Journal für Mathematik-Didaktik (2024) 45:9
https://doi.org/10.1007/s13138-024-00232-w

ORIGINAL ARTICLE

Fostering Learning from Errors—Computer-Based Adaptivity at the Transition Between Problem Solving and Explicit Instruction

Antje Boomgaarden · Katharina Loibl · Timo Leuders

Received: 3 April 2023 / Accepted: 7 May 2024 / Published online: 14 June 2024

Abstract When learners acquire new content by working on a problem-solving task prior to explicit instruction, their attempts to solve the problem usually represent only partial steps on the way to the target concept. Both, theoretical assumptions on conceptual change as well as empirical findings on effective instructional formats with incorrect solutions, suggest that it is beneficial to address incorrect student solutions in a (subsequent) instruction phase by comparing incorrect and correct solutions. There is initial evidence that learning is most successful when learner compare the correct solution to an incorrect solution that reflects the learners' conceptual understanding from the problem-solving phase. In the present study, we investigated in a highly controlled experimental design the relevance of this fit between the learners' individual solution type from the problem-solving phase and the incorrect solution type in the instruction phase for learning success. In a computer-based learning environment, sixth graders worked on a problem-solving task to compare fractions. In the subsequent instruction phase, students in three conditions were given 1) an adaptive comparison, 2) a contra-adaptive comparison, 3) only the correct solution. Overall, there were no differences across conditions regarding the learning success. Further exploratory analyses revealed that only learners with an intermediate prior knowledge benefited from the adaptivity. This finding can be interpreted as indicator that our short intervention only induces conceptual change when basic knowledge is already available.

✉ Antje Boomgaarden · Katharina Loibl · Timo Leuders
University of Education Freiburg, Freiburg, Germany
E-Mail: antje.boomgaarden@gmx.net

Katharina Loibl
E-Mail: katharina.loibl@ph-freiburg.de

Timo Leuders
E-Mail: leuders@ph-freiburg.de

Keywords Problem-solving prior to instruction (PS-I) · Adaptivity · Solution comparison · Fractions · Conceptual change · Computer-based learning environment

Lernen aus Fehlern fördern – Computerbasierte Adaptivität am Übergang zwischen Problemlösen und expliziter Instruktion

Zusammenfassung Wenn sich Lernende vor einer expliziten Instruktion neue Inhalte durch die Bearbeitung einer Problemlöseaufgabe aneignen, stellen ihre Lösungsversuche meist nur Teilschritte auf dem Weg zum Zielkonzept dar. Sowohl theoretische Annahmen zum Konzeptwechsel als auch empirische Befunde zu effektiven Instruktionen mit fehlerhaften Lösungen legen nahe, dass es vorteilhaft ist, fehlerhafte Schülerlösungen in einer (anschließenden) Instruktionsphase durch den Vergleich von fehlerhaften und korrekten Lösungen aufzugreifen. Es gibt erste Hinweise darauf, dass der Lernerfolg am größten ist, wenn die Lernenden die korrekte Lösung mit einer fehlerhaften Lösung vergleichen, die das konzeptuelle Verständnis der Lernenden aus der Problemlösephase widerspiegelt. In der vorliegenden Studie untersuchten wir in einem stark kontrollierten experimentellen Design die Relevanz dieser Passung zwischen dem individuellen Lösungstyp der Lernenden aus der Problemlösephase und dem fehlerhaften Lösungstyp in der Instruktionsphase für den Lernerfolg. In einer computergestützten Lernumgebung bearbeiteten Lernende der sechsten Klasse eine Problemlöseaufgabe zum Vergleich von Brüchen. In der anschließenden Instruktionsphase wurde den Lernenden in drei Bedingungen 1) ein adaptiver Vergleich, 2) ein kontra-adaptiver Vergleich, 3) nur die richtige Lösung vorgelegt. Insgesamt zeigten sich keine Unterschiede zwischen den Bedingungen hinsichtlich des Lernerfolgs. Weitere explorative Analysen ergaben, dass nur Lernende mit einem mittleren Vorwissen von der Adaptivität profitierten. Dieser Befund kann als Indikator dafür gewertet werden, dass unsere kurze Intervention nur dann einen Konzeptwechsel initiiert, wenn bereits ein Basiswissen vorhanden ist.

Schlüsselwörter Problemlösen vor der Instruktion (PS-I) · Adaptivität · Vergleich von Lösungen · Brüche · Konzeptwechsel · Computerbasierte Lernumgebung

1 Introduction

In mathematics or science classes, learners are often asked to complete a problem-solving task before the topic is covered in class (e.g., Prediger et al. 2021, Van den Heuvel-Panhuizen and Drijvers 2020; Loibl et al. 2017). In such settings, the success of learning depends crucially on a subsequent explicit instruction phase (Loibl and Rummel 2014a). For example, a teacher who intends to introduce fractions via a part-whole concept (e.g., Lamon 2007), might set the following problem-solving task embedded in an accessible context (Loibl and Leuders 2018; Prediger 2011):

A group of ten boys compete against a group of five girls in hitting a trash basket with a ball. Each child attempts to score once. The boys score five times, and the girls score three times. Which group has won in a fair comparison?

In line with the principle of "Realistic Mathematical Education" (Van den Heuvel-Panhuizen and Drijvers 2020), a word problem is used here as a vehicle for building mathematical concepts on informal mathematical reasoning (*horizontal mathematization*, Freudenthal 1983, 1991; Liljedahl and Cai 2021; Schukajlow et al. 2022; Verschaffel et al. 2020). In the above problem-solving task, the learners are asked to generate a graphical solution in the form of fraction bars. Learners activate their prior knowledge to solve the task. However, the solutions that emerge from the learners' independent work in the problem-solving phase are usually incomplete and erroneous and thus only represent partial steps on the way to the target concept. Finja, for example, argues merely based on the number of goals: "The boys won, because they scored five times and the girls only three times" (Fig. 1).

To support learners in recognizing the flaws and errors in their solutions and thus becoming aware of their own knowledge gaps (Loibl et al. 2017; VanLehn et al. 2003), it is promising to offer learners a comparison of an incorrect solution with the correct solution in the subsequent instruction phase (Booth et al. 2013; Corral and Carpenter 2020; Durkin and Rittle-Johnson 2012; Heemsoth and Heinze 2014; Pillai et al. 2020; Tsovaltzi et al. 2012; Van Peppen et al. 2021). One possible comparison is illustrated in Fig. 2. Similar to Finja's solution, Till's solution is also incomplete. He argues merely based on the number of children in a group: "The boy won because they were more kids." Ole's solution is correct: He doubles both, the number of girls and their goals, to make the scores of the two teams comparable.

However, the question arises to what extent the incorrect solution in the instruction phase should reflect the learner's conceptual understanding from the problem-solving phase. It is plausible that the recognition and the processing of knowledge gaps is simplified if learners compare their individual solution type with the correct solution. Such a comparison situation is illustrated in Fig. 3. Till's solution corresponds to the conceptual understanding that reflects Finja's solution from the problem-solving phase (focus on the number of goals) and is compared with Ole's correct solution.

Learning approaches with an initial problem-solving phase followed by an explicit instructional phase (PS-I) have been proven effective for supporting conceptual understanding (DeCaro and Rittle-Johnson 2012; Kapur 2010, 2012, 2014; Kapur und Bielaczyc 2012; Loibl and Rummel 2014a, b) and transfer (e.g., Belenky and Nokes-Malach 2012; Schwartz et al., 2011). While the learning effects have been widely documented, the underlying learning mechanisms of PS-I learning scenarios are still poorly understood. At the theoretical level, Loibl and colleagues (2017) have identified three learning mechanisms of PS-I, with the recognition and processing

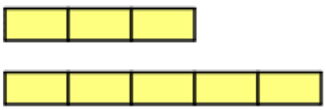


Fig. 1 Finja's solution for the problem-solving task to develop the part-whole concept

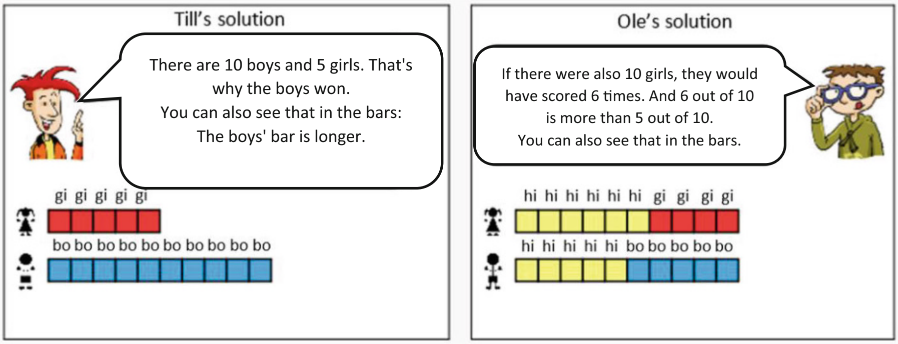


Fig. 2 Comparison of an incorrect solution with the correct solution in the instruction phase

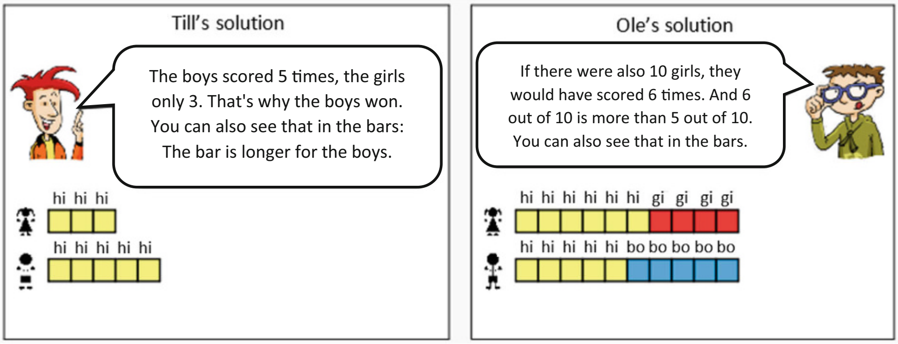


Fig. 3 Comparison of an incorrect solution that reflects the individual conceptual understanding from the problem-solving phase with the correct solution

of knowledge gaps in particular playing a crucial role. It is this mechanism that we address in the present study.

Previous studies have revealed that, addressing typical incorrect student solutions in the instruction phase by comparing them with the correct solution has a crucial impact on the development of conceptual knowledge (Loibl and Rummel 2014a; Loibl and Leuders 2018). The cognitive process of comparing solutions is a powerful learning mechanism (e.g., Rittle-Johnson and Star 2011). By comparing correct and incorrect solutions, learners can identify differences between solutions (Booth et al. 2013; Durkin and Rittle-Johnson 2012; Siegler 2002; Siegler and Chen 2008) and thus specify their knowledge gaps (Loibl and Rummel 2014a). Previous studies on comparing correct and incorrect solutions differ with regard to the type of incorrect solutions: some studies used general incorrect solutions (Booth et al. 2013; Corral and Carpenter 2020; Durkin and Rittle-Johnson 2012; Heemsoth and Heinze 2014; Pillai et al. 2020; Tsovaltzi et al. 2012; Van Peppen et al. 2021), other studies used students' individual errors (Asterhan and Dotan 2018; Gadgil et al. 2012; Heemsoth and Heinze 2016; Siegler 2002; Siegler and Chen 2008). However, a direct comparison of these two types of incorrect solutions is lacking. Post-hoc analyses by Loibl and Leuders (2019) indicate that learners who compared their individual solution type with the correct solution after a problem-solving activity experienced the greatest learning outcomes. These results support the assumption that the identification of individual knowledge gaps is facilitated when learners compare their individual solution type generated in the problem-solving phase with the correct solution in the subsequent instruction phase.

