Learning and Instruction 65 (2020) 101264

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Learning fractions with and without educational technology: What matters for high-achieving and low-achieving students?

Frank Reinhold[a,*], Stefan Hoch[a], Bernhard Werner[b], Jürgen Richter-Gebert[b], Kristina Reiss[a]

[a] Technical University of Munich, TUM School of Education, Heinz Nixdorf-Chair of Mathematics Education, Germany
[b] Technical University of Munich, Department of Mathematics, Chair for Geometry and Visualization, Germany

ABSTRACT

Instructional design research promotes interactive and adaptive scaffolds as features of educational technology. Mathematics education research can guide elaborated fractions curricula to develop basic fraction concepts while challenging the natural number bias. Thus, we developed theory-grounded interactive material for learning fractions providing scaffolds in an eBook. Evaluating both, curriculum and scaffolds, we split 745 high-achieving and 260 low-achieving 6th graders into three groups: Scaffolded Curriculum group (using the eBook on iPads), Curriculum group (using a paper copy of our developed material), and Traditional group (using conventional textbooks). Generalized linear mixed models revealed diverse positive effects on the achievement of students in the experimental conditions: Results showed that high-achieving students did benefit from the curriculum, regardless of whether it was presented with or without scaffolds, while for low-achieving students using scaffolds was decisive. This suggests that interactive and adaptive scaffolds can support students in learning mathematical concepts, especially for low-achieving students.

1. Introduction

Digital media—i.e., features that educational technology can provide—can be beneficial for learning, as recent meta-analyses support (Hillmayr, Ziernwald, Reinhold, Hofer, & Reiss, submitted for publication; Arroyo et al., 2014; Cheung & Slavin, 2013; Ma, Adesope, Nesbit, & Liu, 2014; Steenbergen-Hu & Cooper, 2014). Here, a commonly agreed upon argument within the media debate—initiated by R.E. Clark (1994) and Kozma (1994)—is that it is not the mere medium that does have an effect on learning outcomes, but rather the appropriate way of implementing it into the classroom as well as certain features that technology enhanced learning environments can offer. In detail, adaptivity, feedback, and the use of hands-on activities seem to be promising for the development of suitable interactive material (Alibali & Nathan, 2012; A. Clark, 1999; Mayer, 2014; Moreno, 2004; Moreno, Reisslein, & Delgoda, 2006; Wilson, 2002). All the mentioned possible educational benefits can be realized in interactive and digital learning environments on tablet PCs. But it seems still unclear whether such interactive learning environments developed with regards to well-established design principles—that have proven to be beneficial for learning within several short-term experimental studies—can also work within real classroom situations (e.g., Kucirkova, 2014; see also; de Jong, 2010) and whether they are beneficial for teaching mathematical concepts in school contexts.

For creating well-designed interactive learning environments we suggest that the development should rely on three pillars that will be introduced here and will be illustrated in detail within the following sections: (1) knowledge about educational aspects regarding the *content*—here grounded on research from mathematics education; (2) implications from psychological theories regarding the *instructional design* of multimedia learning environments; and (3) the *technological implementation* of interactive aspects and adaptive scaffolds integrated in interactive and digital learning environments that have shown to be beneficial for the acquisition of new concepts. Therefore, we introduce a framework for the development of digital learning environments for mathematics education, composed of the three mentioned pillars *Content*, *Instructional Design*, and *Technological Implementation*. This framework allows for designing learning environments with regard to evidence-based practices from mathematics education, educational psychology, and instructional design. Following the rationale of the suggested framework, effects of content design can be disentangled from effects regarding its implementation in intervention studies, as displayed throughout this article. For the purpose of our studies, the framework—as shown in Fig. 1—is applied for basic fraction concepts using interactive and adaptive scaffolds within an electronic textbook on tablet PCs (commonly referred to as *tablets*, e.g., iPads).

Within such development of interactive learning environments, content should play a distinctive role (Pepin, Choppin, Ruthven, &

* Corresponding author. Technical University of Munich, TUM School of Education, Arcisstraße 21, 80333, Munich, Germany.
E-mail addresses: frank.reinhold@tum.de (F. Reinhold), stefan.hoch@tum.de (S. Hoch), werner@ma.tum.de (B. Werner), richter@ma.tum.de (J. Richter-Gebert), kristina.reiss@tum.de (K. Reiss).

https://doi.org/10.1016/j.learninstruc.2019.101264
Received 1 February 2019; Received in revised form 2 September 2019; Accepted 22 September 2019
Available online 10 October 2019

