor of (β=.36) for the successful solution process for one of the two tasks. Despite eliciting a note-taking-related predictor of successful solving of reality-based tasks, no clear conclusions can be drawn about the quality of note-taking. Therefore, teachers should be aware of the possible influence of notes when dealing with reality-based tasks in math classes. They should deal with the topic both from a technical as well as a didactical perspective to adapt it for their lessons.

Author notes: This article was written in the context of extensive research work by the group of educational scientists subject and language didacticians and educational researchers. Working Group on Subject and Language (AG FuS). Since 2011, this working group has been since 2011, has been researching various dimensions of subject-related tasks, with the linguistic dimension plays a prominent role. This explains why in this article focuses on text-heavy tasks (rather than, say, informational texts). For more details, see http://ww w.fach-und-sprache.de/.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: Authors stated that an ethical statement and presentation to such a committee is not required in Germany. If the study has been approved by the state education authorities, it can be conducted. After the official approval of the study, the consent of the students and their parents was obtained. The data were then processed only by persons who signed a confidentiality agreement.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Atkinson, R. K., Levin, J. R., Kiewra, K. A., Meyers, T., Kim, S.-I., Atkinson, L. A., Renandya, W. A., & Hwang, Y. (1999). Matrix and mnemonic text-processing adjuncts: Comparing and combining their components. *Journal of Educational Psychology, 91*(2), 342-357. https://doi.org/10.1037/0022-0663.91.2.342

Baghaei, P., & Grotjahn, R. (2014). Establishing the construct validity of conversational C-Tests using a multidimensional Rasch model. *Psychological Test and Assessment Modeling, 56*(1), 60-82.

Blomhøj, M., & Jensen, T. H. (2007). What's all the fuss about competencies? Experiences with using a competence perspective on mathematics education to develop the teaching of mathematical modelling. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education* (pp. 45-56). Springer. https://doi.org/10.1007/978-0-387-29822-1_3

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling: Education, engineering and economics* (pp. 222-231). Horwood Publishing. https://doi.org/10.1533/9780857099419.5.221

Blum, W., Galbraith, P. L., Henn, H.-W., & Niss, M. (Eds.) (2007). *Modelling and applications in mathematics education*. Springer. https://doi.org/10.1007/978-0-387-29822-1

Boaler, J. (2022). *Mathematical mindsets: Unleashing students' potential through creative mathematics, inspiring messages and innovative teaching*. John Wiley & Sons.

Boch, F., & Piolat, A. (2005). Note taking and learning: A summary of research. *The WAC Journal, 16*(1), 101-113. https://doi.org/10.37514/WAC-J.2005.16.1.08

Boonen, A. J. H., van Wesel, F., Jolles, J., & van der Schoot, M. (2014). The role of visual representation type, spatial ability, and reading comprehension in word problem solving: An item-level analysis in elementary school children. *International Journal of Educational Research, 68*, 15-26. https://doi.org/10.1016/j.ijer.2014.08.001

Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of the phases in the modelling process. *ZDM, 38*(2), 1-8. https://doi.org/10.1007/BF02655883

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: Comparing task-specific self-reports to traces. *Metacognition and Learning, 2*(1), 1-20. https://doi.org/10.1007/s11409-007-9004-y

Brinkmann, A. (2003). Mind mapping as a tool in mathematics education. *The Mathematics Teacher, 96*(2), 96-101. https://doi.org/10.5951/MT.96.2.0096

Brinkmann, A. (2005). Knowledge maps: Tools for building structure in mathematics. *International Journal for Mathematics Teaching and Learning, 1473*, 111.

Brown, J. P. (2019). Real-world task context: Meanings and roles. In G. A. Stillman, & J. P. Brown (Eds.), *Lines of inquiry in mathematical modelling research in education* (pp. 53-81). Springer. https://doi.org/10.1007/978-3-030-14931-4_4

Buschmeier, G. (2017). *Denken und Rechnen 2* [*Schülerband*] *(*[*Grundschule, Bremen, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Schleswig-Holstein*]*, Druck A)* [*Thinking and arithmetic 2nd* [*student volume*] *(*[*Elementary school, Bremen, Hamburg, Hesse, Lower Saxony, North Rhine-Westphalia, Rhineland-Palatinate, Saarland, Schleswig-Holstein*]*, Print A)*]. Westermann.

Buzan, T. (1983). *Use both sides of your brain*. EP Dutton.