The findings of Loibl and Leuders (2019) suggest an adaptive design for the transition between the problem-solving phase and the instruction phase in the sense of a fit between the individual student solution generated in the problem-solving phase and the incorrect example solution that is compared with the correct solution in the instruction phase. However, a direct comparison of the use of typical incorrect solutions and the use of the individual solutions from the problem-solving phase is still missing.

Consequently, the aim of the present study is to investigate the relevance of adaptivity at the transition between the problem-solving phase and the instructional phase. Adaptivity here refers to the fit between the individual student solution generated in the problem-solving phase and the incorrect example solution that is compared with the correct solution in the subsequent instruction phase. Such an investigation provides insights regarding the postulated learning mechanism in PS-I of recognizing and processing knowledge gaps. For the investigation, the study uses a computer-based learning environment for the development of the part-whole concept. Thus, one of the goals of our study is the replication of previous findings, specifically that the comparison of students' typical incorrect student solutions with the correct solution has a crucial impact on the development of conceptual knowledge (Loibl and Leuders 2019), which we now investigate in the transposed computer-based learning environment.

2 Theoretical Background

In order to systematically investigate the underlying mechanisms in PS-I learning scenarios, it is necessary to consider three perspectives: (1) the instruction design, (2) the learning processes, and (3) the knowledge components (Koedinger et al. 2012; specified for PS-I in Loibl et al. 2024). With regard to the instructional design (perspective 1), there is already substantial evidence for the effective use of incorrect and correct solutions to foster learning. However, the mechanisms (perspective 2 and 3) induced by the fit between the individual solution from the problem-solving phase and the incorrect solution in the subsequent instruction phase are still poorly understood. In this study, learning (perspective 2) is conceptualized as conceptual change by taking into account existing misconceptions (perspective 3). The part-whole concept (of fractions) was selected as the learning topic for the present study since the pertaining cognitive structures have already been widely investigated and understood.

2.1 Instructional Perspective: Problem-Solving Prior to Instruction

Instructional approaches with an initial problem-solving phase followed by an explicit instruction phase (PS-I) gain attention in educational research (for reviews, see Loibl et al. 2017; Sinha and Kapur 2021). In the problem-solving phase, students attempt to solve a problem that requires the application of a concept that has not been learned before. Therefore, students usually fail at solving the problem correctly (Kapur 2010, 2012). For example, in a study by Kapur (2012) on variability, students were given data on different athletes and asked to identify the most consistent athlete in the problem-solving phase. Afterwards, students were explicitly taught the correct solution in an explicit instruction phase.

The majority of studies in PS-I research chooses the domain of mathematics, and utilizes problems in rich contexts (Kapur 2010, 2012, Loibl and Rummel 2014a). The aim of such a choice is not to represent the utility of mathematics in authentic applications (Blum and Niss 1991; Kaiser 2017). The problem context rather serves to enable students to access their prior knowledge for describing situations with mathematical means and to construct tentative individual solution approaches. This amount to the principle underlying "Realistic Mathematics Education" (Van den Heuvel-Panhuizen and Drijvers 2020), where word problems are seen as vehicle for building mathematical concepts on yet informal mathematical reasoning (*horizontal mathematization,* Freudenthal 1983, 1991; Liljedahl and Cai 2021; Schukajlow et al. 2022; Verschaffel et al. 2020). In this sense, problems for the PS-phase in PS-I should allow students to discover and construct relevant aspects of the target knowledge by referring to features of the problem situation.

Previous studies have repeatedly shown beneficial effects of problem-solving prior to instruction (PS-I) in comparison to direct instruction on conceptual understanding (DeCaro and Rittle-Johnson 2012; Kapur 2010, 2012, 2014; Kapur and Bielaczyc 2012) and transfer (e.g., Belenky and Nokes-Malach 2012; Schwartz et al. 2011). Despite this evidence on the effectiveness of PS-I, the cognitive learning mechanisms are poorly understood.

In a theoretical review, Loibl and colleagues (2017) identified three mechanisms to explain the effectiveness of PS-I[1]: (1) activation of prior knowledge, (2) recognition and processing of knowledge gaps, and (3) recognition of the deep features of the target concept. Despite these initial theoretical approach to explain the learning effectiveness of PS-I, the assumed learning mechanisms remain mostly speculative. There is a lack of studies that isolate and systematically investigate the mechanisms in a highly controlled manner (Loibl et al. 2017). While the activation of prior knowledge is a very general instructional principle and thus can explain the specific learning effectiveness of PS-I approaches only to a limited extent (Loibl et al., 2017), the recognition and processing of knowledge gaps seems to play a crucial role in the PS-I approach (e.g., Loibl and Rummel 2014a). In a 2×2 design, Loibl

[1] Further mechanisms have been proposed by Kapur and Bielaczy (2012) for productive failure. Productive failure is one specific implementation of PS-I with specific criteria (Sinha and Kapur 2021). However, we focus on the main mechanisms that are not only relevant for productive failure, but also for PS-I more generally.

and Rummel (2014a) varied the form of instruction (standard instruction or instruction that compares and contrasts typical solutions to the correct solution) and the timing of instruction (problem-solving prior to or after instruction). In their study, the order of the problem-solving phase and the instruction phase had no influence on conceptual understanding in the posttest if instruction did not build on student solutions (standard instruction). The PS-I condition only outperformed the conditions that started with explicit instruction, when instruction built on typical student solutions. Loibl and Rummel (2014a) argued that learners could benefit most from instruction if they were aware of their own knowledge gaps, which is facilitated by building instruction on student solutions.

Loibl and Leuders (2019) provided further evidence for the importance of knowledge gap awareness as a learning mechanism in PS-I. In their study, Loibl and Leuders (2019) compared three PS-I conditions that differed in the form of instruction: (1) learners received only correct solutions in the instruction phase, (2) learners received correct and incorrect solutions, (3) learners received correct and incorrect solutions with additional comparison prompts. Students who were prompted to compare correct and incorrect solutions in the instruction phase outperformed their counterparts who were not prompted or were not even presented with incorrect solutions. Post-hoc analyses provide evidence that learners who compared their individual solution type (and not other incorrect solutions) to the correct solution experienced the greatest learning outcome (Loibl and Leuders 2019). These results support the assumption that the recognition of individual knowledge gaps is facilitated when learners compare the individual solution type generated in the problem-solving phase with the correct solution in the instruction phase. Thus, the results call for adaptively taking students' individual solution attempts from the problem-solving phase into account during a comparison with the correct solution in the subsequent instruction phase. Such adaptivity is highly demanding for a teacher with about 30 students in a class. However, a digital environment that assesses students' individual solution attempts in a problem-solving phase can provide adaptive comparisons to all individual students (Boomgaarden et al. 2023).

2.2 Learning Perspective: Error Processing During Solution Comparison

Learning theories on learning from errors also include the recognition of knowledge gaps as a central element. Comparing solutions has a high potential in terms of identifying knowledge gaps. Regarding comparison processes in problem solving, Rittle-Johnson and Star (2011) identified five types of comparisons including the comparison of erroneous solutions with the correct solution. This comparison corresponds to the type of comparisons used in PS-I (e.g., Kapur 2010, 2012, 2014; Loibl and Rummel 2014a): The incorrect solutions that learners typically generate during the problem-solving phase are used for comparison with the correct solution in the subsequent explicit instruction phase. In this case, the student solutions represent partial steps on the way to the target concept and enable the identification of misconceptions from everyday life and from classroom experiences. Such misconceptions cannot simply be replaced by the target concept. An instructionally designed conceptual change (e.g., Prediger 2008; Vosniadou and Verschaffel 2004;

Vosniadou 1994) is needed to build sustainable and flexible knowledge. Conceptual change requires dissatisfaction with one's mental model (Chi 2000; Gadgil et al. 2012; Posner et al. 1982). Learners must first recognize the flaws in their mental models (imperfect mental model view, Chi 2000) before they can actively restructure their models (impasse-repair-reflect process, VanLehn 1999; VanLehn et al. 2003). Dissatisfaction with one's mental model arises from a confrontation with the inadequacies (Gadgil et al. 2012), e.g., by comparing incorrect with correct solutions. The direct comparison of student solutions in the instruction phase of PS-I can initiate such a conceptual change: By comparing correct and incorrect solutions, learners can identify differences between solutions. Thereby, learners are confronted with the limitations and weaknesses of their mental model, which in turn enables error processing and restructuring of their own knowledge in terms of conceptual change (Chi 2000; Gadgil et al. 2012.; Posner et al. 1982). Furthermore, in the comparison process, learners can also recognize similarities between the solutions, which fosters the development of abstracted schemata (cf. analogical reasoning, Gentner et al. 2003).

Overall, previous research reveal that higher learning outcomes are achieved when correct and incorrect solutions are compared than when only one type of solution is processed (e.g., Booth et al. 2013; Corral and Carpenter 2020). In their in vivo experiment, Booth et al. (2013) investigated whether correct and incorrect examples with prompts for self-explanation can improve students' conceptual understanding and procedural skills in algebra. The results suggest that incorrect examples, in combination with correct examples, can be particularly beneficial for promoting conceptual understanding. A similar effect was found by Barbieri and Devlin (2024) for learning fractions in a digital environment. In their study, a combination of correct and incorrect worked examples resulted in higher learning outcomes than studying either correct or incorrect worked examples. Studies that exclusively stimulated engagement with erroneous solutions merely indicated effects with strong support (Tsovaltzi et al., 2012) and investigated error processing only after the correct model was introduced (Tsovaltzi et al. 2012; Heemsoth and Heinze 2014, 2016). In addition, the studies reveal that when comparing correct and incorrect solutions, learners need to engage in sufficiently elaborate comparison (Große and Renkl 2007). Since learners do not automatically compare elaborately, it seems useful to explicitly prompt error processing (Catrambone and Holyoak 1989; Loibl and Leuders 2018, 2019). Previous studies differ in whether the correct solution was compared to a typical incorrect solution (e.g., Pillai et al. 2020; Van Peppen et al. 2021) or whether the correct solution was compared to an incorrect solution that reflected the learner's understanding assessed in a pretest or a previous problem-solving phase (e.g., Asterhan and Dotan 2018; Gadgil et al. 2012; Heemsoth and Heinze 2016). However, a direct comparison of the use of typical incorrect solutions and the use of the individual solutions from the problem-solving phase is still missing. On a theoretical level, only when there is a fit between the incorrect solution and the mental model, the initiated learning processes can be interpreted as conceptual change (cf. Gadgil et al. 2012). In line with this argumentation, Asterhan and Dotan (2018) postulated that the fit to learners' individual models could increase learning success.

2.3 Knowledge Perspective: Levels in the Development of the Part-Whole Concept

An instruction that builds upon previously constructed individual solutions in an adaptive manner has to take into account the cognitive structure of the target concept.