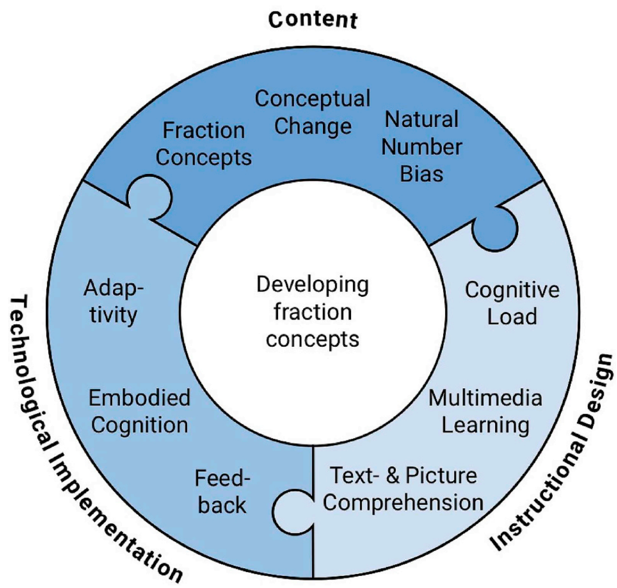

Fig. 1. Framework for developing digital and interactive learning environments for mathematics education; relying on three pillars *Content*, *Instructional Design*, and *Technological Implementation*—exemplarily shown for basic fraction concepts.

Sinclair, 2017), as many previous studies on the use of digital media in the classroom lacked software explicitly developed to support the curriculum (Kucirkova, 2014). In this article, we focus on the understanding of fraction concepts as a key facet of mathematical literacy (Bailey, Hoard, Nugent, & Geary, 2012; Booth & Newton, 2012; Lortie-Forgues, Tian, & Siegler, 2015).

We present an approach on teaching basic fraction concepts with an interactive learning environment, following the suggested framework: The development of this learning environment was based on insights from research in mathematics education, psychology and instructional design research, summarized in the first part of the article. In the second part we evaluate our learning environment with data from two studies with 745 high-achieving students (study 1) and 260 low-achieving students (study 2).

1.1. Challenges of establishing basic fraction concepts

There is empirical evidence that understanding fractions is predictive of later achievement in higher mathematics such as algebra (Bailey et al., 2012). Unfortunately, many students struggle with learning fraction concepts as well as with understanding higher mathematics (Lortie-Forgues et al., 2015), making fractions possible "gatekeepers" of higher mathematics (Booth & Newton, 2012).

One serious source of errors in handling fractions is the *Natural Number Bias* (i.e., NNB). The NNB refers to a robust tendency to use concepts of natural numbers (e.g., the counting scheme) to interpret rational numbers and fractions (Ni & Zhou, 2005). Students can be affected by a natural number bias when they do not master the necessary *conceptual change* from natural numbers to fractions (Meert, Grégoire, & Noël, 2010; Obersteiner, Van Hoof, Verschaffel, & Van Dooren, 2015; Vamvakoussi & Vosniadou, 2004; Van Hoof, Degrande, Ceulemans, Verschaffel, & Van Dooren, 2018). Here, typically four dimensions of necessary conceptual changes are denoted regarding *representations*, *density*, *size* and *operations* (e.g., Obersteiner et al., 2015).

When shifting from natural numbers to rational numbers, the concept of numbers loses the uniqueness of a symbolic *representation*. While there is no other symbolic way to denote the number 2 within the set of natural numbers, there are arbitrary many ways to denote the fraction 3/4 within the set of rational numbers, e.g., $3/4 = 6/8 = 9/12$. Even natural numbers lose their unique representation within the set of rational numbers. Research suggests that even older students struggle with rational numbers having more than one unique symbolic representation (DeWolf, Bassok, & Holyoak, 2015).

While each natural number has one and only one number succeeding it—making counting of numbers possible—the concept of a unique successor must be dropped when shifting to fractions, e.g., there is a countably infinite set of (rational) numbers between the two fractions 5/8 and 6/8. The idea of a discrete set of numbers can hinder students when operating with fractions, resulting in difficulties when developing the concept of *density* of rational numbers (Vamvakoussi & Vosniadou, 2004).

Furthermore, there is empirical evidence that students tend to compare the *size* of fractions by comparing their numerators or denominators—as natural numbers—separately (Meert et al., 2010). Retaining a concept of *size* based on natural numbers can indeed lead to correct answers in specific fraction comparison tasks, e.g., $3/7 < 5/7$. Yet, a conceptual change seems necessary to solve more advanced problems, e.g., $8/9 < 7/6$ although $8 > 7$ and $9 > 6$. Research suggests that developing a concept of the size of fractions can be a first step towards fulfilling a conceptual change of number concepts (Van Hoof et al., 2018).

At last, shifting from natural numbers to fractions alters the effect that arithmetic *operations* have on numbers (Obersteiner et al., 2015). One example is the frequent and common misconception of "multiplication makes bigger" which is true for natural numbers but loses its generality within the set of rational numbers (Prediger, 2008)—e.g., $1/2 \cdot 3/4 = 3/8$ which is smaller than both 1/2 and 3/4.