Chang, W., & Ku, Y. (2015). The effects of note-taking skills instruction on elementary students' reading. *The Journal of Educational Research, 108*(4), 278-291. https://doi.org/10.1080/00220671.2014.886175

Chen, P.-H. (2021). In-class and after-class lecture note-taking strategies. *Active Learning in Higher Education, 22*(3), 245-260. https://doi.org/10.1177/1469787419893490

Clarkson, P. C. (1991). Language comprehension errors: A further investigation. *Mathematics Education Research Journal, 2*(2), 24-33. https://doi.org/10.1007/BF03217225

Cohen, R. A. (1969). Conceptual styles, culture conflict, and nonverbal tests of intelligence. *American Anthropologist, 71*(5), 828-856. https://doi.org/10.1525/aa.1969.71.5.02a00040

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9*(4), 343-363. https://doi.org/10.1016/S0959-4752(98)00051-6

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*(4), 405-438. https://doi.org/10.1016/0010-0285(88)90011-4

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-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology, 83*(1), 61-68. https://doi.org/10.1037/0022-0663.83.1.61

Depaepe, F., De Corte, E., & Verschaffel, L. (2010). Teachers' approaches towards word problem solving: Elaborating or restricting the problem context. *Teaching and Teacher Education, 26*(2), 152-160. https://doi.org/10.1016/j.tate.2009.03.016

Dryvold, A., Berqvist, E., & Österhom, M. (2015). Uncommon vocabulary in mathematical tasks in relation to demand of reading ability and solution frequency. *Nordisk Matematikdidaktikk, 20*(1), 5-31.

Entrekin, V. S. (1992). Mathematical mind mapping. *The Mathematics Teacher, 85*(6), 444-445. https://doi.org/10.5951/MT.85.6.0444

Frejd, P. (2013). Modes of modelling assessment–A literature review. *Educational Studies in Mathematics, 84*(3), 413-438. https://doi.org/10.1007/s10649-013-9491-5

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

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*(1), 30-47. https://doi.org/10.1037/0022-0663.100.1.30

Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., & Martin, B. (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, 22*(2), 152-166. https://doi.org/10.1080/10888438.2017.1398259

Galbraith, P. L., & Stillman, G. A. (2006). A framework for identifying student blockages during transitions in the modelling process. *ZDM, 38*(2), 143-162. https://doi.org/10.1007/BF02655886

Gijsbers, D., de Putter-Smits, L., & Pepin, B. (2020). Changing students' beliefs about the relevance of mathematics in an advanced secondary mathematics class. *International Journal of Mathematical Education in Science and Technology, 51*(1), 87-102. https://doi.org/10.1080/0020739X.2019.1682698

Gillespie, A., Graham, S., Kiuhara, S., & Hebert, M. (2014). High school teachers' use of writing to support students' learning: A national survey. *Reading and Writing: An Interdisciplinary Journal, 27*(6), 1043-1072. https://doi.org/10.1007/s11145-013-9494-8

Graham, S., Kiuhara, S. A., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research, 90*(2), 179-226. https://doi.org/10.3102/0034654320914744

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77*(3), 334-372. https://doi.org/10.3102/003465430303953

Grotjahn, R., & Drackert, A. (2020). *The electronic C-test bibliography*. http://www.c-test.de/deutsch/ctest/pdf/C%20Test%20Bibliography/Grotjahn_Drackert_Electronic_Ctest_Bibliography_10_2020.pdf

Grotjahn, R., Klein-Barley, C., & Raatz, U. (2002). C-test: An overview. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-test* (pp. 93-114). AKS-Verlag.

Hadwin, A. F., & Winne, P. H. (1996). Study strategies have meager support: A review with recommendations for implementation. *The Journal of Higher Education, 67*(6), 692-715. https://doi.org/10.1080/00221546.1996.11774821

Hagen, Å. M., Braasch, J. L. G., & Bråten, I. (2014). Relationships between spontaneous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. *Journal of Research in Reading, 37*(S1), S141-S157. https://doi.org/10.1111/j.1467-9817.2012.01536.x

Harsch, C., & Hartig, J. (2010). Empirische und inhaltliche analyse lokaler Abhängigkeiten im C-test [Empirical and content analysis of local dependencies in the C-test]. *Der C-Test: Beiträge aus der Aktuellen Forschung* [*The C-Test: Contributions from Current Research*], 193-204.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4), 684-689. https://doi.org/10.1037/0022-0663.91.4.684

Jamieson-Noel, D., & Winne, P. H. (2003). Comparing self-reports to traces of studying behavior as representations of students' studying and achievement. *Zeitschrift für Pädagogische Psychologie* [*Journal of Educational Psychology*], *17*(3/4), 159-172. https://doi.org/10.1024//1010-0652.17.34.159

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

Kaiser, G., & Schukajlow, S. (2022). Innovative perspectives in research in mathematical modelling education. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proceedings of the 45th Conference of the International Group for Psychology of Mathematics Education* (pp. 147-176). PME.

