



RISTAL

Research in Subject-matter
Teaching and Learning

Schreiter, S., Vogel, M., Rehm, M., & Dörfler, T. (2021).
Teachers' diagnostic judgment regarding the difficulty of
fraction tasks: A reconstruction of perceived and
processed task characteristics

RISTAL 4 / 2021

Research in Subject-matter Teaching and Learning

Volume 4 – Special Issue edited by Timo Leuders & Katharina Loibl

Citation:

Schreiter, S., Vogel, M., Rehm, M., Dörfler, T. (2021). Teachers' diagnostic judgment regarding the difficulty of fraction tasks: A reconstruction of perceived and processed task characteristics [Special Issue]. *RISTAL*, 4, 127–146.

DOI: <https://doi.org/10.23770/RT1846>

ISSN 2616-7697



This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Teachers' diagnostic judgment regarding the difficulty of fraction tasks: A reconstruction of perceived and processed task characteristics

*Saskia Schreiter, Markus Vogel, Markus Rehm,
& Tobias Dörfler*

Abstract

The selection and modification of mathematical tasks requires teachers to adequately judge task difficulty. Despite such importance, little is known about teachers' cognitive processes that underlie these judgments. Considering possible interrelations with teaching experience and specific aspects of PCK/PK, this study reconstructs teachers' perception and processing of task characteristics while judging task difficulty for students. In the field of fraction calculation, the tasks' difficulty is varied systematically by modifications in the tasks' instructional design (according to cognitive load theory; e.g., split-attention vs. integrated format) and by adjusting the fractions' complexity (e.g., like vs. unlike fractions). The study was conducted with 55 pre- and 35 in-service mathematics teachers. The results suggest that both groups predominantly perceive and process mathematical, but only rarely instructional task characteristics. Contrary to expectations, the findings indicate that participants' specific PCK/PK concerning difficulty-generating task characteristics is high in both areas – fraction and instructional design.

Keywords

Diagnostic judgments, cognitive processing, task difficulty, cognitive load, PCK/PK, teaching experience

1 Introduction

One of mathematics teachers' everyday activities is to select tasks and modify their difficulty (Leuders & Prediger, 2016). To do so, teachers need to be able to adequately judge task demands and difficulty. The sample presented in Figure 1 constitutes a typical 6th grade fraction task. The task's difficulty is determined by several *difficulty-generating task characteristics* (Leuders & Prediger, 2016) that teachers need to consider when judging task difficulty. Firstly, the task requires the addition of a natural number and unlike fractions which makes the task more difficult for students than tasks requiring solely the addition of like or unlike fractions (Padberg & Wartha, 2017). Furthermore, especially for students at the beginning of the learning process, the task's instructional design might cause high cognitive load as the relevant information is not presented close to each other (split-attention effect; Sweller, Ayres, & Kalyuga, 2011). To solve the task, learners are required to split their attention between different information sources (the problem definition, the graphic, and the time information) and to integrate them mentally. This process takes up working memory capacity (Ayres & Sweller, 2014) that is needed to perform mathematical tasks. When teachers judge the task's difficulty, they need to draw on various aspects of professional knowledge in order to identify difficulty-

generating task characteristics. This includes, for example, topic-specific knowledge on task characteristics that influence the difficulty of fraction tasks (= specific aspects of pedagogical content knowledge [PCK]; Shulman, 1986) or more general knowledge of instructional task characteristics that have a beneficial or detrimental effect on learning (= specific aspects of pedagogical knowledge [PK]; Shulman, 1986). Numerous studies in the field of mathematics (e.g., Ostermann, Leuders, & Nückles, 2017) and across disciplines (e.g., McElvany et al., 2009) have investigated and documented the influence of specific PCK on the accuracy of teachers' judgement of task difficulty. Furthermore, teachers may also draw on what they know from their teaching experience to be traditionally difficult for students. Numerous studies show, however, that teachers' judgments vary considerably in accuracy (e.g., Anders, Kunter, Brunner, Krauss, & Baumert, 2010; Karing & Artelt, 2013) and that teachers often fail to adequately consider the positive or detrimental effects of task design (Hellmann & Nückles, 2013). Aiming to develop a better understanding of diagnostic judgments, a stronger research interest in the cognitive processes that underlie teachers' judgments has emerged (Loibl, Leuders, & Dörfler, 2020). Which task characteristics are perceived and processed by teachers? Which factors (e.g., teaching experience, specific PCK/PK) interrelate with teachers' perception and processing of tasks characteristics? Questions like these need to be addressed to gain insights into teachers' judgment processes, which, in turn, can help to design effective training that aims to foster teachers' diagnostic judgment skills.

Considering possible interrelations with teaching experience and specific PCK/PK aspects, this study aims to reconstruct¹ teachers' perception and processing of mathematical and instructional task characteristics while judging the difficulty of fraction tasks for students.



Fig.1. Sample for a fraction task with specific difficulty-generating task characteristics.

¹ Note: *reconstruction* does not refer to a reconstructive methodical approach in the sense of qualitative research. Further details on how perceived and processed task characteristics were reconstructed on the basis of observable diagnostic behavior are described in chapter 4.3.

2 Theoretical Background: Teachers' diagnostic judgments of task difficulty

When a teacher assesses the difficulty of a task, *diagnostic judgment* is a term for both, the process and the result (Artelt & Rausch, 2014). Usually, the quality of diagnostic judgments is evaluated by the accuracy of the result (Südkamp, Kaiser, & Möller, 2012). Therefore, the dominant field of research on teachers' diagnostic judgments has for a long time focused on accuracy (for a review, see Südkamp et al., 2012). Although a new area of interest regarding the cognitive processes that lead to diagnostic judgments has recently emerged, the state of research in this domain is still seen as unsatisfactory (e.g., Herppich et al., 2018).

The framework DiaCoM (*Explaining Teachers' Diagnostic Judgments by Cognitive Modeling*; Loibl et al., 2020) proposes a research strategy for investigating the cognitive processes at play by specifying four components: characteristics of the diagnostic situation (SC), characteristics of the teacher (PC), diagnostic thinking (DT), and diagnostic behavior (DB). This framework has been used to investigate teachers' judgment processes in different instructional contexts and subjects, including mathematics (Rieu, Loibl, Leuders, & Herppich, 2020), foreign language acquisition (Witzigmann & Sachse, 2020) and biology (Hoppe, Renkl, & Rieß, 2020). The framework's components are specified for the present study (see Figure 2) as described below.