The extension from the natural numbers to the rational numbers represents a central learning content at the beginning of the lower secondary level. The topic of fractions is known to cause difficulties for many learners, even far into higher school years. Overall, the topic is very error-prone (Padberg 2002). The transition to rational numbers requires rearrangements of previous concepts, since fractions partly display different characteristics than natural numbers. If natural number concepts are activated when dealing with fractions, this often causes errors (Padberg 2002). When comparing two fractions, learners often erroneously apply the order relation of natural numbers—even after dealing with fraction in class. That is, learner treat the denominator and the numerator as natural numbers; a phenomenon called natural number bias (Ni and Zhou 2005). The natural number bias describes the exclusive focus on the numerator (part) or the denominator (whole) instead of looking at the relation of both quantities. Often learners focus on the denominator and choose the fraction with the larger denominator as larger fraction (*Dominance of the denominator*, Gould 2005; Newstead and Murray 1998; Padberg and Bienert 2000; Stafylidou and Vosniadou 2004). In Newstead and Murray's (1998) study, 38% of sixth graders committed this error. Another error is that learners focus on the numerator and conclude that a fraction is larger because the numerator is larger (McNamara and Shaughnessy 2010; Stafylidou and Vosniadou 2004). Consequently, a variety of studies provide evidence that the learning topic "fractions" requires a conceptual change (Prediger 2008; Stafylidou and Vosniadou 2004; Vamvakoussi and Vosniadou 2002; Van Hoof et al. 2017).

The fraction concept is complex and includes many subconcepts (e.g., part-whole concept, fraction as operator, fraction as ratio, fraction as quotient, fraction as measure, Charalambous and Pitta-Pantazi 2005) with many overlaps between them. In this paper, we focus on the part-whole concept because it is central for understanding fractions (Pitkethly and Hunting 1996) and for the construction of other subconcepts (Kieren 1976). This argumentation is supported by the findings by Charalambous and Pitta-Pantazi (2005) showing, for instance, that the part-whole concept could explain 98% of the variance in the subconcept "fraction as operator".

If a learner has the part-whole concept, he or she understands how to put two quantities into a part-whole relationship instead of only considering the individual quantities (part or whole) separately from each other (Lamon 2007). The formation of units (part-whole relationship) is an important mechanism for the development of the part-whole concept. The ability to decompose a whole into its constituent parts provides the flexibility of reasoning required in the field of fractions (Lamon 2007). For instance, Olive and Steffe (1994) have anchored the mechanisms by which learners construct more complex units (part-whole relationships) in a microworld. The microworld allows learners to experiment with operations (composing and decomposing, iterating, partitioning, measuring) on discrete and continuous objects that can help them develop a part-whole concept.

The following levels of conceptual development of the part-whole concept can be identified in typical student solutions (Loibl and Leuders 2018; Boomgaarden et al. 2019): (0) basic natural number concept, (1) advanced natural number concept, (2) emerging part-whole concept, (3) basic part-whole concept, and (4) flexible part-whole concept. Explicit instruction can build on these levels by selecting incorrect solutions that correspond to these levels. A more detailed description and classification of the levels is given in Chap. 4.2 Problem-solving task and types of solutions.

3 Research Questions

The present study aims to investigate the relevance of adaptivity at the transition between the problem-solving phase and the instruction phase using a computer-based learning environment for the development of the part-whole concept. Due to the content-related focus on the development of the part-whole concept, the study can be linked to previous studies that thematically set the same focus in a paper-based learning environment (Loibl and Leuders 2018, 2019). In order to investigate the relevance of an adaptive instruction phase, in the sense of a fit between the individual solution type from the problem-solving phase and the solution type in the instruction phase, a computer-based learning environment is needed, since a single teacher cannot provide adaptive instruction for 30 students at the same time. Boomgaarden et al. (2023) developed a computer-based learning environment for fraction comparison, which strongly builds on the paper-based version (Loibl and Leuders 2018, 2019; Prediger et al. 2021). The learning environment enables both valid problem-solving processes and an automated diagnosis of the problem-solving products and will be used and enhanced in the present study. Based on this prior work, it can be expected that the learning environment promotes the development of a part-whole concept.

With these considerations in mind, our main research questions are as follows:

- In PS-I-learning scenarios, does the comparison of an erroneous solution with the correct solution after a problem-solving activity foster conceptual learning outcomes?
- Does the effect of comparing an erroneous solution with the correct solution depend on the fit between the students' individual solution attempt from the problem-solving phase and the erroneous example solution in the instruction phase?

To investigate these questions, three conditions are compared that differ in the instruction phase: (1) In an adaptive condition, students receive an erroneous example solution that corresponds to their own individual solution. (2) In a contra-adaptive condition, students receive an erroneous example solution that is complementary to their own individual solution. (3) In a control condition, students receive no erroneous example solution. Students in all conditions receive a correct solution.

We hypothesize that learners who compare correct and incorrect solutions in the instruction phase following the problem-solving phase (adaptive and contra-adaptive condition) reveal higher conceptual learning outcomes than their counterparts in the control condition, who focus exclusively on the correct solution in the instruction phase (hypothesis 1).

This hypothesis focuses on replicating previous findings (Loibl and Leuders 2018, 2019) in the transposed computer-based learning environment (*computational transposition*, Hoyos 2016). Moreover, the comparison of the adaptive and the contra-adaptive condition allows for the investigation of the hypothesis, that the fit between the individual solutions and the example solutions in the instruction is relevant for the learning success:

We hypothesize that the learners in the adaptive condition achieve higher conceptual learning outcomes than their counterparts in the contra-adaptive condition (hypothesis 2).

In addition to these hypotheses, we explore the following questions on interindividual differences of profiles in the learning progression that may produce tentative explanations for variations in the learning processes and outcomes:

How do students' initial misconceptions influence the effect of the conditions (adaptive, contra-adaptive, control) on the conceptual learning outcomes?
How does the fact, whether learners recognize their individual solution in the instruction phase, influence the conceptual learning outcomes?

We approached these questions in a highly controlled computer-based adaptive environment within regular classroom sessions.

4 Methods

4.1 Participants

A group of 353 sixth graders participated in the study. Of these students, 288 had parental consent to participate in the study and were present, that is provided a complete data set, on both test days. Only these learners were included in the analysis (142 girls, 146 boys; age 11–14, $M = 11.43$, $SD = 0.64$). Against the background of previous findings, medium to large effect sizes are expected (e.g., Loibl and Leuders 2019: effect sizes between $d = 0.3$ and $d = 0.5$). A g-power analysis (cf. Faul et al. 2009) indicated that, with a power of $1\text{-}\beta = 0.95$, a sample size of 177 students is required to detect effects of size $d = 0.3$ in a design with three conditions with a one-way analysis of variance. Therefore, our sample size seems appropriate.

At home, 237 learners speak German and 51 learners speak a foreign language. Fractions were already introduced in class, but students had not yet learnt fraction comparison. Thus, the learners are able to use fractions to describe real situations.

4.2 Problem-Solving Task and Types of Solutions

The unit of the study involves comparing fractions using graphical representations. It originates from the KOSIMA project [KOntexte für SInnstiftendes MAthematiklernen] (*contexts for meaningful mathematics learning*; Prediger et al. 2021). The KOSIMA project is a research and development project focusing on the scientific development of innovative learning arrangements in mathematics education. The instructional approach of the learning environments in KOSIMA (inquiry and mathematization—active organization—intelligent practice) is based on Freudenthal's (1983) principle of guided re-invention of mathematical concepts, starting from students' intuitive resources and developing them into more structured ideas (Prediger et al. 2021), thus exhibiting a close affinity to the PS-I approach.

In the initial problem-solving phase, learners are asked to decide in a fair manner which group won a competition in trash ball: The five girls scored three goals and the ten boys scored five goals. It was clarified that each child had only one attempt. The comparison of the absolute number of goals (numerator) of both groups is not fair, because the group sizes (denominator) differ. Rather, the ratio of the number of goals to the group size must be considered. This is feasible using fraction bars of equal lengths (cf. last cell of Table 1 for an example).

The types of solutions that learners generate when working on the problem-solving task described above were categorized by Loibl and Leuders (2018).

The categories of student solutions have a hierarchical structure and increase in their structural complexity, according to literature on fraction (Lamon 2007; Padberg 2002): students consider only one quantity (0: basic natural number concept, 1: advanced natural number concept), students recognize two quantities without connecting them adequately (2: emerging part-whole concept), students relate two quantities multiplicatively (3: basal part-whole concept), even in most general situations (4: flexible part-whole concept). This hierarchy is precisely in line with the categories of the SOLO taxonomy ("structure of observed learning outcomes", Biggs and Collis 1989), where SOLO uses the generic (topic independent) categories: unistructural (see level 0 and 1 above), multi-structural (level 2), relational (level 3) and extended abstract (level 4). Therefore, we chose to also use these terms to make the hierarchy of understanding fractions more explicit (Table 1).

As revealed in the study by Loibl and Leuders (2018), many learners answer the above problem-solving task with a focus on only one quantity. For example, they might argue: "The boys won because they scored more goals" (Focus only on the number of goals; parts). Similarly, learners could reason: "The boys won because they are more kids" (Focus only on the number of children; wholes), or they assume: "The girls won because they had fewer misses than the boys" (Focus only on the missed goals; remainder). Since each of these explanations refers to ONE quantity, they can be assigned to the unistructural level in the SOLO taxonomy. However, there are still differences within the unistructural level. Most likely, all learners have noticed the number of goals, since the number of goals is very salient. However, students who argue based on the number of children or based on the number of missed goals did not integrate the number of goals into their rationale. It can be assumed that these learners have recognized that merely taking the number of goals

Table 1 Levels of the part-whole concept

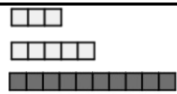
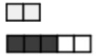
Level	SOLO	Fraction concept resp. misconception	Contextual focus	Indicative solution
0	Unistructural	Basic natural number concept	Focus on parts	
1		Advanced natural number concept	a) Focus on wholes	
			b) Focus on the remainder	
2	Multistructural	Emerging part-whole concept	Recognition of the dependence of *two* quantities: Part AND whole a) No equal whole OR	 
			b) Separating part and whole	
3	Relational	Basal part-whole concept	Consider whole as equal (refine)	
4	Extended abstract	Flexible part-whole concept	More complex schemes: simultaneous refinement (non-symbolic expansion)	

Table 2 Overview of the experimental conditions (level 3 corresponds to the target concept)

Diagnosis: *Individual level to which the learner's solution belongs:*	Control condition *Only the correct solution*	Adaptive condition *Example solution at the individual level*	*Correct solution*	Contra-adaptive condition *Example solution not at the individual level*	*Correct solution*
Non-mathematical	3	0	3	1a	3
0	3	0	3	1a	3
1a	3	1a	3	0	3
1b	3	1b	3	0	3
2a	3	2a	3	0	3
2b	3	2b	3	0	3
3	3	3	4	0	4

is not sufficient and that another quantity is necessary. Therefore, these learners are considered to already having an advanced natural number concept. In contrast to the second level, learners with an advanced natural number concept consider parts and wholes in their solutions attempts, but do not yet know how to combine two quantities in their processing. For instance, they argue: "The boys scored more goals, but the girls had fewer attempts. Thus, both groups performed equally well." Level 3 represents the target concept, a basal part-whole concept. Learners who argue on this relational level understand how to put two quantities in part-whole relation. On this level, learners realize that the wholes must be equal and therefore double the number of girls and their goals (refine). The extended abstract level represents a flexible part-whole concept, which is above what is stimulated by the learning environment in this study but can potentially be achieved by very high achievers. For instance, learners can conduct a non-symbolic expansion as shown in Table 1.