In this paper, we focus on the first three dimensions of *representations*, *density*, and *size*—which can be argued to be of interest in developing basic fraction concepts (see also Section 2.2.1). Regarding these difficulties, learning environments for teaching fractions should not focus only on building up procedural knowledge of fractional arithmetic. Rather, they should focus on the development of conceptual knowledge of fractions and directly address the necessity of altering the concept of numbers (Kainulainen, McMullen, & Lehtinen, 2017). Therefore, we suggest that building up elaborated *concepts of rational numbers* (e.g., the part-whole concept and the concept of fraction magnitude; see Behr, Lesh, Post, & Silver, 1983; Meert et al., 2010) are one appropriate way to introduce fractions in the classroom—and to support students in overcoming an NNB. We follow the idea of Behr et al. (1983) and suggest that in addition to symbolic representations of fractions, iconic depictions—e.g., circle or tape diagrams—may be beneficial for teaching and learning these concepts.

We illustrate this for the concept of expanding and simplifying of fractions, exemplarily. The arithmetic operation of expanding 3/4 with 2 leading to the fraction 6/8 can be considered mere procedural, as no conceptual understanding of fractions seems necessary to follow the rule "multiply both the numerator and the denominator with 2". It is plausible to assume that such arithmetic tasks will not support students in changing the concept of *representation* of numbers: since the operation is based on symbolic representations of fractions, students that are already affected by an NNB—i.e., hold on to the uniqueness of a number's symbolic representation—will probably not develop a conceptual understanding of expanding fractions. We suggest that using iconic representations to introduce expanding and simplifying of fractions as "refining and coarsening of a given division" (e.g., Padberg & Wartha, 2017) first can support a conceptual change, *and*—in addition—can motivate the necessary arithmetic operations: when the whole is divided into twice as much pieces, twice as much pieces are needed to result in the same part of the whole.

While most studies on students' struggle with an NNB focus on short-term measurements of students' misconceptions, a few studies shed light on the specific learning trajectories followed when developing fraction concepts (Kainulainen et al., 2017; McMullen, Laakkonen, Hannula-Sormunen, & Lehtinen, 2015; Van Hoof et al., 2018)—agreeing in two central findings: (1) learning paths taken by students during the acquisition of fraction knowledge are in line with conceptual change theory, and (2) overcoming an NNB is a gradual and slow process which not every student masters despite traditional fractions instruction in school. In more detail, Van Hoof et al. (2018) could show that once students reach an appropriate level of knowledge about operations with fractions, most of them do not develop a deeper understanding of the structure of fractions—challenging traditional approaches of teaching rational numbers with a focus on arithmetic (Vamvakoussi, Christou, Mertens, & Van Dooren, 2011). In their longitudinal study, Kainulainen et al. (2017) could not find shifts from natural-number-biased to mathematically correct concepts; however, gradual shifting from wrong conceptualizations of fractions to more—yet, still not completely—correct conceptualizations with less focus on natural number thinking were indeed present. This suggests that overcoming an NNB should not be seen as an immediate learning goal, but as a rather slow—yet, radical—conceptual change that needs specific guidance within classroom practice. For such conceptual change, McMullen et al. (2015) suggest that grasping the concept of fraction magnitude is necessary—yet, not sufficient—to develop elaborated concepts of fractions, e.g., density. This supports teaching approaches that do not treat fractions as two separate natural numbers, but a fraction as one holistic symbol consisting of a numerator and a denominator.

With regards to both short-term and longitudinal studies, we assume that approaches using iconic representations both to foster conceptual understanding of fractions and to illustrate the semantics of arithmetic procedures within mathematics classrooms can support students in overcoming an NNB—albeit slowly and gradually.

1.2. Designing multimedia learning environments

According to *cognitive load theory*, human working memory is assumed to be limited and therefore effective learning may be hindered when learning environments overload the cognitive capacity of learners (Paas & Sweller, 2014; Sweller, 2010; Sweller, Ayres, & Kalyuga, 2011). Here, cognitive load theory distinguishes between intrinsic and extraneous load—both drawing on the cognitive capacity during learning activities. *Intrinsic load* can be understood as a measure of the fixed complexity of a specific content: content with high element interactivity—i.e., content that cannot be learned in isolation but only when presented within the context of other information, as in our case fraction concepts—is considered to evoke high intrinsic load that cannot be altered by changing the learning environment or the modes of instruction (Sweller, 2010). In contrast, *extraneous load* is additional complexity generated due to the use of specific learning environments: it is assumed that students need to use additional working memory resources when dealing with inappropriately designed learning environments (Paas & Sweller, 2014), which we think is of specific importance for designing *interactive* learning environments. Therefore, cognitive load theory suggests that it is most important to have good instructional design when students need to learn content with high element interactivity—in order to minimize extraneous load given the unalterable high intrinsic load (Sweller et al., 2011).