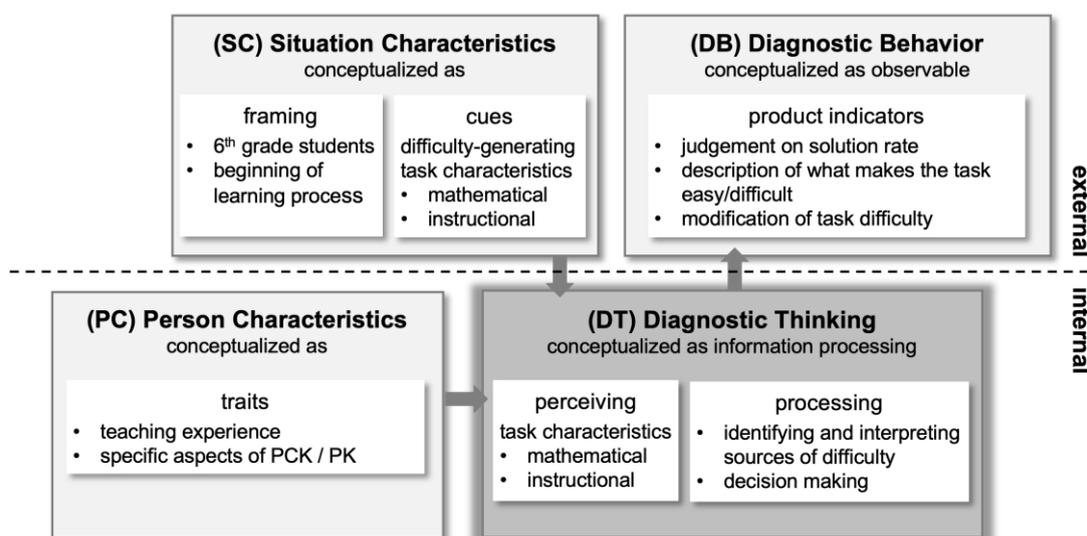


Fig. 2. The present study in light of the adapted DiaCoM framework (Loibl et al., 2020).

2.1 Situation Characteristics

Situation characteristics constitute the context of teachers' diagnostic judgment including pieces of information – so called *cues* (Loibl et al., 2020) – that teachers may (or may not) use in the judgment process. In this study, cues are conceptualized as specific difficulty-generating task characteristics that are systematically varied between fraction

tasks: a) by adapting the fraction's complexity (mathematical task characteristics) and b) by modifying the instructional design (instructional task characteristics). In the presented sample item (Figure 1), the addition of a natural number and unlike fractions (mathematical task characteristic) as well as the separately presented information sources (instructional task characteristic) serve as cues that hold information about the task's difficulty. The usage of these cues might be dependent on the context information presented to the teacher (= the *framing*). In the present study, the tasks' difficulty is to be judged for 6th grade students at the beginning of the learning process.

2.1.1 Mathematical task characteristics

Fractions are a recurrent domain of mathematics teaching programs that often presents a challenge for students (e.g., Lortie-Forgues, Tian, & Siegler, 2015; Smith, 2002). In line with their importance, fractions have been the subject of numerous studies that have tried to identify typical student errors and potential error sources (for an overview, see Padberg & Wartha, 2017). Padberg (1986) was able to establish a sequence of difficulty for the addition of fractions: While a large number of the students could add fractions with a common denominator correctly (85 %), the solution rates dropped when it came to adding fractions with unlike denominators (70 %) and even more when the tasks required the addition of a natural number with a fraction (55 %). More recent studies report similar results or even lower solution rates (e.g., Brown & Quinn, 2006). The most common misconception when adding fractions is that fractions' numerators and denominators can be treated as separate whole numbers ($\frac{a}{b} + \frac{c}{d} = \frac{a+c}{b+d}$) (e.g., Post, 1981). This mistake can be found more often when adding unlike as opposed to like fractions. When adding a natural number with a fraction, mistakes often occur as students add the natural number to the numerator ($n + \frac{a}{b} = \frac{n+a}{b}$) or choose a more complicated as necessary calculation (e.g., $3 + \frac{1}{5} = \frac{3}{1} + \frac{1}{5} = \frac{15}{5} + \frac{1}{5} = \frac{16}{5} = 3\frac{1}{5}$ instead of $3 + \frac{1}{5} = 3\frac{1}{5}$). Padberg and Wartha (2017) argue that one reason for student difficulties with this - on a semantic level generally easy - type of fraction, is that teachers tend to underestimate the difficulty of these tasks for students and thus neglect this type of tasks in class. The difficulty of a fraction task is assumed to be not only determined by the complexity of the fraction, but also by other task characteristics, such as the instructional design.

2.1.2 Instructional task characteristics

According to the Cognitive Load Theory (CLT; e.g., Sweller et al., 2011) the cognitive load imposed on working memory originates from two categories of cognitive load: intrinsic cognitive load (ICL) and extraneous cognitive load (ECL). A third category of cognitive load – germane cognitive load – is, according to latest research, assumed not to contribute to the total cognitive load (Sweller, van Merriënboer, & Paas, 2019). ICL results from the complexity of a task (e.g., like vs. unlike fractions, cf. 2.1.1) and is determined by levels of element interactivity (according to Sweller et al. (2011): the extent different task elements interact and must be processed simultaneously rather than as single) and by the prior knowledge of the student. ECL is imposed by the manner in which the information is presented. If either ICL and/or ECL is high, working memory can become

overloaded and inhibit successful learning (Ayres, 2006). Unlike ICL, that is inherent to the task, ECL can be reduced by changing the instructional design (Sweller et al., 2019). Especially at the beginning of a learning process, when students deal with a wide range of new information that take up large amounts of working memory capacity, it is important that teachers select and modify the instructional design of tasks in a way that ECL is low. Numerous researchers have identified several guidelines for the instructional design of learning material that work successfully to reduce ECL (for an overview, see Sweller et al., 2011). The following three design guidelines were used to vary ECL between tasks in the present study:

A *split-attention effect* occurs when learners are required to split their attention between at least two sources of information (e.g., graphic and text) that have been separated either spatially or temporary (Sweller et al., 2011). In the above presented task (Figure 1), relevant information (problem definition, graphic and corresponding time information) is presented separately. To solve the task, learners have to split their attention, search for corresponding information and mentally integrate them. This process causes ECL which might inhibit learning (e.g., Ayres & Sweller, 2014). The split-attention effect has been investigated extensively in research on subject-matter education and could be found in several different subjects, such as mathematics (Tarmizi & Sweller, 1988), geography (Purnell, Solman, & Sweller, 1992) and economics (Ayres, & Youssef, 2008). Research results suggest that a split-attention design has negative consequences and should be replaced by an integrated-format design where relevant information is presented close to each other (Sweller, van Merriënboer, & Paas, 1998).