Kaiser, G., Blum, W., Stillman, G., & Borromeo Ferri, R. (Eds.). (2011). *Trends in teaching and learning of mathematical modelling*. Springer. https://doi.org/10.1007/978-94-007-0910-2

Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review, 1*(2), 147-172. https://doi.org/10.1007/BF01326640

Kiewra, K. A., Benton, S. L., Kim, S. I., Risch, N., & Christensen, M. (1995). Effects of note-taking format and study technique on recall and relational performance. *Contemporary Educational Psychology, 20*(2), 172-187. https://doi.org/10.1006/ceps.1995.1011

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology, 83*(2), 240-245. https://doi.org/10.1037/0022-0663.83.2.240

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal, 29*(2), 303-323. https://doi.org/10.2307/1163370

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394. https://doi.org/10.1037/0033-295X.85.5.363

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*(2), 242-262. https://doi.org/10.1016/j.cedpsych.2004.10.001

Lahtinen, V., Lonka, K., & Lindblom-Ylänne, S. (1997). Spontaneous study strategies and the quality of knowledge construction. *British Journal of Educational Psychology, 67*(1), 13-24. https://doi.org/10.1111/j.2044-8279.1997.tb01223.x

Laidlaw, E. N., Skok, R. L., & McLaughlin, T. F. (1993). The effects of notetaking and self-questioning on quiz performance. *Science Education, 77*(1), 75-82. https://doi.org/10.1002/sce.3730770105

Lee, K., Ng, S.-F., Ng, E.-L., & Lim, Z.-Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology, 89*(2), 140-158. https://doi.org/10.1016/j.jecp.2004.07.001

Leenaars, F. A. J., van Joolingen, W. R., & Bollen, L. (2013). Using self-made drawings to support modelling in science education: Using self-made drawings to support modelling. *British Journal of Educational Technology, 44*(1), 82-94. https://doi.org/10.1111/j.1467-8535.2011.01272.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

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 Für Mathematik-Didaktik* [*Journal for Mathematics Didactics*]*, 31*(1), 119-141. https://doi.org/10.1007/s13138-010-0006-y

Lonka, K., Lindblom-YlÄnne, S., & Maury, S. (1994). The effect of study strategies on learning from text. *Learning and Instruction, 4*(3), 253-271. https://doi.org/10.1016/0959-4752(94)90026-4

Ludwig, M., & Reit, X. R. (2013). Comparative study about gender differences in mathematical modelling. In G. Nagarjuna, A. Jamakhadia, & E.M. Sam (Eds.), *Proceedings of EPISTEME 5* (pp. 48-54). Cinnamon Teal.

Malone, J., & Dekkers, J. (1984). The concept map as an aid to instruction in science and mathematics. *School Science and Mathematics, 84*(3), 220-231. https://doi.org/10.1111/j.1949-8594.1984.tb09543.x

Matos, J. F., & Carreira, S. (1995). Cognitive processes and representations involved in applied problem solving. In C. W. Sloyer, W. Blum, & I. D. Huntley (Eds.), *Advances and perspectives in teaching of mathematical modelling and applications* (pp. 71-80). Horwood Publishing.

Mayer, R. E., & Hegarty, M. (1996). The process of understanding mathematical problems. In R. J. Sternberg, & T. Bem-Zeev (Eds.), *The nature of mathematical thinking* (pp. 29-54). Lawrence Erlbaum.

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: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Niss, M. (2016). Mathematical competencies and PISA. In K. Stacey, & R. Turner (Eds.), *Assessing mathematical literacy. The PISA experience* (pp. 5-34). Springer.