Focusing on the topic of our studies, developing elaborated concepts of fractions can be considered a learning task with high element interactivity since students do not only have to learn new content but have to radically change their concept of numbers (see Section 1.1). Thus, learning environments for teaching basic fraction concepts should fulfil certain criteria that can—based on empirical evidence—be expected to lower extraneous load. These criteria include, e.g., the use of completion problems (problem completion effect, see Paas, 1992), the combined presentation of interacting elements that have to be learned in cohesion (split-attention effect, see Tarmizi & Sweller, 1988), and avoiding redundant information (redundancy effect, see Sweller et al., 2011). Here, one central assumption in line with the results from different studies is that one and the same learning environment can yield both necessary and redundant information for students with low and high prior knowledge, respectively (expertise reversal effect, see Kalyuga, 2007). It seems reasonable that this can lead to different effects for those two groups of students, motivating adaptive learning scenarios (see Section 1.3).

Following the *cognitive theory of multimedia learning*, words and pictures are thought to be processed within different cognitive structures with separate capacities (Mayer, 2014). Hence, presenting new-to-learn content both as texts and pictures has shown to be more effective than using texts only (multimedia principle; Butcher, 2014). Moreover, texts and pictures informing the same content should be presented close together to accomplish the biggest support of learning (spatial contiguity principle; Mayer & Fiorella, 2014). For the case of fractions, the multimedia principle can be realized when developing material where students need to shift between symbolic and iconic representations of fractions during their development of basic fraction concepts, which refers to general principles regarding different representations—e.g., iconic and symbolic—in mathematics education (e.g., Behr et al., 1983; Bruner, 1960; Duval, 2006; Lesh, 1981; Padberg & Wartha, 2017).

Beyond that, according to the *integrated model of text and picture comprehension*, processing texts and pictures related to the same content can lead to different mental representations, i.e., propositional representations or mental models (Schnotz, 2014; Schnotz & Bannert, 2003). In contrast to cognitive theory of multimedia learning, pictures are not generally thought to be beneficial for learning but only when their inherent structure is analogous to the mental representation that is intended to be built up. In the case of fractions, this can be illustrated comparing the inherent structure of the number line and a circle diagram, which are both iconic representations of fractions. Since many students struggle with identifying the whole on the number line (e.g., Novillis-Larson, 1980), it seems more appropriate to teach the part-whole concept with regards to a circle diagram where the whole can easily be identified, as the circle diagram's inherent structure represents exactly one whole while the number line may be arbitrarily large. In contrast, it seems plausible to compare the size of two fractions using the number line, as its inherent structure allows for the comparison of two numbers in one iconic representation.

Cognitive load theory, cognitive theory of multimedia learning, and the integrated model of text and picture comprehension offer general guidelines for high-quality instructional design that are commonly agreed upon in the field of social sciences and mathematics education. However, when referring to empirical evidence for the effects of cognitive load theory, cognitive theory of multimedia learning, and the integrated model of text and picture comprehension, one should bear in mind that such evidence is usually based on rather short experimental situations (i.e., instructional *message* design; see Mayer, 2014; van Merriënboer & Kester, 2014). Therefore, the relevance for longitudinal classroom interventions (i.e., instructional *curriculum* design; see van Merriënboer & Kester, 2014) may be challenged, as it seems questionable whether time constrains typically used in experimental studies appear in realistic situations, or how students' offloading due to taking notes in classroom situations affects the validity of those effects (de Jong, 2010). Moreover, one can think of additional differences between such experimental situations and real classroom practice that may have an effect, such as, the presence of a teacher, an authentic learning environment, and unobtrusive research during classroom education. Thus, the question can be asked whether following these guidelines can be beneficial for teaching fraction concepts within real school contexts.

1.3. Potential benefits from interactive learning environments

With digital media in educational contexts emerging as technology develops further and is more easily accessible, the media debate initiated by R.E. Clark (1994) and Kozma (1994) more than 25 years ago seems still important when investigating effects of educational technology. In line with Kozma (1994), we argue that educational technology can provide certain features which offer potential benefits for classroom practice, as cognitive load in complex learning scenarios may be lowered using appropriately designed and well-integrated interactive and adaptive learning environments on touchscreen devices (e.g., Moreno, 2004; Moreno et al., 2006; Wilson, 2002). In this article, we focus on three aspects of *technological implementation* that have been shown to be beneficial for learning: embodiment, adaptivity and feedback.