During the subsequent explicit instruction phase, learners of the adaptive condition received a solution that corresponded to their own solution type from the problem-solving phase. They compared this solution to Ole's correct solution (basal part-whole concept). Learners in the contra-adaptive condition compared the correct solution to an erroneous solution that did not correspond to their own solution type. Students in this condition received a solution on level 0 (basic natural number concept), unless they themselves generated a solution on level 0, in which case they received a solution on level 1a (advanced natural number concept) to avoid a fit to their own solution. Solutions on level 1b and level 2 were not selected. Solutions on level 1b may prompt focusing on parts and wholes to determine the remainders; solutions on level 2 explicitly address parts and wholes. Thus, these solution types would partly fit to all solutions that include parts and/or wholes (i.e., level 0, 1a, 1b, 2, 3). Learners in the control condition received only the correct solution. Table 2 provides an overview of the solution types that learners in each condition worked with during explicit instruction. Compared to Table 1, a category that comprises non-mathematical strategies in solving the task (e.g., "Boys always win, therefore it is not fair.") was included. Learners arguing non-mathematical in the problem-solving phase were treated as learners at level 0 because level 0 was considered the most basic and intuitive. Note that the instruction of these learners can, however, not be considered as adaptive or contra-adaptive.

4.3 Learning Environment

To investigate the present research questions, a computer-based learning environment was needed, since a single teacher cannot provide adaptive instruction for 30 students at the same time. The computer-based learning environment for fraction comparison was developed in the digital simulation environment "Cinderella" (Richter-Gebert and Kortenkamp 2012), based on the results of Boomgaarden et al. (2023). The computer-based learning environment strongly builds on the paper-based version (Loibl and Leuders 2018, 2019; Prediger et al. 2021), ensuring that the learning environment enables the same cognitive processes as the paper-based version (*computational transposition*, Hoyos 2016). The computer-based learning environment was found to facilitate both valid problem-solving processes and high diagnostic accuracy (Boomgaarden et al. 2023). Three rectangular shapes of different lengths are the central tool in the learning environment. These shapes can be used to form longer bars. The shapes can be coloured differently by clicking on them. The learning environment consists of three parts: At the beginning, the students get to know the functionality of the learning environment via context-free tasks. In the second part of the learning environment, the learners repeat the concept of fractions. Finally, students work on fraction comparison. During the problem-solving phase, they work on the task in the trash ball context presented above. This task asks for a graphical solution in form of fraction bars and a written answer. Learners also have to tick which group won the competition. The results of Boomgaarden et al. (2023) showed that in this learning environment, the graphical solution is the best indicator of learners' conceptual understanding. Thus, for the current study, an algorithm was integrated that diagnoses the level of conceptual understanding based on the student's graphical solution. According to the three conditions in the present study (adaptive, contra-adaptive, control condition), three versions of the subsequent instruction phase were included in the learning environment. Version one (adaptive condition) presents an erroneous solution that corresponds to the student's own solution type from the previous problem-solving phase and the correct solution. Version two (contra-adaptive condition) presents an erroneous solution that does not correspond to the student's own solution and the correct solution. Both versions include a prompt to compare the erroneous solution to the correct solution (Loibl and Leuders 2019). In addition, both versions ask students to tick on a six-point scale how well they recognize their own solution from the problem-solving phase to check implementation fidelity. This is important for the interpretation of the results because the effect of adaptivity may depend on whether the adaptive or contra-adaptive comparisons are perceived as such. Version three (control condition) only presents the correct solution.

Figure 4 illustrates the concrete task in the problem-solving phase and a potential solution. The graphical solution reflects a conceptual understanding on level 0 (see

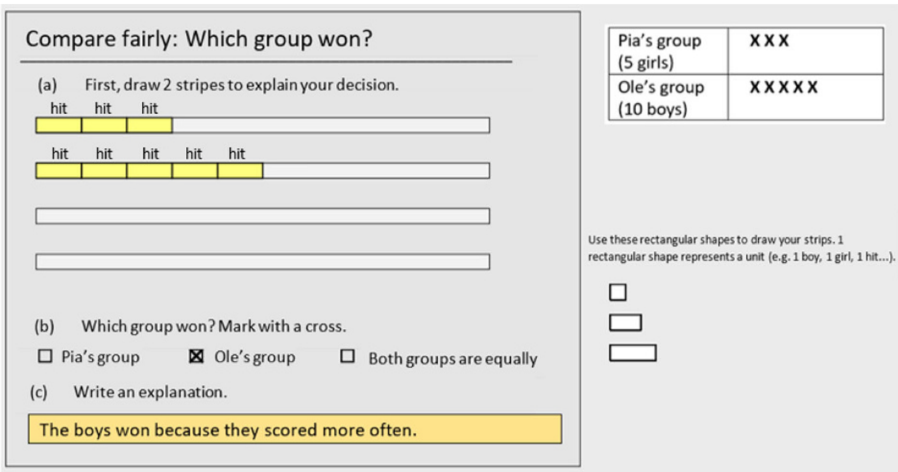


Fig. 4 Potential solution during the problem-solving phase (translated)

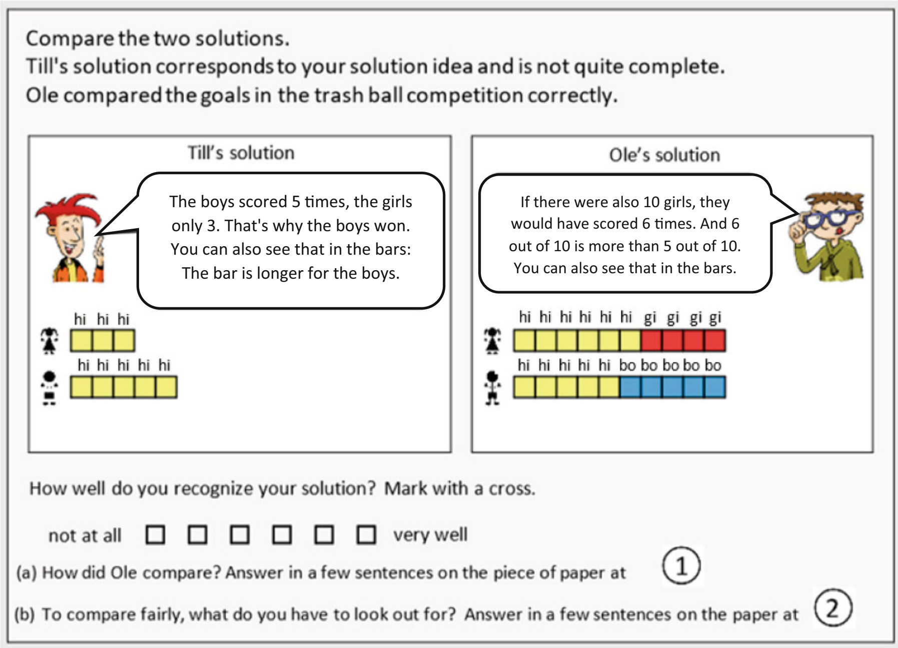


Fig. 5 Comparison of solutions in the instruction phase based on the solution from Fig. 4 (adaptive condition, translated)

Table 1). The learner argues based on the absolute number of goals and ignores the number of children in each group (cf. natural number bias: Vamvakoussi and Vosniadou 2010).

Figure 5 presents the comparison task that a learner on level 0 would receive in the instruction phase in the adaptive condition. The left solution corresponds to the learner's own solution type from the problem-solving phase. The right solution corresponds to the target concept (Table 1, level 3).

Figure 6 presents the comparison task that a learner on level 0 would receive in the instruction phase in the contra-adaptive condition. While the right-hand solution again corresponds to the target concept (Table 1, level 3), the left-hand solution reflects a different conceptual understanding (Table 1, level 1a) than the individual displayed in the problem-solving phase (Table 1, level 0).

Learners in the control condition, regardless of their own solution, receive only the correct solution (Table 1, level 3) in the instruction phase (Fig. 7).

4.4 Procedure

The study[2] covers two lessons of 90 min each (Fig. 8).

At the beginning of the first lesson, the operation comprehension (Schulz et al. 2020), text comprehension (Lenhard et al. 2009), and the information and communi-

[2] A detailed overview of the timeline of the study is given in the appendix.

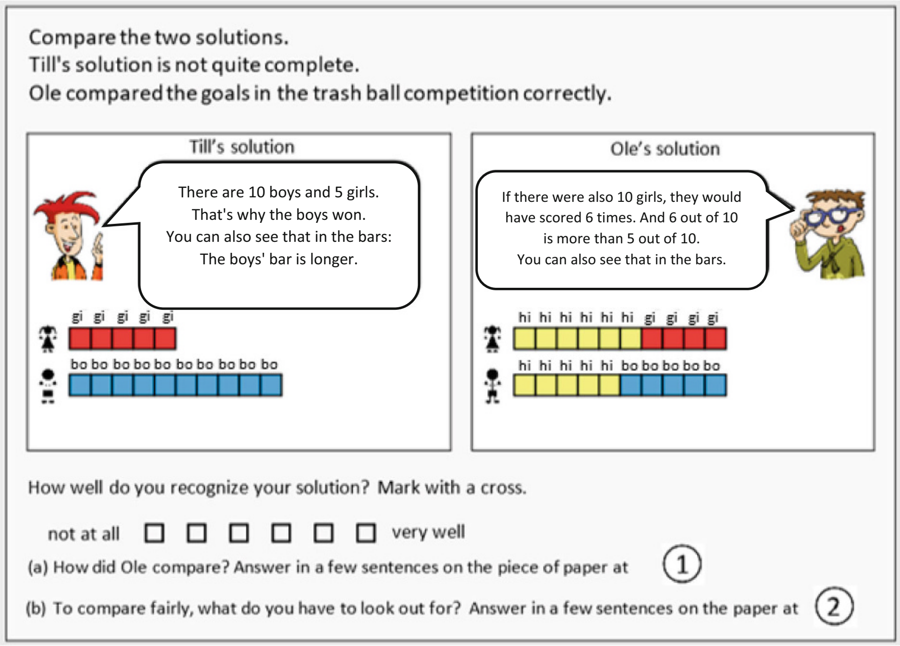


Fig. 6 Comparison of solutions in the instruction phase based on the solution from Fig. 4 (contra-adaptive condition, translated)

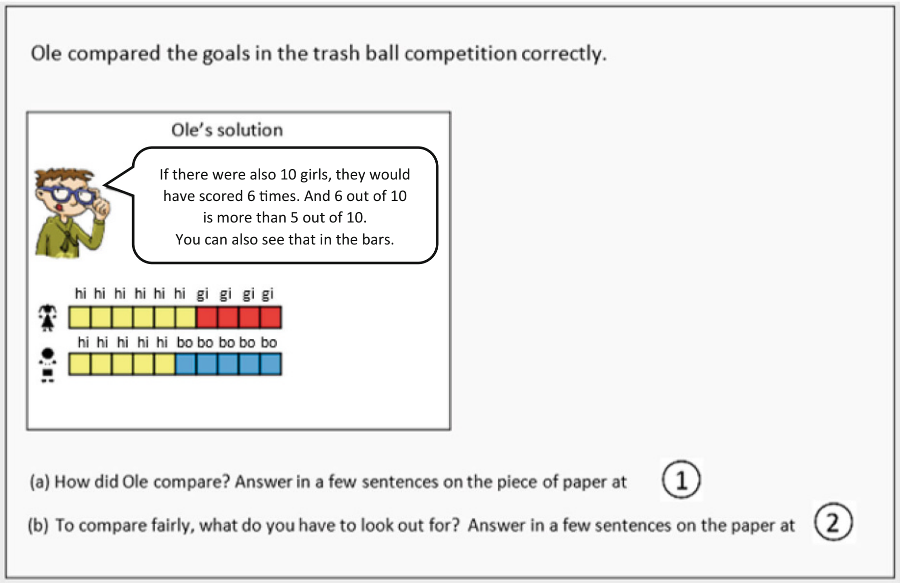


Fig. 7 General instruction in the control condition (translated)

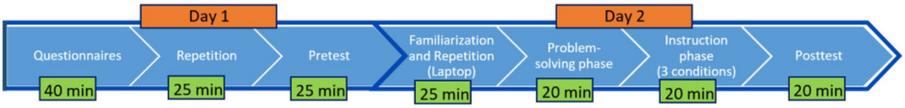


Fig. 8 Procedure of the study

cation technology competencies (ICT-competencies: 1. Computer use for school-related tasks, 2. Personally perceived ICT-competence, Mang et al. 2019, p. 116–128) were assessed as control variables. A repetition of the concept of fractional numbers in class intended to guarantee a basic understanding of symbolic and graphical representations of fractions, as this is a prerequisite for the comparison task. The learners broke $\frac{3}{5}$ of a real bar of chocolate. They marked on paper $\frac{3}{4}$ of a round cake, and they were asked to justify why in different square cakes of the same size, a quarter is still left, even though the parts look different. The solutions were discussed in class. Afterwards students filled in the prior knowledge test.