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* (pp. 3-32). Springer. https://doi.org/10.1007/978-0-387-29822-1_1

Novac, J. D. (1996). Concept mapping: A tool for improving science teaching and learning. In D. F. Treagust, R. Duit, & B. J. Fraser (Eds.), *Improving teaching and learning in science and mathematics* (pp. 32-43). Teachers College Press.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*(10), 937-949. https://doi.org/10.1002/tea.3660271003

OECD. (2013). *PISA 2012 released mathematics items*. OECD Publishing.

OECD. (2019). *PISA 2018 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/b25efab8-en

Palm, T. (2002). *The realism of mathematical school tasks. Features and consequences* [Unpublished doctoral dissertation]. University of Umea.

Pape, S. J. (2004). Middle school children's problem-solving behavior: A cognitive analysis from a reading comprehension perspective. *Journal for Research in Mathematics Education, 35*(3), 187-219. https://doi.org/10.2307/30034912

Paredes, S., Cáceres, M. J., Diego-Mantecón, J. M., Blanco, T. F., & Chamoso, J. M. (2020). Creating realistic mathematics tasks involving authenticity, cognitive domains, and openness characteristics: A study with pre-service teachers. *Sustainability, 12*(22), 9656. https://doi.org/10.3390/su12229656

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement, 53*(3), 801-813. https://doi.org/10.1177/0013164493053003024

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*(3), 291-312. https://doi.org/10.1002/acp.1086

Plath, J. (2017). Das Anfertigen von Notizen bei der Bearbeitung von realitätsbezogenen Mathematikaufgaben [Taking notes while working on reality-based mathematics tasks]. In D. Leiss, M. Hagena, A. Neumann, & K. Schwippert (Eds.), *Mathematik und Sprache: Empirischer Forschungsstand und unterrichtliche Herausforderungen* [*Mathematics and language: Empirical research and teaching challenges*]. Waxmann.

Plath, J. (2020). Verstehensprozesse bei der Bearbeitung realitätsbezogener Mathematikaufgaben: Klassische Textaufgaben vs. Zeitungstexte [Comprehension processes during the processing of reality-based mathematics tasks: Classical text tasks vs. newspaper texts]. *Journal für Mathematik-Didaktik* [*Journal for Mathematics Didactics*], *41*(2), 237-266. https://doi.org/10.1007/s13138-019-00148-w

Plath, J., & Leiss, D. (2018). The impact of linguistic complexity on the solution of mathematical modelling tasks. *ZDM, 50*(1-2), 159-171. https://doi.org/10.1007/s11858-017-0897-x

Pollak, H. O. (1979). The interaction between mathematics and other school subjects. In UNESCO (Ed.), *New trends in mathematics teaching.* United Nations Educational, Scientific and Cultural Organization.

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, 52*(1), 33-44. https://doi.org/10.1007/s11858-019-01118-9

Ray, A. B., Graham, S., Houston, J. D., & Harris, K. R. (2016). Teachers' use of writing to support students' learning in middle school: A national survey in the United States. *Reading and Writing, 29*(5), 1039-1068. https://doi.org/10.1007/s11145-015-9602-z

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. (2020). 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

Rellensmann, J., Schukajlow, S., & Leopold, C. (2020). Measuring and investigating strategic knowledge about drawing to solve geometrical modelling problems. *ZDM, 52*(1), 97-110. https://doi.org/10.1007/s11858-019-01085-1

Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. *Instructional Science, 17*(4), 309-338. https://doi.org/10.1007/BF00056219

Rinkens, H.-D., & Dingemans, S. (Eds.). (2014). *Welt der Zahl: Mathematisches Unterrichtswerk für die Grundschule. 2* [*Schülerband*]*/erarb. von Steffen Dingemans* (*Für die Grundschule,* [*Nordrhein-Westfalen, Hessen, Rheinland-Pfalz, Saarland*], *Druck A*) [*World of numbers: Mathematical instruction book for elementary school. 2,* [*student volume*]*/edited by Steffen Dingemans* (*for elementary school,* [*North Rhine-Westphalia, Hesse, Rhineland-Palatinate, Saarland*], *print A*)]. Schroedel.