The underlying idea of the *embodied cognition theory* can be traced back to Piaget's theory of developmental psychology, “which emphasized the emergence of cognitive abilities out of a groundwork of sensorimotor abilities” (Wilson, 2002, p. 625). We follow the idea of a “simple” embodied cognition theory (A. Clark, 1999), which is considered to be in line with a classical theoretical framework of cognitive science: embodiment is expected to affect the inner organization and processing of knowledge (e.g., Ballard, Hayhoe, Pook, & Rao, 1997). Within this theory, one aspect that promotes the use of touchscreen devices is that gestures congruent to the context are expected to be beneficial for the development of elaborated concepts—especially in the field of mathematics (Alibali & Nathan, 2012). For example, cutting through pizza in terms of swiping over a touchscreen can be considered a congruent gesture (cf. Fig. 3b later in this article), while cutting through pizza by clicking and pointing with a mouse on a computer can be considered an incongruent gesture. In fact, Black, Segal, Vitale, and Fadjo (2012) could show that first and second graders obtained better results in addition tasks after learning the rules of addition on an iPad compared to students who worked with the same software on a computer. They argue that the difference in the learning outcome might be justified by how students handle the two different digital devices: Students operate iPads by using natural gestures while the input via a mouse is more or less artificial. This argumentation follows the claim that cognitive demands can be offloaded onto the environment (Wilson, 2002): referring to the given example of cutting pizza, we argue that cutting and sharing pizza using hands-on activities on touchscreen devices can result in lower cognitive load during the acquisition of the part-whole concept than using a mouse on a computer, since the congruent gestures used on touchscreen devices are analogous to the mental processes of partitioning and distributing, while clicking and pointing with a mouse are artificial processes. Regarding this, we see potential benefit in using such interactive learning environments designed (and developed) to make explicit use of touchscreen technology, as depicted within the given example of “sharing pizza” given above.

Adaptivity of digital learning environments can be described as an automatic variation of parameters of the learning environment—e.g., task difficulty—based on students' responses while working within the environment—e.g., correct or false answers—(Leutner, 2004). It is argued that an adaptive adjustment of task difficulties can lower cognitive load (e.g., Moreno et al., 2006). For the case of fractions, where conceptual changes seem necessary, adaptive task difficulty might play a crucial role: if tasks are too easy, students might lack the necessity to overcome their concepts of natural numbers—especially when tasks are solvable by using natural number concepts. If tasks are too hard, they might not illustrate new-to-learn concepts in an intelligible and plausible way. Here, adaptive task difficulty can be used to match students' individual needs—resulting in better outcomes when students work with adaptive learning environments than with traditional instruction (Hillmayr et al., submitted for publication; Arroyo et al., 2014; Cheung & Slavin, 2013; Ma et al., 2014; Steenbergen-Hu & Cooper, 2014), and accounting for a potential expertise reversal effect (Kalyuga, 2007).

Feedback can be conceptualized as external feedback given to a student by a teacher or a digital learning environment as a direct reaction to his or her solution of a specific task—i.e., task-level feedback (Hattie & Timperley, 2007). Such feedback can be corrective—i.e., merely stating “correct” or “incorrect”—or explanatory—i.e., giving explanations for possible misconceptions relying on the individual answer (Hattie & Timperley, 2007). Research suggests that individual explanatory feedback can support students in closing subject-related gaps, in overcoming misconceptions, and in developing elaborated concepts (Hattie & Timperley, 2007). We follow the idea of Moreno (2004) that such feedback can additionally lower cognitive load. Here, fractions are a suitable topic for building interactive learning environments with individual feedback, since problems and common errors are well known and documented over the last decades (e.g., Behr, Wachsmuth, & Post, 1985; Obersteiner et al., 2015; Stafylidou & Vosniadou, 2004). Moreover, both individual feedback and adaptivity are key features of intelligent tutoring systems, which have proven to have an overall moderate to high positive effect on students' learning, as reported by recent meta-analyses (Hillmayr et al., submitted for publication; Cheung & Slavin, 2013; Ma et al., 2014; Steenbergen-Hu & Cooper, 2014).

1.4. The present studies

One serious issue of many studies on the use of features of educational technology in schools is the lack of software explicitly developed to support the curriculum (Kucirkova, 2014). Here, digital curriculum resources (Pepin et al., 2017)—such as electronic textbooks—offer one way to overcome this essential shortcoming. In line with these authors, we introduced a framework for the development of interactive learning environments relying on three pillars: content, instructional design, and the technological implementation of interactive aspects and scaffolds that are known to be beneficial for learning. This framework was initially given in Fig. 1 and theoretically founded within an interdisciplinary context (i.e., mathematics education, educational psychology, instructional design) throughout Section 1.

Yet, making the content one main aspect in the development, educational research should make distinctions between effects of the design of the content (e.g., tasks and explanations of mathematical concepts) and effects regarding its technological implementation within digital learning environments (e.g., interactivity, adaptivity, and feedback) possible, as reliable empirical evidence whether students really can benefit from features that educational technology can provide seems crucial for decisions concerning media in education (R.E. Clark, 1994; Kozma, 1994; Kucirkova, 2014). This question seems of explicit importance, since several studies supporting the benefit of such features rely on rather short experimental designs (i.e., instructional message design), challenging their external—or ecological—validity (see Kucirkova, 2014 for iPads; and de Jong, 2010 for cognitive load theory).