The second lesson started with the work in the digital learning environment (Boomgaarden et al. 2023). The learners became familiar with the handling in the computer-based learning environment and worked on tasks to repeat part-whole situations. Subsequently, the learners worked on the problem-solving task in the context of a trash ball competition. This was followed by the explicit instruction phase in three conditions. All students were randomly assigned to one of these conditions. During the explicit instruction phase, learners in the adaptive and contra-adaptive conditions had to indicate on a six-point scale how well they recognized their own solution from the problem-solving phase (1: I do not recognize my own solution at all, 6: I recognize my own solution completely). Finally, a posttest assessed the learning outcome.

4.5 Measures

In the following, the conception of the pretest and posttest is presented. The prior knowledge test consisted of three tasks. All tasks used the context of a trash ball competition. Task 1 and 2 assessed whether the learners had the prerequisites to be able to complete the problem-solving task (understanding of the context and the question) and produced ceiling effects (as intended). Task 3 was isomorphic to the task in the problem-solving phase (see description above) with different numbers. Thus, this third task covered the target concept (Table 1, level 3) and was therefore exclusively included in the data analyses as pretest for prior knowledge on the target concept (Table 1, level 3). The posttest consisted of 4 tasks within the trash ball context. Like the pretest, not all tasks covered the target concept (Table 1, level 3). The first two tasks reproduced the expected ceiling effects. Only two tasks (termed 'Posttest 1' and 'Posttest 2') were isomorphic to the task in the problem-solving phase and are therefore included in the analyses. With Cronbach's $\alpha = 0.93$, the internal consistency of the posttest is satisfying.

4.6 Analyses

To control whether the random assignment to the conditions was successful, we tested for differences between the groups in prior knowledge, text comprehension, understanding of arithmetic operations, and information and communication technology competencies (ICT-competencies).

For the further analysis, all cases are excluded that already indicated the target level in the pretest (Table 1, level 3; 10.4%) and thus could no longer achieve a learning increase through the instruction. In addition, those cases were excluded that argued non-mathematically in the problem-solving phase, since adaptive or contra-adaptive instruction was not feasible here, which was, in fact, the central intention of this study (17,7%). Therefore, the following analyses refer to a sample of $N = 206$.

In a first step, the student solutions from the pretest and the posttest (Posttest 1 and Posttest 2) were coded (Table 1). All solutions could be assigned to the codes (see Table 1). Thus, no further codes had to be developed inductively. The next step was to observe how frequently each level occurred within each condition. In a third step, the levels were scored (non-mathematical strategies and level 0: 0 pts., level 1: 1 pt., level 2: 2 pts., level 3: 3 pts.) in order to compare the mean values of the conditions descriptively.

To test hypothesis 1, that learners who compared an incorrect and a correct solution (i.e., learners in the adaptive and contra-adaptive conditions) achieve higher learning gains than their counterparts who studied a correct solution only (i.e., control condition) and hypothesis 2, that learners in the adaptive condition achieve higher learning gains than their counterparts in the contra-adaptive condition, we conducted a mixed ANOVA and pairwise comparisons.

In order to explore the learning outcomes of learners in the different conditions at a more fine-grained level the cases were divided according to the level of conceptual understanding at which the learners started in the pretest. Furthermore, the learning outcomes of learners who appropriately recognized their own solution type in the instruction in the adaptive condition are compared with those learners in the adaptive condition who indicated that they did not recognize their solution type. Similarly, the learning outcomes of learners in the contra-adaptive condition who appropriately did not recognize their individual solution type in the instruction are compared with those learners in the contra-adaptive condition who reported recognizing their solution type despite being in the contra-adaptive condition.

5 Results

5.1 Prior Knowledge

The data presented in Table 3 indicates that the learners ($N = 288$) in the three conditions did not differ regarding prior knowledge about comparing fractions, their text comprehension, their operation comprehension or their ICT-competencies (1. Computer use for school-related tasks, 2. Personally perceived ICT-competence). We

Table 3 Overview of the control variables and prior knowledge. The values in brackets refer to the sample included in the final analyses

	Adaptive condition $n=95$ ($n=67$)		Contra-adaptive condition $n=98$ ($n=67$)		Control condition $n=95$ ($n=72$)				
	M	*SD*	*M*	*SD*	*M*	*SD*	$F(2,285)$ $F(2,203)$	*p*	η^2
Prior knowledge test (Pretest)	1.09 (0.84)	1.05 (0.83)	0.92 (0.78)	1.08 (0.92)	1.20 (1.07)	0.96 (0.85)	1.83 (2.29)	0.16 (0.10)	0.01 (0.02)
Text comprehension	18.08 (18.22)	2.26 (2.04)	17.67 (17.66)	2.75 (2.63)	17.95 (17.65)	2.51 (2.68)	0.67 (1.20)	0.51 (0.30)	0.01 (0.01)
Operation comprehension	5.39 (5.28)	2.15 (1.99)	5.43 (5.28)	2.29 (2.21)	5.57 (5.53)	2.45 (2.47)	0.16 (0.28)	0.85 (0.76)	0.00 (0.00)
Computer use for school-related tasks	1.90 (1.84)	0.66 (0.68)	2.03 (2.06)	0.76 (0.77)	2.03 (2.00)	0.83 (0.79)	0.96 (1.96)	0.38 (0.19)	0.01 (0.02)
Personally perceived competence	2.08 (2.09)	0.58 (0.56)	2.13 (2.13)	0.65 (0.64)	2.13 (2.11)	0.66 (0.67)	0.22 (0.08)	0.80 (0.93)	0.00 (0.00)

therefore conclude that the randomized assignment of learners to conditions was successful.

5.2 Generated Solution Types Per Test and Condition

Table 4 displays the absolute frequency of learners at each level of the development of the part-whole concept. A distinction is made between the tests of the study and the conditions. The descriptive values in Table 4 reveal that—as expected—most students show a basic or an advanced natural number bias (level 0 and 1, see Table 1) at the pretest. A smaller, but still substantial part of the students show an emerging part-whole concept (level 2) at the pretest. Fewer learners already argued in the pretest according to the target concept (basal part-whole concept, level 3). These values reveal the heterogeneity of the learners when entering our learning environment. When comparing the pretest and the posttest, the number of students with a basic or advanced naturally number bias is substantially lower and the number of students at the target level is substantially higher at posttest than at pretest. The abstract level 4, which is above the target concept, was not reached by any learner. Overall, the descriptive values suggest a learning effect in all conditions. However, it seems that the conditions do not differ substantially.

For the further (inferential) analyses of this descriptive tendencies, all students were excluded that already showed the target level in the pretest (Table 1, level 3) and those students that argued non-mathematically in the problem-solving phase.

Table 5 displays the average scores per condition for the pre- and posttest. The descriptive values in Table 5 show learning gains from pre- to posttest for all con-

Table 4 Absolute frequencies of the solution types in the respective tests and conditions

Levels of understanding fractions	Conditions	Pretest	Posttest 1	Posttest 2
Non-mathematical concept (0 points)	a	9	2	3
	ca	6	2	2
	c	4	2	3
Basic natural number concept: Only Focus on the numerator (0 points)	a	27	13	16
	ca	45	17	17
	c	24	8	8
Advanced natural number concept: Only focus on the denominator (1 point)	a	25	17	19
	ca	13	20	25
	c	27	19	13
Advanced natural number concept: Focus on the rest (1 point)	a	1	1	1
	ca	1	2	4
	c	1	3	3
Emerging part-whole concept: Focus on numerator and denominator (2 points)	a	21	26	19
	ca	23	25	20
	c	31	28	34
Basal part-whole concept: Consider whole as equal (3 points)	a	12	36	37
	ca	10	32	30
	c	8	35	34

a adaptive condition, *ca* contra-adaptive condition, *c* control condition

Table 5 Scores of the solution types in all three conditions and tests in comparison

	Pretest		Posttest 1		Posttest 2		Posttest (mean)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Adaptive condition ($n=67$)	0.84	0.83	1.78	1.09	1.72	1.15	1.75	1.09
Contra-adaptive condition ($n=67$)	0.78	0.90	1.61	1.10	1.52	1.08	1.57	1.04
Control condition ($n=72$)	1.07	0.85	1.86	1.00	1.90	1.00	1.88	0.97

ditions. While students in the adaptive condition show the highest increase of about 0.9 points from pretest to mean posttest (in comparison to 0.8 in the other two conditions), the difference between conditions is very small.

5.3 Learning Outcomes in the Different Conditions

As an implementation check, we tested whether learners recognized their solution in the instruction phase. Learners in the adaptive condition recognized their own solution very well in the example solution during the instruction ($M=4.44$, $SD=1.46$). Learners in the contra-adaptive condition recognized their solution type less ($M=3.03$, $SD=1.81$). There is a significant difference between recognition in the adaptive and the contra-adaptive condition, $U=2610.50$, $p<0.01$. This indicates that it was successful to create an adaptive and a contra-adaptive learning environment, which is a requirement for the further analyses.

To test hypothesis 1, that learners in the adaptive and contra-adaptive condition achieve higher learning gains than their counterparts in the control condition and hypothesis 2, that learners in the adaptive condition achieve higher learning gains than their counterparts in the contra-adaptive condition, we conducted a mixed ANOVA and pairwise comparisons.

The Levene's test is not significant for the pretest ($F(2,203)=1.858$, $p=0.159$) or the posttest ($F(2,203)=1.479$, $p=0.230$). Thus, variance homogeneity can be assumed. There was homogeneity of covariance, as assessed by Box's test ($p=0.422$). Therefore, Wilks-Lambda is reported.

The mixed ANOVA revealed a significant main effect of the time ($F(1,203)=105.61$, $p<0.01$, $\eta_p^2=0.34$). However, the ANOVA revealed no significant main effect of condition ($F(2,203)=2.93$, $p=0.06$, $\eta_p^2=0.03$). There was no significant interaction effect of time and condition ($F(2,203)=0.20$, $p=0.82$, $\eta_p^2=0.002$). Thus, all conditions benefited equally from the instruction.