The *redundancy effect* suggests that learning is hindered when learners are presented with the same information in two or more forms (cf. sample items, Figure 4b) and/or with additional information that is not relevant for solving the task (Sweller et al., 2011). Requiring learners to process redundant information takes up precious working memory capacity that is needed to solve the task. Research shows that eliminating redundant information from tasks results in enhanced learning (e.g., Mayer, Heiser, & Lonn, 2001). The redundancy effect is pervasive and could be found in a wide variety of instructional contexts and subjects, including mathematics and science, but also foreign language acquisition or social science and humanities (for an overview, see Sweller et al., 2011)

Step-by-step guidance (following e.g., Kalyuga, Chandler, & Sweller, 2001; Hellmann & Nückles, 2013,) provides learners with a guideline by segmenting a task into sub-tasks. This procedure reduces cognitive load and hence frees working memory capacity. Without this instructional guidance, learners have to hold the problem definition, the aim and all the solution steps needed to solve the task in their working memory at the same time, struggling to put them into a correct order. Research shows, however, that segmenting a complex task is only effective under certain conditions (cf. Blayney, Kalyuga, & Sweller, 2015): while novice learners profit significantly as their working memory might be overloaded if they deal with the task in its entity, the same step-by-step instruction might cause increased cognitive load for more experienced learners.

The benefit of the above presented instructional guidelines depends on learners' level of expertise and is especially effective for learners with no or only little prior knowledge (Kalyuga, Ayres, Chandler, & Sweller, 2003). For teachers or for more experienced learners, the positive effect may be eliminated or even have negative consequences – this is what Kalyuga et al. (2003) call the *expertise reversal effect*. In order to perceive the beneficial or detrimental effect of task design when judging task difficulty for students, teachers can thus not recur to own perceived difficulties and hence need to draw on specific knowledge and/or teaching experience (see 2.2).

2.2 Person characteristics

Person characteristics include *traits* that may influence teachers' diagnostic judgment. In order to identify difficulty-generating task characteristics when judging task difficulty, teachers need to draw on various aspects of professional knowledge. This includes topic-specific knowledge on task characteristics that influence task difficulty, for example in the domain of fractions. According to Shulman (1986, p. 9), such “an understanding of what makes the learning of specific topics easy or difficult” constitute key components of PCK. Moreover, teachers may also draw on more general knowledge of instructional task characteristics that have a beneficial or detrimental effect on learning. An understanding of general methods and strategies that are proven to work well for teaching can be conceived as PK (Shulman, 1986). Research found that the accuracy of pre-service teachers' judgment of task difficulty could be improved by providing relevant PCK about task characteristics and students' misconceptions in the area of functions and graphs (Ostermann et al., 2017). Based on these empirical findings, there is good reason to assume that specific aspects of PCK interrelate with teachers' perception and processing of task characteristics. Hellmann and Nückles (2013) found that pre- as well as in-service teachers fail to adequately consider task design when judging task difficulty for students. However, it remains unclear if this finding is interrelated with missing PK concerning difficulty-generating task characteristics in instructional design.

Besides specific knowledge about difficulty-generating task characteristics that can already be imparted during teacher education at university, teachers gain additional explicit and implicit knowledge aspects by personal teaching experience at school (van Ophuysen, 2006). When judging a task's difficulty, teachers may therefore also draw on what they know from their teaching experience to be traditionally difficult for students. Most research that illuminates the potential influence of teaching experience on diagnostic judgments has hitherto focused on judgment accuracy. On the one hand, studies found that in-service teachers in comparison to pre-service teachers seem to have a “better mental model of what a student is able to achieve” (Hellmann & Nückles, 2013, p. 2518), which enables them to make better judgments on task difficulty and student performance. Likewise, Ostermann, Leuders, and Nückles (2015) report that the accuracy of teachers' judgment of task difficulty improves with increasing teaching experience. On the other hand, however, numerous studies failed to find a positive influence of teaching experience on judgment accuracy (e.g., Dünnebier, Gräsel, & Krolak-Schwerdt, 2009). Against this inhomogeneous theoretical background, a more detailed examination of underlying cognitive processes might contribute to further illuminate the

potential influence of teaching experience on the process as well as the result of diagnostic judgments. According to van Ophuysen (2006), it can be expected that teachers' judgment processes change with increasing teaching experience because of daily assessment practice.

2.3 Diagnostic behavior

Diagnostic behavior refers to teachers' observable behavior. When judging the difficulty of tasks for students, different kinds of diagnostic behavior may be directly observed, such as: teachers' judgment of task solution rate, a description of what makes the task easy or difficult for students and, relating thereto, the modification of the task to make it easier or more difficult. Most studies in this domain have measured the accuracy of teachers' judgments by comparing them with objective data, for example by correlating teachers' estimation of solution rates with actual student results on specific tasks (e.g., Anders et al., 2010; Karing & Artelt, 2013; Ostermann et al., 2015, 2017). The results of these studies show two overarching trends: Firstly, teachers' judgments generally vary considerably in accuracy and secondly, teachers tend to overestimate students' achievement or underestimate the difficulty of tasks. In the study by Hellmann and Nückles (2013), teachers estimated solution rates for specific tasks that varied in instructional design according to CLT guidelines. Comparing teachers' judgments with empirical student solution rates, the authors found that teachers failed to adequately consider the positive or detrimental effects of task design. However, in most studies focusing on the result of diagnostic judgments, it remains unclear how teachers get to their result. Which difficulty-generating task characteristics teachers perceive and process in the judgement's genesis is crucial to understand its result. A better understanding of these cognitive processes contributes to a more sophisticated picture of teachers' diagnostic behavior which, in turn, can help to design training to foster teachers' diagnostic judgment skills.

2.4 Diagnostic thinking

The component of diagnostic thinking comprises internal cognitive processes during diagnostic judgments and is located at the heart of the DiaCoM framework. When teachers judge the difficulty of a task, several cognitive processes are activated, such as *perceiving* task characteristics that hold information about the difficulty of the task, *identifying* sources of difficulty, *interpreting* them in light of the given context information and finally *making a decision* about the task's difficulty (adapted from Loibl et al., 2020). For example, when judging the difficulty of the task presented in Figure 1, a teacher may (or may not) perceive mathematical task characteristics (e.g., addition of a natural number and unlike fractions) and instructional task characteristics (e.g., relevant information is not presented close to each other) that hold information about the task's difficulty. If the task characteristics are perceived, they can be further processed: sources of difficulty may be identified and interpreted with regard to the given context information (the task's difficulty is to be judged for 6th graders at the beginning of the learning process). However, the positive effect of CLT task design guidelines may be eliminated for teachers (see expertise-reversal effect, 2.1) and hence specific instructional task

characteristics may not be perceived or, if perceived, may be misinterpreted regarding their difficulty. Also, teachers often tend to underestimate the difficulty of adding a natural number and a fraction for students (see 2.1). It can thus be assumed that teachers need to draw on specific PCK/PK (see 2.2) and/or teaching experience in order to perceive and adequately process the above-mentioned task characteristics which is an important prerequisite in order to come to an adequate judgment of task difficulty.