With regard to the framework introduced in Fig. 1, we developed an electronic textbook on basic fraction concepts, integrating interactive and adaptive scaffolds for the use on iPads (eBook). This eBook draws on intuitive and perceptual abilities with non-symbolic representations of fractions and links them to symbolic fractions, following an elaborated fractions curriculum, as described in detail within Section 2.2.1. In order to distinguish between effects of the curriculum and effects of interactive and adaptive scaffolds—i.e., interactive explanatory tasks making specific use of touchscreen technology, adaptive task difficulty, and individual explanatory feedback implemented into the eBook—a paper-based textbook was created as a one-to-one copy with identical explanations and the same exercises. Relying on our framework (cf. Fig. 1), we ask:

1. Are interactive and adaptive scaffolds and/or an elaborated fractions curriculum beneficial for establishing fraction concepts during regular classroom instruction when compared to traditional classroom settings?
2. Do potential benefits of interactive and adaptive scaffolds and/or an elaborated fractions curriculum have different effects on the development of procedural and conceptual knowledge of fractions?
3. Do potential effects of interactive and adaptive scaffolds and/or an

elaborated fractions curriculum on the development of fraction concepts differ for low-achieving and high-achieving students?

We assume the curriculum to have a general positive effect on the development of conceptual knowledge of fractions (see Section 1.1), which may again inform procedures. As we do not focus on only one feature educational technology can provide within a rather short experiment, but on a combination of different features within a content-specific learning environment for the use in a rather longitudinal intervention (i.e., real classroom scenarios), assessing the effectiveness of the scaffolds within school contexts seems a valid question. Here, we assume the scaffolds to be more beneficial for low-achieving than for high-achieving students, as they may deal with more misconceptions resulting in the need for more individualized learning paths and more interactive depictions to support them in developing fraction concepts—which may be phrased as an expansion of the expertise reversal effect (Kalyuga, 2007) on interactive and adaptive scaffolds.

2. Method

Both classroom intervention studies were conducted as cluster randomized controlled trials with classrooms as clusters. They followed the same pre-post-control, constructive research design with two treatment groups and one control group learning basic fraction concepts: The *Scaffolded Curriculum group* (abbreviated in tables and figures as *Scaffolded Curriculum*) worked with our newly-developed interactive environment on iPads. We created this digital learning environment as a interactive textbook, which will be described in detail below (see Section 2.2). Students from the Scaffolded Curriculum group were able to benefit both from the developed material and the use of interactive and adaptive scaffolds—as features of educational technology—during the intervention. The *Curriculum group* (abbreviated in tables and figures as *Curriculum*) worked with the same material as a paper-based book. Hence, students from the Curriculum group received the same guidance and were therefore able to benefit from the developed material but not from the use of educational technology features. In order to control for effects of this material, the *Traditional group* (abbreviated in tables and figures as *Traditional*) used conventional textbooks, where the teachers were free to choose how to achieve the administered educational objectives (see Section 2.4 for a detailed description).

Therefore, the three intervention groups differed first in whether they were exposed to our newly-designed elaborated fractions curriculum, i.e., a holistic material to learn basic fraction concepts, which was developed taking account of both insights into students' learning of fractions from a mathematics education perspective and guidelines to design learning environments from an instructional psychology perspective. They also differed in whether they worked with interactive and adaptive scaffolds in a digital learning environment, which was able to give feedback to students' answers and made use of hands-on activities to foster fraction concepts. This constructive research design is summarized in Table 1.

2.1. Sample

In total, $N = 1005$ sixth grade students from German (Bavarian) public schools took part in our two studies. To investigate effects on both high-achieving students and low-achieving students, we included sixth-graders from different school tracks from the three-track public school system in Bavaria, Germany. Study 1 included 745 (46% female) high-achieving students from German *Gymnasium* (i.e., the highest school track with students showing above the average grades in grade4). Study 2 included 260 (42% female) low-achieving students from German *Hauptschule* or *Mittelschule* (i.e., the lowest school track with students showing below the average grades in grade 4). In general, differences in mathematical achievement in standardized tests between students from these two school tracks are very high: e.g., Götz, Lingel, and Schneider (2013) report differences at the end of grade 6 as high as $d = 2.68$ using the DEMAT 6+; and Sälzer, Reiss, Schiepe-Tiska, Prenzel, and Heinze (2013) find very large differences in the mathematical competency of fifteen-year-olds in PISA 2012 between those school tracks. Therefore, the a priori operationalization of high-achieving and low-achieving students made for the present studies seems meaningful. We will provide additional validation of this operationalization in terms of differences in pretest scores and posttest outcomes within the present studies in Section 3.