5.4 Exploratory Analyses of Prior Knowledge and Recognition of Solution

Overall, the hypotheses cannot be confirmed. However, the descriptive data in Table 4 indicated that learners may benefit differently from the learning opportunities in the different conditions, depending on their prior knowledge. In the following, learners are divided according to the level of conceptual understanding at which they start in the pretest to investigate whether the differences formulated in the hypotheses occur in relation to different levels of prior knowledge.

For learners who have already paid attention to both—the goals and the number of children in a group (i.e., level 2 in Table 1)—in the pretest, learners in the adaptive condition show the highest learning outcomes in the posttest (Fig. 9). Learners in the contra-adaptive condition perform the worst. Their performance in the posttest is even slightly worse than in the pretest. Statistically, the difference across conditions in the posttest performance is significant ($F(2,64) = 4.93$, $p = 0.01$, $\eta_p^2 = 0.134$). Bonferroni-corrected post-hoc analysis reveals a significant difference between the adaptive and the contra-adaptive condition ($p < 0.01$), but not the adaptive condition and the control condition ($p = 0.56$) and the control condition and the contra-adaptive condition ($p = 0.13$).

Learners who argued based on the absolute number of scores (Table 1, level 0) or on the absolute number of children per group (Table 1, level 1a) in the pretest show learning gains from pretest to posttest (level 0: $F(1,70) = 110.95$, $p < 0.001$, $\eta_p^2 = 0.61$; level 1a: $F(1,45) = 48.74$, $p < 0.001$, $\eta_p^2 = 0.52$). However, for both groups, there are no differences in the learning outcomes between the conditions (level 0: $F(2,70) = 0.06$, $p = 0.94$, level 1a: $F(2,45) = 0.8$, $p = 0.46$).

Thus, only differences between the conditions for learners who already argued based on an emerging part-whole concept (level 2) in the pretest are revealed. As this is only a small subsample, the overall results for the whole sample reported above only showed a small descriptive tendency but failed to reach statistical significance.

Furthermore, there is a descriptive difference in the learning outcomes depending on how well learners recognize their own solution during the instruction phase (adaptive condition: Fig. 10; contra-adaptive condition: Fig. 11). This assessment was made on a six-point scale (1: I do not recognize my own solution at all, 6: I recognize my own solution completely), with two learners ticking nothing.

Learners in the adaptive condition show no difference in learning outcome, whether or not they appropriately recognize their individual solution type in the

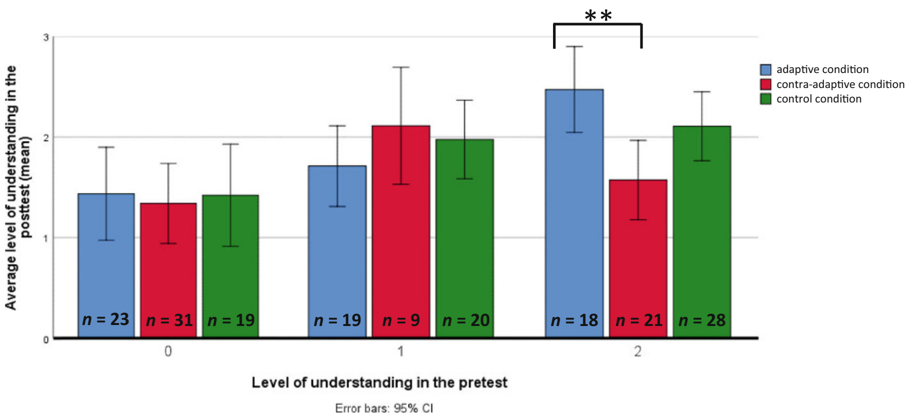


Fig. 9 Average level of conceptual understanding achieved in the posttest (mean) depending on the level of understanding indicated in the pretest

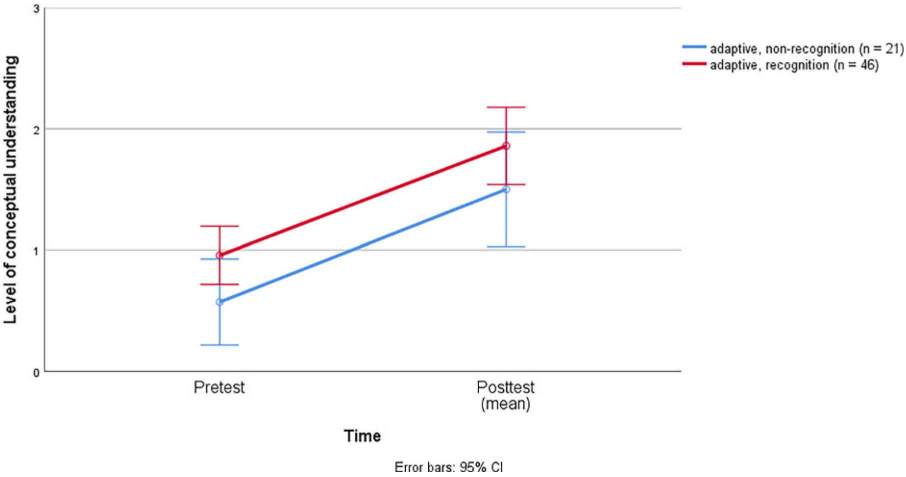


Fig. 10 Learning profiles of learners in the adaptive condition who correctly recognized their own solution type (six-point scale: 4–6) compared to learners in the adaptive condition who nevertheless did not recognize their solution type in the instruction (six-point scale: 1–3)

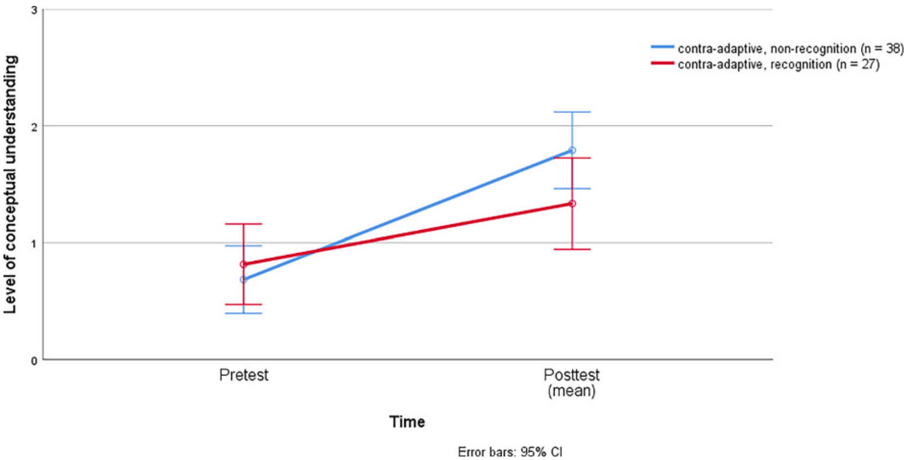


Fig. 11 Learning profiles of learners in the contra-adaptive condition who correctly did not recognize their own solution type (six-point scale: 1–3) compared to learners in the contra-adaptive condition who nevertheless recognized their solution type in the instruction (six-point scale: 4–6)

instruction phase. In contrast, appropriate non-recognition in the contra-adaptive condition seems to have a positive influence on learning success. However, this difference fails to reach statistical significance (contra-adaptive condition: $t(63) = 1.79$, $p = 0.08$).

6 Discussion

The research presented here is based on instructional approaches with an initial problem-solving phase followed by an explicit instruction phase (PS-I, Loibl et al. 2017). The student solutions from the problem-solving phase only represent partial steps on the way to the target concept. For the subsequent instruction, the findings on conceptual change and learning from errors suggest beneficial effects of comparing incorrect solutions with the correct solution. The results of Loibl and Leuders (2019) indicate that an adaptive fit between the individual solution type generated by the learners in the problem-solving phase and the type of the erroneous solution in the instruction phase may be relevant for learning success. The present study investigated the effect of the adaptive fit systematically by comparing three variants of a computer-based learning environment (called adaptive, contra-adaptive, and control condition).

We hypothesized a) that the comparison of incorrect solutions with the correct solution (adaptive and contra-adaptive condition) leads to greater conceptual learning outcomes than the exclusive consideration of the correct solution (control condition), and b) that the comparison of one's own solution type from the problem-solving phase with the correct solution (adaptive condition) leads to greater conceptual learning outcomes than the comparison of another solution type to the correct solution (contra-adaptive condition). In addition, we exploratively analysed how a) learners' prior level of conceptual understanding and b) learners' recognition of their individual solutions in the instruction phase influence the effects of conditions on learning outcomes.

In the following, the results of the present study are discussed with regard to the three underlying perspectives: (1) instructional design, (2) learning processes, and (3) knowledge. We include the discussion of limitations with regard to each of these perspectives.

6.1 Instructional Perspective: Review of the Design of the Study

While we did find a learning effect in all conditions (suggesting that the transposition to a digital learning environment was adequate), we did not find the expected interaction between time and condition. Thus, it could not be replicated that the learners of the adaptive and contra-adaptive condition learned more effectively than the learners of the control condition (in contrast to Loibl and Leuders 2019). It could also not be confirmed that the learners of the adaptive condition achieve a higher learning success than the learners of the contra-adaptive condition.

This unexpected result could, of course, indicate that the theoretical assumptions on error processing in PS-I do not hold: Comparisons and adaptivity of the instruction with respect to processing erroneous solutions may be less relevant for learning fractions than expected. However, there may be other, more specific features of the realization of the instructional approach, which are responsible for the lack of influence of the conditions.

6.1.1 Limitations and Implications for Future Research Regarding the Instructional Perspective

In order to investigate individually and instantaneous adaptive instruction, a computer-based learning environment was used. One potential reason for the missing effect of condition might be the fact that working on the laptop had a stronger influence on the learning success than the conditions. For example, it was not feasible to fix the time spent for comparing solutions during the instruction phase, because they were able to click on “Next” independently. In the study by Loibl and Leuders (2019), the instruction phase was of equal length for each learner and tended to be longer than in the present study. As time on task usually is a relevant predictor for learning outcomes (Karweit and Slavin 1981), this feature of the instructional design may have affected the effects.

Moreover, working in a computer-based learning environment in a highly controlled setting implies that learners do not interact with each other or with the teacher. This limits external validity, as most studies in the PS-I context have been implemented in collaborative settings (e.g., Kapur 2012; Kapur and Bielaczyc 2012; Loibl and Rummel 2014a). However, other studies have successfully implemented PS-I in an individual setting (e.g., Kapur 2014) and previous studies have not yet provided evidence for beneficial effects of collaborative learning over individual learning in PS-I (Brand et al. 2023; Mazziotti et al. 2015, 2019).

To avoid that working on the computer obscures the effects of the conditions, a further study could benefit from closely examining the materials used in the prior study by Loibl and Leuders (2019) and implement adaptivity without a computer-based learning environment.

6.2 Learning Perspective: Comparison Processes and Recognition of Solution Type as Prerequisites for Learning

The finding that only learners who already argued based on an emerging part-whole concept (level 2) in the pretest benefitted from an adaptive comparison of solutions in the explicit instruction phase match the findings by Große and Renkl (2007) showing that only learners with high prior knowledge benefit from comparing correct and incorrect worked examples as only these learners engage in elaborative comparisons. Learners with lower prior knowledge may need more support in order to draw relevant comparisons (cf. Loibl and Leuders 2019).