Teachers' perception and processing of task characteristics constitute internal cognitive processes that cannot be directly observed. The present study aims to reconstruct the perception and processing of mathematical and instructional task characteristics. The prerequisite for this reconstruction is a systematic variation of mathematical and instructional task characteristics that is intended to allow conclusions from the observable diagnostic behavior to the task characteristics that teachers have perceived and processed.

3 Research Questions

Considering possible interrelations with teaching experience and specific PCK/PK aspects (see 2.2), this study investigates teachers' perception and processing of mathematical and instructional task characteristics while judging task difficulty for students. For the tasks used in this study we chose fractions, an important and recurrent domain in the mathematics curriculum (see 2.1). The tasks' difficulty is varied systematically by adjusting the fraction's complexity and by modifications in instructional design according to the above described CLT design guidelines (see 2.1). The design of the study includes pre-service as well as in-service teachers in order to investigate potential interrelations between teaching experience (see 2.2) and the perception and processing of task characteristics. In line with the need for research pointed out in the previous section, this study aims to address the following research questions:

1. Which task characteristics (mathematical vs. instructional) do teachers perceive and process when judging task difficulty for students?
2. Do specific aspects of PCK/PK interrelate with teachers' perception and processing of task characteristics?
3. Do pre-service and in-service teachers differ in perceiving and processing task characteristics?

4 Methodology

4.1 Sample

A total of 55 pre-service teachers majoring in mathematics (mean study time being 5.43 semesters, $SD = 1.74$) and 35 in-service mathematics teachers (mean time of teaching experience being 11.43 years, $SD = 9.01$) participated in the study. The in-service teachers were recruited from three different secondary school types located in large cities as well as rural areas across the state of Baden-Württemberg, Germany. The participants

completed a paper-and-pencil test (cf. the following section) at their university or their schools in the presence of the test administrator. They have been given sufficient time to finish the test without time pressure.

4.2 Materials and design

Corresponding to our research questions, a paper-and-pencil test comprising two main sections (diagnostic test and test of specific PCK/PK aspects) was designed.

The diagnostic test includes six tasks that involve the addition of fractions. Ecological validity was ensured by drawing upon tasks from conventional 6th grade school books. Between these tasks, difficulty-generating task features were varied systematically by modifying the instructional design (instructional task characteristics) and by adapting the fraction's complexity (mathematical task characteristics). High (+) vs. reduced (-) ECL was caused by applying one of the following CLT guidelines for instructional design: split-attention effect, redundancy effect, step-by-step guidance (see 2.1). Three of the tasks were presented in an ECL (+) version and three were presented in an ECL (-) version, balanced within the test. Every task presented in its ECL (+) version had a corresponding item in its ECL (-) version, with the same mathematical level of difficulty but a different context, elsewhere in the test (see sample items, Figure 3). This design was chosen to address and avoid the following two issues: Presenting the same task twice except for one difference in task design might firstly lead to repetition effects, and secondly, it might implicitly direct participants' perception to the instructional design and thus revealing the study's objective. This methodological approach is based on the study by Hellmann and Nückles (2013), who investigated teachers' judgment accuracy considering the difficulty of tasks that differ in instructional design. To address the scope of the present study, this methodological design was developed further by integrating a systematic variation of mathematical task features. In accordance to Padberg and Wartha (2017), the mathematical difficulty between pairs of corresponding items was varied in three levels (starting with the highest difficulty): addition of a natural number and unlike fractions – addition of unlike fractions – addition of like fractions.

The theoretically derived difficulties of the tasks used in this study were further validated in a study with $n = 44$ secondary school students from grade 6. Findings show that the participating students solved significantly more ECL (-) than ECL (+) tasks correctly ($F(1, 43) = 3.74, p = .03, \eta^2 = .08$). Furthermore, significant differences could be found between solution rates of tasks with different mathematical levels of difficulty ($F(2, 42) = 25.21, p < .001, \eta^2 = .38$). Bonferroni-adjusted post-hoc analysis revealed a significant difference ($p < .001$) in student solution rates between tasks that involve the addition of like fractions ($M = .74, SD = .06$) in comparison to tasks involving the addition of unlike fractions ($M = .41, SD = .07$) or a natural number and unlike fractions ($M = .30, SD = .06$). Differences in the solution rates between tasks involving the addition of unlike fractions and tasks that include a natural number and unlike fractions did not reach statistical significance ($p = .115$).

Each task in the diagnostic test was presented on a separate double page. First, participants were asked to solve the task by themselves, which, according to Leuders and Leuders (2013), leads to richer and more specific judgments. Then, they were given some information about the situational context that they should consider for their judgment (e. g., the tasks' difficulty is to be judged for a class with 30 students of grade 6 at the beginning of the learning process). In the following, three different kinds of diagnostic behavior (DB) were captured for each task: DB1: Predict how many students would presumably solve this task correctly? (Numerical entry between 0 and 30.); DB2: What makes this task easy or difficult for students? and DB3: Please describe how this task could be modified to make it easier/more difficult for students (DB2 and DB3 as open-ended questions).

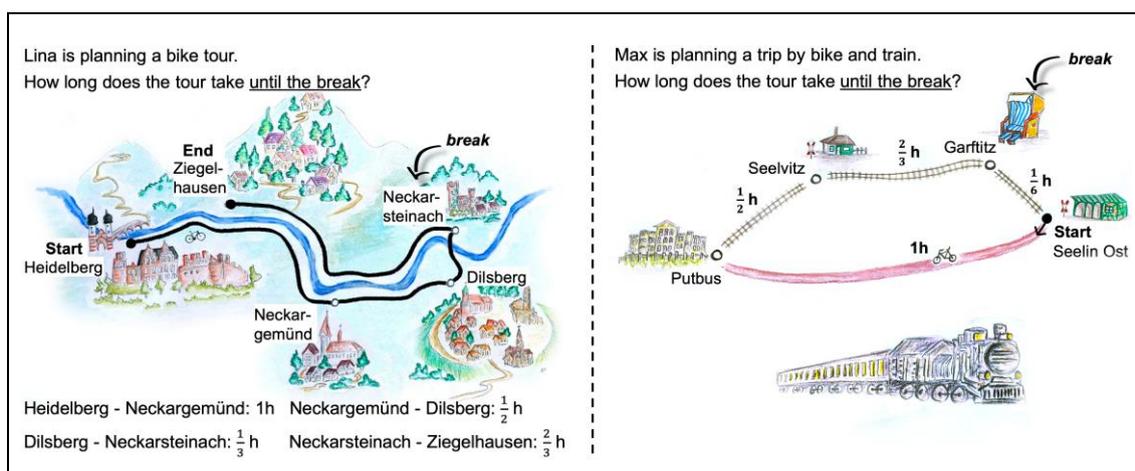


Fig. 3. Sample items of the diagnostic test: high vs. reduced Extraneous Cognitive Load according to the split-attention effect.