Both studies followed the same research design: Only after teachers agreed in taking part in the study, whole classes were assigned to one of the three intervention groups randomly, while taking care not to concentrate certain groups on specific schools. We were able to provide the equipment necessary for the Scaffolded Curriculum group (i.e., tablet PCs and management container for storing and charging the tablet PCs) for six classrooms with up to 33 students for each study. Regarding these restrictions, we first assigned classrooms to the Scaffolded Curriculum group randomly and split the remaining classrooms equally to the Curriculum group and the Traditional Group. The resulting number of students in each group for high-achieving and low-achieving students can be seen in Table 2, where sample sizes for the pretest and the posttest are reported separately, since not all students took part in both tests for different reasons. As this procedure led to an unbalanced design, we chose a statistical method of analysis that takes this into account, appropriately (see Section 2.6), and report post-hoc power analyses for each reported effect (see Section 3).

2.2. Material

The development of our learning environment followed the framework that is summarized in Fig. 1 and presented throughout Section 1.

2.2.1. Content

On the basis of the reported findings on how students learn and should learn fractions we created an *elaborated fractions curriculum* to teach basic fraction concepts (*Content* in Fig. 1). This curriculum covered (1) the part-whole concept with one and many wholes, using iconic, continuous, discrete as well as symbolic representations (Behr et al., 1983); (2) fraction magnitude as an intuitive understanding of the “size” of unique parts (Meert et al., 2010); (3) expanding and simplifying fractions exemplified as divisions becoming more refined or coarser (Lamon, 2012); (4) fractions on the number line (Kieren, 1976); (5) fractions representing more than one whole and mixed numbers; and (6) comparing fractions using a variety of strategies (Clarke & Roche, 2009). Thus, the four basic arithmetical operations—i.e., adding, subtracting, multiplying and dividing fractions—are not in the

Table 1
Overview of the constructive research design of both studies with graduated treatment in two treatment groups and one control group.

Intervention groups	Elaborated fractions curriculum: Students were exposed to newly designed holistic material to learn basic fraction concepts	Interactive and adaptive scaffolding: Students worked with an interactive and adaptive digital learning environment with feedback and hands-on activities
Traditional	No	No
Curriculum	Yes	No
Scaffolded Curriculum	Yes	Yes

Table 2
Number of students participating in the pretest and posttest during the intervention, given for both studies: high-achieving students (study 1 , left) and low-achieving students (study 2, right).

	Study 1: High-achieving students		Study 2: Low-achieving students	
Intervention groups	Pretest	Posttest	Pretest	Posttest
Scaffolded Curriculum	156	159	105	107
Curriculum	318	320	64	71
Traditional	263	266	82	82
Total	737	745	251	260

scope of the reported studies. A validity check of the elaborated fractions curriculum, performed a priori with regards to the mandatory curriculum for grade six in Bavaria, Germany, ensured that students within all three intervention groups were exposed to the same six content domains mentioned. Yet, our elaborated fractions curriculum focusses primarily on conceptual knowledge of fractions while traditional textbooks focus largely on procedural knowledge of fractions. We want to clarify this with three examples: Firstly, the elaborated fractions curriculum focusses on a conceptual understanding of expanding and reducing fractions as refining and coarsening of a given division (both in iconic and symbolic representation, see Fig. 2d for a widget), while the prevalent approaches present in traditional textbooks focus on the mere arithmetic procedure of multiplying or dividing enumerator and denominator by the same natural number—yielding no appropriate learning opportunities to overcome a NNB regarding *representation*, as argued in Section 1.1. Secondly, we exemplify the curricular differences with the content of comparing fractions: The elaborated fractions curriculum focusses on a variety of strategies that need to be chosen appropriately with regard to the given pair of fractions (Clarke & Roche, 2009), e.g., benchmarking to 1 when comparing 7/6 vs. 8/9 (Reinhold, Reiss, Hoch, Werner, & Richter-Gebert, 2018). In contrast, the prevalent approach present in traditional textbooks is finding a common denominator and then comparing the fractions referring solely to their numerators—which may yield no appropriate learning opportunities to overcome a NNB regarding *size*, as applying the common misconception "the fraction with the larger enumerator is the larger fraction" to the resulting pair of fractions after finding a common denominator results in correct solutions, yet may not introduce the necessary conceptual change.