Comparing correct and incorrect solutions may initiate a conceptual change, if the incorrect solution corresponds to the own erroneous mental model (Gadgil et al. 2012). Although the learners in the control condition only received the correct solution in the instruction phase, they still may have cognitively recalled their own (incorrect) solution due to the proximity in time[3]. Thus, they could have compared their own incorrect solution to the correct solution mentally which would initiate a conceptual change. Similarly, learners in the contra-adaptive condition also may

[3] Note that in contrast to our study in previous PS-I studies, the instruction phase often does take place a couple of days later (e.g., Kapur 2010; Loibl and Leuders 2018).

have integrated their own solution into the comparison process mentally. If students in the control condition and the contra-adaptive condition mentally included their own solution in the comparison, the intended differences in the learning processes across conditions are reduced.

The explorative results provide some hints on the learning process. Learners in the adaptive condition learn equally well regardless of whether they recognize their solution type in the instruction phase or not. This may be explained as follows: In the adaptive condition, the correct solution is always compared to the individual solution type—even if students do not recognize their solution. According to the underlying assumptions, comparing the individual solution type to the correct solution facilitates conceptual change (Gadgil et al. 2012).

In contrast, learners in the contra-adaptive condition tend to learn better when they appropriately recognize that the presented erroneous solution does not correspond to their own solution in comparison to learners who inappropriately state that they recognize their own solution. This tendency may be explained as follows: Learners who correctly realize that their solution differed from the presented solution, can obviously remember their own solution, and thus are able to mentally integrate their own solution from the problem-solving phase into the instruction phase. By comparing their remembered solution with the presented solutions, they can revise their flawed mental model similarly as students in the adaptive condition.

While these explanations of the exploratory findings fit the theoretical assumptions, they still need to be investigated in further experimentally controlled study, e.g., by capturing and integrating more process data on the comparison processes into the analysis.

6.2.1 Limitations and Implications for Future Research Regarding the Learning Perspective

Although the study was carefully designed to investigate the envisioned learning mechanisms, the empirical analysis has shown that some features of the study design must be reconsidered.

One of these limitations is the fact that recalling the own solution may have influenced the learning processes. Future studies could reduce this recall effect, for example by increasing the time span between the two phases or by including a distractor task.

Unexpectedly, and in contrast with the intended implementation, some learners in the contra-adaptive condition indicated that they recognized their solutions. This "paradox" recognition could be caused by a focus on surface features (color, rectangular shapes, bars; Boomgaarden et al. 2023). Future studies should strive to reduce this focus on surface features, e.g., by changing the item design.

Furthermore and conversely, some learners in the adaptive condition did not recognize their solution type. Presenting a screenshot of the individual solution could potentially improve the veracity of the students' comparison and thus the adaptivity.

6.3 Knowledge Perspective: Prior Knowledge as Prerequisite for Learning Effects

The descriptive results indicate that learning effects may differ depending on the initial conceptual level of understanding as measured in the pretest. These interindividual differences may have contributed to the missing effects of conditions on learning outcomes. Our explorative results suggest that the learning success for this learning topic is in fact very dependent on the prior knowledge of the learners.

Learners who already focus on the two components—number of goals and number of children—in the pretest benefit from adaptive instruction (consistent with hypothesis H2). This result was statistically significant in our sample. However, due to the exploratory nature of our analyses, it needs to be tested again with a new sample. Nevertheless, this finding fits to the theoretical assumptions on the learning process: It is reasonable that learners who already indicated an emerging part-whole concept in the pretest benefit most in the adaptive condition, since both components—number of goals and number of children—are also addressed in both solutions (the correct and the incorrect one) provided in the explicit instruction phase.

Learners in the contra-adaptive condition with an emerging part-whole-concept, on the other hand, again focus on a basic natural number concept in the instruction phase because the erroneous solution focuses on the number of goals (numerator) only. Since the natural number concept is very intuitive and persistent (Ni and Zhou 2005; Vamvakoussi and Vosniadou 2010), these learners may regress to a basic natural number concept in the posttest. This finding also underpins the persistence of the natural number bias, which in this case could be a barrier for learning in the contra-adaptive condition.

Furthermore, the theoretical assumptions on conceptual change suggest that a certain level of prior knowledge is required to successfully recognize knowledge gaps, which is the basis for conceptual change (Vosniadou 2007). Similarly, in his work on ‘unproductive failure’, Kapur (2016) argues that learners need a certain level of prior knowledge to benefit from problem solving and a subsequent comparison of solutions (PS-I).

In summary, the result suggests that a specific level of prior knowledge (i.e., an emerging part-whole concept) is required in order to achieve a conceptual change by comparing the individual solution with the correct solution in the instruction phase (adaptive condition).

6.3.1 Limitations and Implications for Future Research Regarding the Knowledge Perspective

At this point we would like to note that 65 learners did not have a full data set as they did not have parental consent ($n = 18$) or missed at least one test day ($n = 47$). These learners were excluded from the analysis. Given that the missings were equally distributed across conditions and the still adequate sample size in our study, we refrained from imputing missing values as it did not seem crucial for the results.

Nevertheless, we need to acknowledge the limitation that we cannot guarantee that these missing were completely at random.

The interpretation regarding the persistence of the natural number bias and the prerequisites for conceptual change highlight that the results of the present study are limited to the part-whole concept, thus the results are not broadly generalizable.

More importantly, as already stated, the findings discussed in this section are based on exploratory analyses and, thus, need to be tested with a new sample.

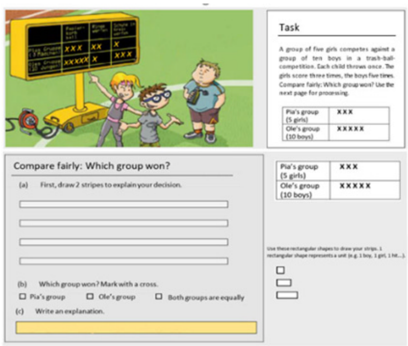
7 Conclusion

PS-I-learning scenarios reveal beneficial learning effects for the development of conceptual understanding (Kapur and Bielaczyc 2012; Loibl and Rummel 2014a, b) and transfer (e.g., Belenky and Nokes-Malach 2012; Schwartz et al. 2011). Despite the substantial evidence of the effectiveness of this learning approach, the cognitive learning mechanisms have not been sufficiently investigated so far. In particular, the recognition and processing of knowledge gaps appears to be of great relevance for learning success (Loibl et al. 2017). Although the overall results of the present study could not confirm the underlying assumptions in a computer-based learning environment, the results provided evidence that learning effects depend on students' prior knowledge, that is, students only benefit from adaptive instruction when they already have a certain level of prior knowledge. Overall, the results presented here can be considered an important step on the way to investigate the learning mechanisms of PS-I and they highlight the need for focussing even more on the learning processes in the two phases in future research (cf. Loibl et al. 2024).

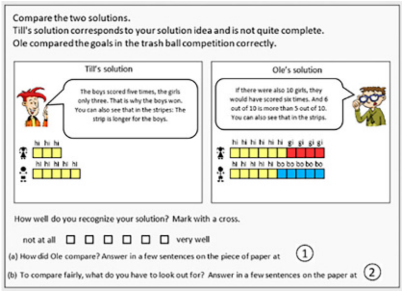
8 Appendix

Table 6 Overview of the timeline of the study

<table>
<tr><td>1. Pretest</td><td>25 min</td></tr>
<tr><td>2. Familiarization with the handling in the computer-based learning environment</td><td>15 min</td></tr>
<tr><td>3. Repetition of part-whole situations</td><td>10 min</td></tr>
<tr><td>4. Problem-solving phase</td><td>20 min



cf. Fig. 4</td></tr>
<tr><td>5. Instruction phase</td><td>20 min



There were three versions of the instruction: cf. Fig. 5 for the adaptive version
cf. Fig. 6 for the contra-adaptive version
cf. Fig. 7 for the version of the control condition</td></tr>
<tr><td>6. Posttest</td><td>20 min</td></tr>
</table>

Acknowledgements This work was supported by the German Research Foundation (DFG) under Grant No 456976731.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest A. Boomgaarden, K. Loibl and T. Leuders declare that they have no competing interests.

Ethical standards The study was implemented in alignment with the principles of ethical and professional conduct of the German Psychological Society (DGPs). The data was collected and saved anonymously. Informed consent was obtained from all individual participants included in the study.

References

Asterhan, C.S., & Dotan, A. (2018). Feedback that corrects and contrasts students' erroneous solutions with expert ones improves expository instruction for conceptual change. *Instructional Science*, *46*, 337–355.

Barbieri, C.A., & Devlin, B.L. (2024). Targeting fraction misconceptions and reducing high confidence errors in an online tutor. *Journal of Computer Assisted Learning*, *40*(1), 254–268.

Belenky, D.M., & Nokes-Malach, T.J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences*, *21*(3), 399–432. https://doi.org/10.1080/10508406.2011.651232.

Biggs, J., & Collis, K. (1989). Towards a model of school-based curriculum development and assessment using the SOLO taxonomy. *Australian journal of education*, *33*(2), 151–163.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects—State, trends and issues in mathematics instruction. *Educational studies in mathematics*, *22*(1), 37–68.

Boomgaarden, A., Loibl, K., & Leuders, T. (2019). Computer-based Problem Solving to Prepare for Adaptive Consolidation. In Lund, K., Niccolai, G. P., Lavoué, E., Hmelo-Silver, C., Gweon, G., & Baker, M. (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019* (Vol. 2, pp. 863–864). Lyon, France: International Society of the Learning Sciences.

Boomgaarden, A., Loibl, K., & Leuders, T. (2023). The trade-off between complexity and accuracy. Preparing for computer-based adaptive instruction on fractions. *Interactive Learning Environments*, 31(10), 6379–6394.

Booth, J.L., Lange, K.E., Koedinger, K.R., & Newton, K.J. (2013). Using example problems to improve student learning in algebra: differentiating between correct and incorrect examples. *Learning and Instruction*, *25*, 24–34. https://doi.org/10.1016/j.learninstruc.2012.11.002.

Brand, C., Hartmann, C., Loibl, K., & Rummel, N. (2023). Do students learn more from failing alone or in groups? Insights into the effects of collaborative versus individual problem solving in productive failure. *Instructional Science*. https://doi.org/10.1007/s11251-023-09619-7.

Catrambone, R., & Holyoak, K.J. (1989). Overcoming contextual limitations on problemsolving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1147–1156.