The second section of the paper-and-pencil test, a test of specific PCK/PK aspects, was designed to measure knowledge regarding difficulty-generating task characteristics when adding fractions (specific PCK aspects, two items, see sample item Figure 4a) and regarding instructional task characteristics that have a beneficial or detrimental effect on learning according to CLT design guidelines (specific PK aspects, three items, see sample item Figure 4b). With this specific focus, the PCK/PK test captures participants' knowledge regarding the cues that have been systematically varied between the tasks in the diagnostic test. In contrast to the diagnostic test, two tasks that only differ in one difficulty-generating characteristic are presented next to each other and participants have to judge which task is easier for students ("The left task is easier for students than the right task") on a five-point Likert scale from "I fully agree" to "I do not agree at all". Ticking the option in the middle should give participants the option to express that both tasks are considered equally easy or difficult for students. Opposing two tasks that only differ in one characteristic aims to capture specific knowledge without the necessity that participants are familiar with the terminology of CLT.

A pilot study with 36 pre-service teachers was conducted to evaluate feasibility of both the diagnostic test and test of specific PCK/PK aspects. On this basis, minor changes in terms of wording were made to ensure and improve comprehensibility.

a)	$\frac{2}{8} + \frac{3}{4}$	$1 + \frac{2}{8}$																
b)	<p>24 students voted on where to go for this year's school trip. The results showed that 10 students want to go to a climbing park, 8 students want to go to the zoo and 6 students want to go to a museum. What proportion of the students voted for each trip? Write the proportions as fractions.</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Trip</th> <th>Votes</th> </tr> </thead> <tbody> <tr> <td>Climbing park</td> <td>10</td> </tr> <tr> <td>Zoo</td> <td>8</td> </tr> <tr> <td>Museum</td> <td>6</td> </tr> </tbody> </table>	Trip	Votes	Climbing park	10	Zoo	8	Museum	6	<p>24 students voted on where to go for this year's school trip. What proportion of the students voted for each trip? Write the proportions as fractions.</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Trip</th> <th>Votes</th> </tr> </thead> <tbody> <tr> <td>Climbing park</td> <td>10</td> </tr> <tr> <td>Zoo</td> <td>8</td> </tr> <tr> <td>Museum</td> <td>6</td> </tr> </tbody> </table>	Trip	Votes	Climbing park	10	Zoo	8	Museum	6
Trip	Votes																	
Climbing park	10																	
Zoo	8																	
Museum	6																	
Trip	Votes																	
Climbing park	10																	
Zoo	8																	
Museum	6																	

Fig. 4. Sample items: test of a) specific PCK aspects: variation of the mathematical level of difficulty caused by the addition of unlike fractions vs. addition of a natural number and a fraction and (b) specific PK aspects: variation of ECL caused by including redundant vs. non-redundant information.

4.3 Reconstruction of perceived and processed task characteristics

For each kind of captured diagnostic behavior, two new variables were created (reconstructed perception and processing of a) mathematical and b) instructional task characteristics) and binary coded (0 = not perceived/processed, 1 = perceived and processed).

DB1: Teachers' estimated solution rates for corresponding tasks that either differ in instructional or in mathematical difficulty were compared with each other. If, for example, the solution rate for an ECL (-) task was estimated higher than for the corresponding ECL (+) task (see Figure 3), we concluded that the instructional task characteristic was perceived and processed by the participant. The same was done with tasks that differ in their mathematical difficulty.

DB2 and DB3: Participants' answers were first assigned to one of four categories for addressed mathematical and/or instructional task characteristics (cf. Table 1). Depending on the assigned category, conclusions were drawn with regard to perceived and processed mathematical and/or instructional task characteristics. Additional described task characteristics (neither mathematical nor instructional) were not evaluated further as they were not part of the study's scope and hence were not systematically varied between the tasks.

The answers were double-coded by the first author and a student research assistant with high interrater reliability (Cohen's Kappa .97). Discrepancies were resolved through discussion.

Tab. 1. Categories for addressed task characteristics, sample answers (related to items in Figure 3) and codes for reconstructed perception and processing of mathematical vs. instructional task characteristics

DB2: What makes this task easy or difficult for students?			
Category	Sample answer (difficult)	Codes	
		mathematical	instructional
Mathematical only	" - Unlike fractions - Whole numbers such as 1h"	1	0
Instructional only	"The time data is not integrated in the map and needs to be matched"	0	1
Mathematical and instructional	"- Sketch is not directly labelled - Fractions and whole numbers are mixed"	1	1
Neither mathematical nor instructional	" <i>Until the break</i> might be overread "	0	0

DB3: How can the task be modified to make it easier or more difficult for students?			
Category	Sample answer (easier version)	Codes	
		mathematical	instructional
Mathematical only	" - Like fractions - No mixed fractions or whole numbers that need to be converted into a fraction"	1	0
Instructional only	"Include time information into the map. It is easier to read this way"	0	1
Mathematical and instructional	"- Write times directly next to track sections - No mix of fractions and natural numbers"	1	1
Neither mathematical nor instructional	"How long does the tour take <i>until the end</i> ?"	0	0

5 Results

For each kind of diagnostic behavior, mean scores for the perceived and processed task characteristics were calculated (mathematical vs. instructional, see Figure 5). Differences between the perception and processing of mathematical vs. instructional task characteristics were addressed by conducting three Bonferroni-adjusted t-tests for dependent samples. The results suggest that the participants have on average predominantly perceived the mathematical difficulty that derives from the fraction, but only rarely the difficulty caused by the instructional design (see Figure 5). The same pattern can be found in all three kinds of diagnostic behavior: participants perceived significantly more mathematical than instructional task characteristics with high effect sizes (DB1: $t(89) = 10.76$, $p < .001$, $d = 1.62$; DB2: $t(89) = 12.60$, $p < .001$, $d = 1.74$; DB3: $t(89) = 8.64$, $p < .001$, $d = 1.45$).