Rogiers, A., Merchie, E., & Van Keer, H. (2020). Opening the black box of students' text-learning processes: A process mining perspective. *Frontline Learning Research, 8*(3), 40-62. https://doi.org/10.14786/flr.v8i3.527

Rovers, S. F., Clarebout, G., Savelberg, H. H., de Bruin, A. B., & van Merriënboer, J. J. (2019). Granularity matters: comparing different ways of measuring self-regulated learning. *Metacognition and Learning, 14*, 1-19. https://doi.org/10.1007/s11409-019-09188-6

Samuelstuen, M. S., & Bråten, I. (2007). Examining the validity of self-reports on scales measuring students' strategic processing. *British Journal of Educational Psychology, 77*(2), 351-378. https://doi.org/10.1348/000709906X106147

Schnotz, W., & Dutke, S. (2004). Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen [Cognitive psychology of reading competence: Multilevel processing using multiple sources of information]. In U. Schiefle, C. Artelt, W. Schneider & P. Stanat (Eds.), *Struktur, Entwicklung und Förderung von Lesekompetenz* [*Structure, development and promotion of reading skills*](pp. 61-99). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-81031-1_4

Schukajlow, S., Kaiser, G., & Stillman, G. (2018). Empirical research on teaching and learning of mathematical modelling: A survey on the current state-of-the-art. *ZDM, 50*(1-2), 5-18. https://doi.org/10.1007/s11858-018-0933-5

Schukajlow, S., Kaiser, G., & Stillman, G. (2021). Modeling from a cognitive perspective: Theoretical considerations and empirical contributions. *Mathematical Thinking and Learning*, 1-11. https://doi.org/10.1080/10986065.2021.2012631

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest and self-efficacy expectations. *Educational Studies in Mathematics, 79*(2), 215-237. https://doi.org/10.1007/s10649-011-9341-2

Slotte, V., Lonka, K., & Lindblom-Ylänne, S. (2001). Study-strategy use in learning from text. Does gender make any difference? *Instructional Science, 29*(3), 255-272. https://doi.org/10.1023/A:1017574300304

Spires, H. A., & Stone, P. D. (1989). The directed notetaking activity: A self-questioning approach. *Journal of Reading, 33*(1), 36-39.

Staub, F. C. (2006). Notizenmachen: Funktionen, Formen und Werkzeugcharakter von Notizen [Note-taking: Functions, forms and tool character of notes]. In H. Mandl, & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* [*Learning strategies handbook*] (pp. 59-71). Hogrefe.

Stillman, G., Brown, J., & Galbraith, P. L. (2010). Identifying challenges within transition phases of mathematical modeling activities at year 9. In R. Lesh, P. L. Galbraith, C. R. Haines, & A. Hurford (Eds.), *Modeling students' mathematical modeling competencies* (pp. 385-398). Springer. https://doi.org/10.1007/978-1-4419-0561-1_33

Strohmaier, A. R. (2020). *When reading meets mathematics: Using eye movements to analyze complex word problem solving* [Doctoral dissertation, Technical University of Munich]. https://mediatum.ub.tum.de/?id=1521471

Trabasso, T., & Bouchard, E. (2002). Teaching readers how to comprehend text strategically. *Comprehension Instruction: Research-based Best Practices*, 176-200.

UNESCO. (2017). *Education for sustainable development goals–Learning objectives*. United Nations Educational, Scientific and Cultural Organization. https://doi.org/10.54675/CGBA9153

van den Heuvel-Panhuizen, M. (2003). The didactical use of models in realistic mathematics education: An example from a longitudinal trajectory on percentage. *Educational Studies in Mathematics, 54*(1), 9-35. https://doi.org/10.1023/B:EDUC.0000005212.03219.dc

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*. Swets & Zeitlinger Publishers.

Verschaffel, L., Schukajlow, S., Star, J., & van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM, 52*(1), 1-16. https://doi.org/10.1007/s11858-020-01130-4

Vorhölter, K. (2018). Conceptualization and measuring of metacognitive modelling competencies: Empirical verification of theoretical assumptions. *ZDM, 50*(1-2), 343-354. https://doi.org/10.1007/s11858-017-0909-x

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology, 105*(4), 932-945. https://doi.org/10.1037/a0031882

Walkington, C., Clinton, V., & Shivraj, P. (2018). How readability factors are differentially associated with performance for students of different backgrounds when solving mathematics word problems. *American Educational Research Journal, 55*(2), 362-414. https://doi.org/10.3102/0002831217737028

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315-327). Macmillan.

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

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

Witherby, A. E., & Tauber, S. K. (2019). The current status of students' note-taking: Why and how do students take notes? *Journal of Applied Research in Memory and Cognition, 8*(2), 139-153. https://doi.org/10.1016/j.jarmac.2019.04.002