Charalambous, C. Y., & Pitta-Pantazi, D. (2005). Revisiting a theoretical model on fractions: Implications for teaching and research. *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*, *2*(2), 233–240.
Chi, M. T. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in instructional psychology*, *5*, 161–238.
Corral, D., & Carpenter, S. K. (2020). Facilitating transfer through incorrect examples and explanatory feedback. *Quarterly Journal of Experimental Psychology*, *73*(9), 1340–1359. https://doi.org/10.1177/1747021820909454.
DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of experimental child psychology*, *113*(4), 552–568. https://doi.org/10.1016/j.jecp.2012.06.009.
Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, *22*(3), 206–214. https://doi.org/10.1016/j.learninstruc.2011.11.001.
Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.
Freudenthal, H. (1983). *Didactical Phenomenology of mathematical structures*. Dodrecht: Kluwer.
Freudenthal, H. (1991). *Revisiting mathematics education*. China Lectures. Dodrecht: Kluwer.
Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction*, *22*(1), 47–61. https://doi.org/10.1016/j.learninstruc.2011.06.002.
Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of educational psychology*, *95*(2), 393.
Gould, P. (2005). Year 6 students' methods of comparing the size of fractions. In *Proceedings of the annual conference of the mathematics education research group of Australasia* (pp. 393–400). Sydney: MERGA.
Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and instruction*, *17*(6), 612–634. https://doi.org/10.1016/j.learninstruc.2007.09.008.
Heemsoth, T., & Heinze, A. (2014). The impact of incorrect examples on learning fractions: a field experiment with 6th grade students. *Instructional Science*, *42*(4), 639–657. https://doi.org/10.1007/s11251-014-9313-x.
Heemsoth, T., & Heinze, A. (2016). Secondary school students learning from reflections on the rationale behind self-made errors: a field experiment. *The Journal of Experimental Education*, *84*(1), 98–118. https://doi.org/10.1080/00220973.2014.963215.
Hoyos, V. (2016). Distance technologies and the teaching and learning of mathematics in the era of MOOC. In M. Niess, S. Driskell & K. Hollebrands (Eds.), *Handbook of research on transforming mathematics teacher education in the digital age* (pp. 137–164).
Kaiser, G. (2017). The teaching and learning of mathematical modeling. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 267–291). Reston: National Council of Teachers of Mathematics.
Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional science*, *38*(6), 523–550. https://doi.org/10.1007/s11251-009-9093-x.
Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*(4), 651–672. https://doi.org/10.1007/s11251-012-9209-6.
Kapur, M. (2014). Productive failure in learning math. *Cognitive Science*, *38*(5), 1008–1022. https://doi.org/10.1111/cogs.12107.
Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, *51*(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457.
Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, *21*(1), 45–83. https://doi.org/10.1080/10508406.2011.591717.
Karweit, N., & Slavin, R. E. (1981). Measurement and modeling choices in studies of time and learning. *American Educational Research Journal*, *18*(2), 157–171.
Kieren, T. E. (1976). On the mathematical, cognitive, and instructional foundations of rational numbers. In R. A. Lesh (Ed.), *Number and measurement* (pp. 101–150). Columbus: Eric/SMEAC.

Koedinger, K.R., Corbett, A.T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. *Cognitive science*, *36*(5), 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x.

Lamon, S.J. (2007). Rational numbers and proportional reasoning: toward a theoretical framework for research. In *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 629–667).

Lenhard, W., Lenhard, A., & Schneider, W. (2009). Diagnose und Förderung von Leseverständnis mit ELFE 1–6 und ELFE-Training. In W. Lenhard & W. Schneider (Eds.), *Diagnose und Förderung des Leseverständnisses* (pp. 79–112). Göttingen: Hogrefe.

Liljedahl, P., & Cai, J. (2021). Empirical research on problem solving and problem posing: a look at the state of the art. *ZDM–Mathematics Education*, *53*(4), 723–735.

Loibl, K., & Leuders, T. (2018). Errors during exploration and consolidation—The effectiveness of productive failure as sequentially guided discovery learning. *Journal für Mathematik-Didaktik*, *39*(1), 69–96. https://doi.org/10.1007/s13138-018-0130-7.

Loibl, K., & Leuders, T. (2019). How to make failure productive: fostering learning from errors through elaboration prompts. *Learning and Instruction*, *62*, 1–10. https://doi.org/10.1016/j.learninstruc.2019.03.002.

Loibl, K., & Rummel, N. (2014a). Knowing what you don't know makes failure productive. *Learning and Instruction*, *34*, 74–85. https://doi.org/10.1016/j.learninstruc.2014.08.004.

Loibl, K., & Rummel, N. (2014b). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science*, *42*(3), 305–326. https://doi.org/10.1007/s11251-013-9282-5.

Loibl, K., Roll, I., & Rummel, N. (2017). Towards a theory of when and how problem solving followed by instruction supports learning. *Educational Psychology Review*, *29*(4), 693–715. https://doi.org/10.1007/s10648-016-9379-x.

Loibl, K., Leuders, T., Glogger-Frey, I., & Rummel, N. (2024). CID: a framework for the cognitive analysis of composite instructional designs. *Instructional Science*, 1–25.

Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A., & Reiss, K. (2019). *PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

Mazziotti, C., Loibl, K., & Rummel, N. (2015). Collaborative or individual learning within productive failure. Does the social form of learning make a difference? In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine & S. Ludvigsen (Eds.), *Exploring the material conditions of learning: the computer supported collaborative learning (CSCL) conference 2015* (Vol. 2, pp. 570–575). Gothenburg: ISLS.

Mazziotti, C., Rummel, N., Deiglmayr, A., & Loibl, K. (2019). Probing boundary conditions of productive failure: does the productive failure effect transfer to young students, and what is the role of collaboration? *npj Science of Learning*, *4*, 1–9.

McNamara, J., & Shaughnessy, M.M. (2010). *Beyond pizzas & pies: 10 essential strategies for supporting fraction sense, grades 3–5*. Math Solutions.

Newstead, K., & Murray, H. (1998). Young students' constructions of fractions. In A. Olivier & K. Newstead (Eds.), *Proceedings of the twenty-second international conference for the psychology of mathematics education*. Stellenbosch. (pp. 295–302).

Ni, Y., & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: the origins and implications of whole number bias. *Educational Psychologist*, *40*(1), 27–52. https://doi.org/10.1207/s15326985ep4001_3.

Olive, J., & Steffe, L.P. (1994). *TIMA Bars*. Acton: William K. Bradford. Computer software

Padberg, F. (2002). Anschauliche Vorerfahrungen zum Bruchzahlbegriff zu Beginn der Klasse 6. *Praxis der Mathematik in der Schule*, *44*(3), 112–117.

Padberg, F., & Bienert, T. (2000). Didaktik der Bruchrechnung-Zur Entwicklung des Bruchzahlverstandnisses und der Rechenoperationen mit gemeinen Brüchen innerhalb eines Schuljahres. *Mathematikunterricht*, *46*(2), 24–37.

Pillai, R.M., Loehr, A.M., Yeo, D.J., Hong, M.K., & Fazio, L.K. (2020). Are there costs to using incorrect worked examples in mathematics education? *Journal of Applied Research in Memory and Cognition*, *9*(4), 519–531. https://doi.org/10.1016/j.jarmac.2020.06.007.

Pitkethly, A., & Hunting, R. (1996). A review of recent research in the area of initial fraction concepts. *Educational Studies in Mathematics*, *30*(1), 5–38. https://doi.org/10.1007/bf00163751.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: toward a theory of conceptual change. *Science education*, *66*(2), 211–227.

Prediger, S. (2008). The relevance of didactic categories for analysing obstacles in conceptual change: revisiting the case of multiplication of fractions. *Learning and instruction*, *18*(1), 3–17. https://doi.org/10.1016/j.learninstruc.2006.08.001.

Prediger, S. (2011). Vorstellungsentwicklungsprozesse initiieren und untersuchen. *Der Mathematikunterricht*, *57*(3), 5–14.
Prediger, S., Barzel, B., Hußmann, S., & Leuders, T. (2021). Towards a research base for textbooks as teacher support: the case of engaging students in active knowledge organization in the KOSIMA project. *ZDM–Mathematics Education*. https://doi.org/10.1007/s11858-021-01245-2.
Richter-Gebert, J., & Kortenkamp, U.H. (2012). *The Cinderella. 2 manual: working with the interactive geometry software*. Springer.
Rittle-Johnson, B., & Star, J.R. (2011). The power of comparison in learning and instruction: learning outcomes supported by different types of comparisons. In *Psychology of learning and motivation* (Vol. 55, pp. 199–225). Academic Press.
Schukajlow, S., Blomberg, J., Rellensmann, J., & Leopold, C. (2022). The role of strategy-based motivation in mathematical problem solving: the case of learner-generated drawings. *Learning and Instruction*, *80*, 101561.
Schulz, A., Leuders, T., & Rangel, U. (2020). The use of a diagnostic competence model about children's operation sense for criterion-referenced individual feedback in a large-scale formative assessment. *Journal of Psychoeducational Assessment*, *38*(4), 426–444. https://doi.org/10.1177/0734282918823590.
Schwartz, D.L., Chase, C.C., Oppezzo, M.A., & Chin, D.B. (2011). Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759–775. https://doi.org/10.1037/a0025140.
Siegler, R.S. (2002). Microgenetic studies of self-explanation. In N. Garnott & J. Parziale (Eds.), *Microdevelopment: transition processes in development and learning* (pp. 31–58). Cambridge: Cambridge University Press.
Siegler, R.S., & Chen, Z. (2008). Differentiation and integration: Guiding principles for analyzing cognitive change. *Developmental Science*, *11*(4), 433–448. https://doi.org/10.1111/j.1467-7687.2008.00689.x.
Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: evidence for productive failure. *Review of Educational Research*, *91*(5), 761–798. https://doi.org/10.3102/00346543211019105.
Stafylidou, S., & Vosniadou, S. (2004). Students' understanding of the numerical value of fractions: a conceptual change approach. *Learning and Instruction*, *14*(5), 503–518. https://doi.org/10.1016/j.learninstruc.2004.06.015.
Tsovaltzi, D., McLaren, B.M., Melis, E., & Meyer, A.K. (2012). Erroneous examples: effects on learning fractions in a web-based setting. *International Journal of Technology Enhanced Learning*, *4*(3-4), 191–230.
Vamvakoussi, X., & Vosniadou, S. (2002). Conceptual change in mathematics: from the set of natural to the set of rational numbers. In S. Lehti & K. Merenluoto (Eds.), *Proceedings of the Third Symposium on Conceptual Change*. Turku. (pp. 201–204).
Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and instruction*, *28*(2), 181–209. https://doi.org/10.1080/07370001003676603.
Van den Heuvel-Panhuizen, M., & Drijvers, P. (2020). Realistic mathematics education. In *Encyclopedia of mathematics education* (pp. 713–717). https://doi.org/10.1007/978-3-030-15789-0_170.
Van Hoof, J., Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2017). The transition from natural to rational number knowledge. In *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 101–123). https://doi.org/10.1111/bjep.12134.
VanLehn, K. (1999). Rule-learning events in the acquisition of a complex skill: an evaluation of cascade. *The Journal of the Learning Sciences*, *8*(1), 71–125.
VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. (2003). Why does only some events cause learning during human tutoring? *Cognition and Instruction*, *21*(3), 209–249. https://doi.org/10.1207/S1532690XCI2103_01.
Van Peppen, L.M., Verkoeijen, P.P.J.L., Heijltjes, A.E.G., Janssen, E.M., & Van Gog, T. (2021). Enhancing students' critical thinking skills: Is comparing correct and erroneous examples beneficial? *Instructional Science*, *49*(6), 747–777. https://doi.org/10.1007/s11251-021-09559-0.
Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: a survey. *ZDM*, *52*, 1–16.
Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and instruction*, *4*(1), 45–69.

Vosniadou, S. (2007). The cognitive-situative divide and the problem of conceptual change. *Educational Psychologist*, *42*(1), 55–66. https://doi.org/10.1080/00461520709336918.

Vosniadou, S., & Verschaffel, L. (2004). The conceptual change approach to mathematics learning and teaching. *Learning and instruction*, *14*(5), 445–548.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.