In order to measure to what degree the perception and processing of task characteristics between the three kinds of captured diagnostic behavior are interrelated, Pearson's correlation coefficients were calculated. Regarding the perception and processing of

instructional task characteristics, correlations could be found between DB1 and DB2 ($r(88) = .40, p < .001$), DB1 and DB3 ($r(88) = .38, p < .001$), as well as between DB2 and DB3 ($r(88) = .76, p < .001$). These results suggest that the participants who perceived and processed the difficulty that derives from the instructional design when estimating the solution rate (DB1) mostly also described the task's difficulty with instructional characteristics (DB2) and modified the task's difficulty by changes in the instructional design (DB3). Regarding the perception and processing of mathematical characteristics, correlations could be found between DB2 and DB3 ($r(88) = .55, p < .001$).

Addressing our second research question, mean scores for participants' PCK and PK regarding specific mathematical and instructional difficulty-generating task characteristics were calculated and compared with the results of the diagnostic test. Contrary to expectations, the results suggest that participants' PCK/PK concerning difficulty-generating task characteristics is high in both areas: fraction and instructional design (see Figure 5).

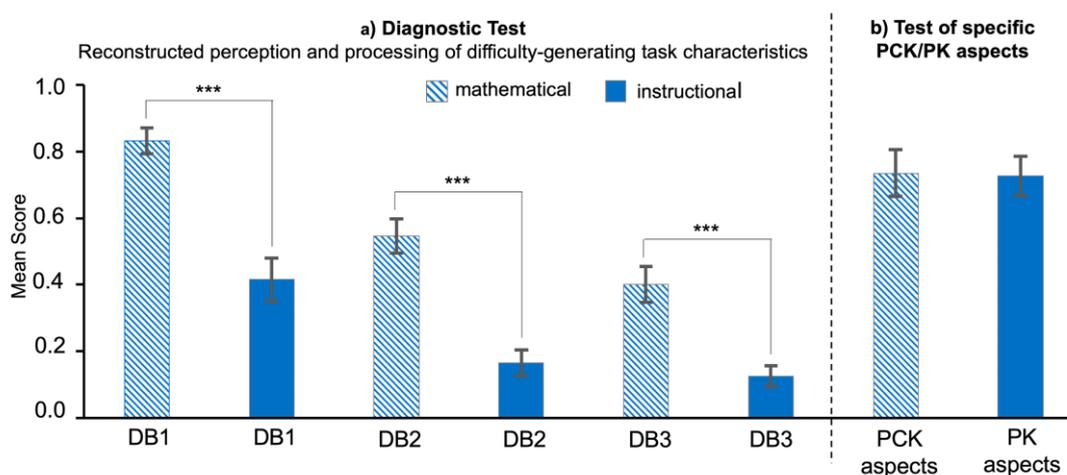


Fig. 5. Means and Standard Error for participants' a) perceived and processed difficulty-generating task characteristics (mathematical and instructional) for each type of captured diagnostic behavior and b) specific PCK/PK regarding difficulty-generating mathematical and instructional task characteristics. DB1: Estimation of solution rate, DB2 = Explanation of what makes the task easy or difficult for students, DB3 = Description of how the task can be modified to make it easier or more difficult.

*** $p < .001$.

Corresponding to the third research question, differences between pre- and in-service teachers were addressed by conducting a repeated-measures analysis of variance (ANOVA). The results show that there are no significant differences between pre- and in-service teachers' perception and processing of mathematical and instructional task characteristics (DB1: $F(1, 88) = .04, p = .838, \eta^2 = .00$; DB2: $F(1, 88) = 3.70, p = .058, \eta^2 = .04$; DB3: $F(1, 88) = .34, p = .560, \eta^2 = .00$). The tendency to predominantly perceive and process mathematical, but only rarely instructional task characteristics could be determined for both pre- and in-service teachers throughout all three kinds of captured diagnostic behavior (see Table 2). Comparing the scores for specific PCK and PK aspects, no significant differences occurred between pre- and in-service teachers ($F(1, 85) = 2.31, p = .132, \eta^2 = .03$).

Tab. 2. Means and Standard Deviations for in-service and pre-service teachers' scores on the diagnostic test (DB1 - DB3) and the test of specific PCK/PK aspects

Category	DB1		DB2		DB3		PCK / PK aspects		
	in-service teachers	pre-service teachers							
	<i>M (SD)</i>	<i>M(SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	
Mathematical	.86 (.17)	.82 (.20)	.51 (.26)	.57 (.23)	.40 (.25)	.40 (.27)	PCK	.64 (.36)	.80 (.30)
Instructional	.43 (.39)	.41 (.25)	.20 (.21)	.14 (.18)	.15 (.13)	.11 (.13)	PK	.72 (.30)	.73 (.27)

Note. In-service teachers: $n = 35$, pre-service teachers: $n = 55$.

6 Discussion

6.1 Interpretation of results

The findings of this study contribute to a more detailed picture of teachers' perception and processing of mathematical and instructional task characteristics while judging the difficulty of fraction tasks. Furthermore, they can help to enhance our understanding of interrelations between these cognitive processes and teachers' specific PCK/PK concerning difficulty-generating task characteristics. The comparison of pre-service and in-service teachers allows the investigation of possible differences between diagnostic judgments of teachers with and without teaching experience within the scope of this study empirically.

Regarding the perception and processing of specific difficulty-generating task characteristics, findings show that participants have on average predominantly perceived and processed mathematical, but only rarely instructional task characteristics. The same tendency could be determined consistently for all three kinds of captured diagnostic behavior (DB1: Estimation of solution rate, DB2 = Explanation of what makes the task easy or difficult for students, DB3 = Description of how the task can be modified to make it easier or more difficult). The results concerning the lack of perceived and processed instructional task characteristics are in line with findings reported by Hellman and Nückles (2013) who found that teachers failed to adequately consider the positive or detrimental effects of task design when judging task difficulty. However, especially at the beginning of a learning process, when the cognitive load imposed on students' working memory is generally high, it is important that teachers select or modify tasks in a way that they do not impose an additional and unnecessary ECL on students. If teachers do not consider the difficulty that derives from task design, they might risk that students' working memory becomes overloaded which, in turn, might inhibit successful learning. This empirically validated insight is derived from cognitive load theory research and applies to a wide variety of instructional contexts and subjects (for an overview, see Sweller et al., 2011). The findings of this study hence indicate a need for fostering the perception and processing of instructional task characteristics during teacher education and training.

In light of the fact that many participants' answers have revealed an adequate perceiving and processing of mathematical task characteristics, but a lack of perceiving and processing instructional task characteristics, we focus next on participants' specific knowledge regarding difficulty-generating mathematical (specific PCK aspects) and instructional task characteristics (specific PK aspects). Against this background, we were surprised to find that participants' PCK/PK regarding difficulty-generating task characteristics was on average high in both areas: fraction and instructional design. These results imply that most participants do actually know how tasks should be designed to reduce task difficulty. This raises the question of why teachers mainly perceive and process mathematical, but only rarely instructional task characteristics although they possess distinct knowledge in both areas. A possible explanation for this finding may result from the different design of the diagnostic test and the PK/PCK test. In the diagnostic test, teachers do not receive any instructions on which characteristics they should focus when judging the difficulty of individual tasks. In the test of specific PCK/PK aspects, however, teachers might be implicitly informed about what task characteristics should be considered as they have two tasks that only differ in one characteristic presented next to each other. This interpretation would indicate that the lack of perceiving and processing instructional task characteristics is not a matter of missing PK, but a matter of where teachers do and do not put their focus of attention. This explanation may be seen in line with findings reported by Südkamp et al. (2012) who found that judgments were more accurate when teachers were informed about the standard their judgment was compared with.

Comparing pre- and in-service teachers' perception and processing of mathematical and instructional task characteristics, no significant differences occurred. On average, the participating teachers – with and without teaching experience – tend to perceive and process mainly mathematical, but only rarely instructional task characteristics. Furthermore, comparing the scores for specific PCK and PK aspects, no significant differences occurred between pre- and in-service teachers. Surprisingly both groups mainly perceive and process mathematical, but only rarely instructional task characteristics, although they possess distinct knowledge in both areas. These results imply that the focus on mathematical task characteristics and the lack of perceiving and processing instructional task characteristics remains mainly unaffected by teaching experience.

6.2 Limitations and implications for further research

We would like to emphasize that the findings of this study should be considered in the light of some limitations and hence be interpreted with caution.

Firstly, the data gathering method using a paper-and-pencil test does not allow to reconstruct all perceived task characteristics, but only the ones that participants perceived, adequately processed (i.e., successfully identified sources of difficulty, adequately interpreted in light of the given framing), and eventually addressed in their written diagnostic judgments. For further research, it would be desirable to gain more detailed insights into teachers' perception and processing of task characteristics by

collecting direct process indicators, which can be realized, for example, by using eye tracking technology.

Furthermore, the study focused on two specific categories of difficulty-generating task characteristics, restricted to the domain of fractions. Of course, the difficulty of fraction tasks is influenced by many other aspects, such as linguistic complexity and numerous other mathematical aspects. Given this narrow focus and also the small number of items in the PCK/PK test, the study's findings can only give first insights that need to be interpreted carefully and investigated further by studies that focus on other difficulty-generating task characteristics and on different mathematical content domains.

Our findings show that instructional task characteristics are only rarely perceived and processed by pre- and in-service mathematics teachers when judging fraction tasks. CLT design guidelines have been investigated extensively in research on subject-matter education and have been found to be highly effective in a wide variety of instructional contexts and subjects. For further research, it would therefore be desirable to investigate the perceiving and processing of instructional task characteristics in other subjects and content domains.

The present study should be seen as a first step towards enhancing our understanding of teachers' perception and processing of task characteristics when judging the difficulty of fraction tasks. It allows answering the addressed research questions, however, at the same time, it has given rise to other questions, such as why teachers mainly do not perceive and process instructional task characteristics although they possess distinct knowledge in this area. Possible explanations discussed above should be addressed and investigated in future research.

Acknowledgments:

We are very thankful for the teachers cooperating in this research project. Furthermore, we thank Julie Pusch for her artistic support with drawing the images of the tasks used in this study.

Funding:

This study was carried out as part of the graduate school "DiaKom" funded by the Ministry of Science, Research and the Arts in Baden-Wuerttemberg, Germany.

References

- Anders, Y., Kunter, M., Brunner, M., Krauss, S., & Baumert, J. (2010). Diagnostische Fähigkeiten von Mathematiklehrkräften und ihre Auswirkungen auf die Leistungen ihrer Schülerinnen und Schüler [Mathematics teachers' diagnostic skills and their impact on students' achievements]. *Psychologie in Erziehung und Unterricht*, 57, 175–193.
- Artelt, C., & Rausch, T. (2014). Accuracy of teacher judgments. When and for what reasons? In S. Krolak-Schwerdt, S. Glock, & M. Böhmer (Eds.), *The future of education research: Volume 03. Teachers' professional development: Assessment, training, and learning* (pp. 27–43). Rotterdam, Boston, Taipei: Sense Publishers.
- Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology*, 20(3), 287–298. <https://doi.org/10.1002/acp.1245>
- Ayres, P., & Youssef, A. (2008). Investigating the influence of transitory information and motivation during instructional animations. In P. A. Kirschner, F. Prins, V. Jonker, & G. Kanselaar (Eds.), *Published Proceedings of the Eighth International Conference for the Learning Sciences* (Vol. 1, pp. 68–75). Utrecht, The Netherlands.
- Ayres, P., & Sweller, J. (2014). The Split-Attention Principle in Multimedia Learning. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 206–226). Cambridge: Cambridge University [Press. https://doi.org/10.1017/CBO9781139547369.011](https://doi.org/10.1017/CBO9781139547369.011)
- Blayney, P., Kalyuga, S., & Sweller, J. (2015). Using Cognitive Load Theory to Tailor Instruction to Levels of Accounting Students' Expertise. *Educational Technology & Society*, 18(4), 199–210.
- Brown, G., & Quinn, R. J. (2006). Algebra Students' Difficulty with Fractions: An Error Analysis. *Australian Mathematics Teacher*, 62(4), 28–40.
- Dünnebier, K., Gräsel, C., & Krolak-Schwerdt, S. (2009). Urteilsverzerrungen in der schulischen Leistungsbeurteilung [Biases in teachers' assessments of student performance: An experimental study of anchoring effects]. *Zeitschrift für Pädagogische Psychologie*, 23(34), 187–195. <https://doi.org/10.1024/1010-0652.23.34.187>
- Hellmann, K., & Nückles, M. (2013). Expert Blind Spot in Pre-Service and In-Service Mathematics Teachers: Task Design moderates Overestimation of Novices' Performance. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Cooperative minds: Social interaction and group dynamics; proceedings of the 35th Annual Meeting of the Cognitive Science Society*. Austin, Tex.: Cognitive Science Soc.
- Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Südkamp, A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education*, 76, 181–193. <https://doi.org/10.1016/j.tate.2017.12.001>
- Hoppe, T., Renkl, A., & Rieß, W. (2020). Förderung von unterrichtsbegleitendem Diagnostizieren von Schülervorstellungen durch Video- und Textvignetten [Fostering on-the-fly judgements of students' conceptions using video and text vignettes]. *Unterrichtswissenschaft*, 71(6), 382. <https://doi.org/10.1007/s42010-020-00075-7>
- Kalyuga, S., Chandler, P., & Sweller, J. (2001). Learner Experience and Efficiency of Instructional Guidance. *Educational Psychology*, 21(1), 5–23. <https://doi.org/10.1080/01443410124681>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, 38(1), 23–31. https://doi.org/10.1207/s15326985Sep3801_4
- Karing, C., & Artelt, C. (2013). Genauigkeit von Lehrpersonenurteilen und Ansätze ihrer Förderung in der Aus- und Weiterbildung von Lehrkräften [Accuracy of teacher judgments and considerations on their improvement via teacher education]. *Beiträge zur Lehrerbildung*, 31(2), 166–173.

- Leuders, J., & Leuders, T. (2013). Improving diagnostic judgement of preservice teachers by reflective task solution. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education*. Kiel.
- Leuders, T., Dörfler, T., Leuders, J., & Philipp, K. (2018). Diagnostic Competence of Mathematics Teachers: Unpacking a Complex Construct. In T. Leuders, K. Philipp, & J. Leuders (Eds.), *Diagnostic Competence of Mathematics Teachers* (Vol. 3, pp. 3–31). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-66327-2_1
- Leuders, T., & Prediger, S. (2016). *Flexibel differenzieren und fokussiert fördern im Mathematikunterricht. Sekundarstufe I + II* [Flexible differentiation in secondary mathematics]. Berlin: Cornelsen.
- Loibl, K., Leuders, T. & Dörfler, T. (2020). A Framework for Explaining Teachers' Diagnostic Judgements by Cognitive Modeling (DiaCoM). *Teaching and Teacher Education*, 91. <https://doi.org/10.1016/j.tate.2020.103059>
- Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review*, 38, 201–221. <https://doi.org/10.1016/j.dr.2015.07.008>
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93(1), 187–198. <https://doi.org/10.1037/0022-0663.93.1.187>
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Ullrich, M., et al. (2009). Diagnostische Fähigkeiten von Lehrkräften [Teachers' diagnostic skills]. *Zeitschrift für Pädagogische Psychologie*, 23(34), 223–235. <https://doi.org/10.1024/1010-0652.23.34.223>.
- Ostermann, A., Leuders, T., & Nückles, M. (2015). Wissen, was Schülerinnen und Schülern schwer fällt. Welche Faktoren beeinflussen die Schwierigkeitseinschätzung von Mathematikaufgaben? [Knowing what students know. Which factors influence teachers' estimation of task difficulty?]. *Journal für Mathematik-Didaktik*, 36(1), 45–76. <https://doi.org/10.1007/s13138-015-0073-1>
- Ostermann, A., Leuders, T., & Nückles, M. (2017). Improving the judgment of task difficulties: prospective teachers' diagnostic competence in the area of functions and graphs. *Journal of Mathematics Teacher Education*, 21(6), 579–605. <https://doi.org/10.1007/s10857-017-9369-z>
- Padberg, F. (1986). Über typische Schülerschwierigkeiten in der Bruchrechnung - Bestandsaufnahme und Konsequenzen [Typical difficulties of pupils with fractions – survey and consequences]. *Der Mathematikunterricht*, 32(3), 58–77.
- Padberg, F., & Wartha, S. (2017). *Didaktik der Bruchrechnung* (5. Auflage) [Didactics of fractions]. *Mathematik Primarstufe und Sekundarstufe I + II*. Berlin: Springer Spektrum. Retrieved from <http://dx.doi.org/10.1007/978-3-662-52969-0>.
- Post, T. (1981). Fractions: Results and Implications from National Assessment. *The Arithmetic Teacher*, 28(9), 26–31.
- Purnell, K. N., Solman, R. T., & Sweller, J. (1992). The effects of technical illustrations on cognitive load. *Instructional Science*, 20(5-6), 443–462. <https://doi.org/10.1007/BF00116358>.
- Rieu, A., Loibl, K., Leuders, T., & Herppich, S. (2020). Diagnostisches Urteilen als informationsverarbeitender Prozess – Wie nutzen Lehrkräfte ihr Wissen bei der Identifizierung und Gewichtung von Anforderungen in Aufgaben? [Judging task difficulty: Effects of PCK and time pressure on cognitive processes]. *Unterrichtswissenschaft*, 61(5), 738. <https://doi.org/10.1007/s42010-020-00071-x>
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2), 4. <https://doi.org/10.2307/1175860>
- Smith, J. P. (2002). The development of students' knowledge of fractions and ratios. In B. H. Litwiller & G. W. Bright (Eds.), *Yearbook / National Council of Teachers of Mathematics: Vol. 2002. Making sense of fractions, ratios, and proportions*. Reston, Va.: National Council of Teachers of Mathematics.

- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743–762. <https://doi.org/10.1037/a0027627>
- Sweller, J., Ayres, P., & Kalyuga, S. (Eds.). (2011). *Cognitive Load Theory*. New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4419-8126-4>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296. <https://doi.org/10.1023/A:1022193728205>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80(4), 424–436. <https://doi.org/10.1037/0022-0663.80.4.424>
- Van Ophuysen, S. (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahneempfehlung [Comparison of diagnostic decisions between novices and experts: The example of school career recommendation]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38(4), 154–161. <https://doi.org/10.1026/0049-8637.38.4.154>
- Witzigmann, S., & Sachse, S. (2020). Verarbeitung von Hinweisreizen beim Beurteilen von mündlichen Sprachproben von Schülerinnen und Schülern durch Hochschullehrende im Fach Französisch [Processing of cues in the assessment of oral language samples from pupils by university teachers in French]. *Unterrichtswissenschaft*, 12, 239. <https://doi.org/10.1007/s42010-020-00076-6>

Saskia Schreiter

worked as a secondary school teacher for the subjects Mathematics, English and Religious Education. Since 2018 she is working on her doctoral thesis in the project DiaCoM (Diagnostic Judgements by Cognitive Modeling) at the University of Education in Heidelberg.

Prof. Dr. Markus Vogel

is Professor for Mathematics and Mathematics Education at the University of Education Heidelberg and Vice Dean of the Faculty of Natural and Sociological Sciences. He is the leader of the mathematical part of the project DiaCoM in Heidelberg and the PhD advisor of Saskia Schreiter.

Prof. Dr. Markus Rehm

is Professor for Chemistry Education and speaker of the Doctoral Training Program 'Effective Competence Diagnosis in Teacher Training' at University of Education Heidelberg, Germany

Prof. Dr. Tobias Dörfler

is a full professor for psychological assessment at the University of Education in Heidelberg. He is an expert in reading development of children and teachers' professional development.