Regarding the cited literature on teaching and learning fractions our elaborated fractions curriculum allowed students to explore non-symbolic fractions (e.g., circle diagrams depicted as pizzas, or tape diagrams) before introducing more formal representations (e.g., fractions on the number line; symbolical representation of fractions), as providing intuitive pathways to core fraction concepts (Behr et al., 1983) was our first key goal in the development of the material. In addition, we paid special attention to include teaching rationales and tasks that *do* yield appropriate learning opportunities to overcome an NNB, according to the literature in Section 1.1 and as depicted in the two examples given above. We close with a third example, concerning the task of challenging the dimension of *density* within the NNB: The task to represent a fraction in continuous circle or bar diagrams *does not* allow for counting strategies, but can support the concept of fraction magnitude (Meert et al., 2010)—and, therefore, may be beneficial for a shift from thinking about fractions as *discrete* numbers to an age-appropriate interpretation of *density* and *continuity* (i.e., "being not discrete"). Moreover, we assume that exposing students to representation tasks with both prepartitioned *and* continuous diagrams—i.e., relying on *and* suppressing counting strategies—can introduce cognitive conflict and therefore may elaborate their concept of *density* (Merenluoto & Lehtinen, 2004).

2.2.2. Instructional design

We designed our material with regard to the implications of cognitive load theory, cognitive theory of multimedia learning, and integrated model of text and picture comprehension for developing valuable learning environments (*Instructional Design* in Fig. 1). Here, we understand iconic representations of fractions as pictures and symbolic representations of fractions as text—thus assuming that these different representations of fractions can be processed in different cognitive structures. Therefore, we assume that teaching fractions using both non-symbolic and symbolic representations can lead to better learning outcomes than using solely symbolic representations (multimedia principle). Furthermore, we think of exercises that focus on the transitions between various non-symbolic and symbolic representations of fractions as fruitful tasks to develop conceptual knowledge of fractions (e.g., Behr et al., 1983). Yet, as such tasks yield high element interactivity resulting in high intrinsic cognitive load, we assume that they are of high cognitive demand for students—which ascribes importance to reducing extraneous cognitive load in the learning environment. Thus, lowering extraneous cognitive load was the second key goal in the development of our material.

2.2.3. Technological Implementation

We created our digital learning environment as an electronic textbook (eBook) for use on tablet PCs using iBooks Author (Apple Inc., 2017) for the framework and CindyJS (von Gagern, Kortenkamp, Richter-Gebert, & Strobel, 2016) as the programming environment for the interactive content (*Technological Implementation* in Fig. 1). This eBook (Hoch et al., 2018) allowed for hands-on activities and included a total of 90 interactive and adaptive exercises (i.e., *Widgets*). Referring to findings portrayed above, we assume that tangible hands-on activities and tutoring with adaptive task difficulty as well as individual feedback can reduce cognitive load in learning tasks—and therefore lower the high cognitive demand of learning basic fraction concepts—especially for low-achieving students.

As the learning environment was developed mainly for the use within real classroom scenarios, we designed widget-based *adaptivity*, since teachers would most frequently choose the widget to work with for the students. We implemented levels of increasing difficulty for each widget, based not on empirically measured task difficulties, but on evidence-based difficulty-generating characteristics of the specific task type, as illustrated for number line tasks: Variation in length (i.e., number line of length 1 or longer) and equivalence (i.e., unit segment is separated into segments according to the exact value of the denominator of the to-represent fraction, or not) can be seen as difficulty-generating characteristics for number line tasks (Novillis-Larson, 1980). Each widget generates sets of tasks within one level of difficulty, randomly. Students have to complete these sets before they are allowed to proceed to the next-higher difficulty level. Tasks answered incorrectly have to be repeated until a heuristically determined threshold of correct answers is reached. For a detailed description of the technical development of adaptivity within the eBook, see Hoch, Reinhold, Werner, Richter-Gebert, & Reiss (2018).

Feedback is designed and technologically implemented for each widget, based on student-specific or task-specific characteristics. We provide an example for both kinds of explanatory feedback: Firstly, in iconic representation tasks, feedback for an incorrect answer is given based on students' answers, e.g., the correct solution will be shown, and the fraction the student depicted will be named, in addition. Secondly, in comparison tasks, feedback for an incorrect answer is given based on an algorithm choosing an appropriate strategy, e.g., for 7/6 vs. 8/9 benchmarking to 1 would be suggested by the algorithm, while for 4/9 vs. 3/5, benchmarking to 1/2 would be suggested—regardless of the students' answer. For a detailed description of the technical development of feedback within the eBook, see Hoch et al. (2018).

Different aspects of interactivity are integrated within the eBook. One of them are *exploratory tasks* to get involved with basic fraction concepts. For example, students are asked to “distribute pizza” using congruent gestures—as described in Section 1.3—or “use sliders” to explore what happens to the iconic representation of a given fraction, when the enumerator or the denominator is changed, or the fraction is raised. A second aspect of interactivity integrated in certain widgets is self-regulated *graded assistance*, where students can choose to get constructive hints for solving the problem—referring to the problem completion effect (Paas, 1992).

To summarize: making use of the touchscreen technology, and implementing adaptivity and automatic individual feedback as *scaffolds* was the third key goal in the development of our material. The eBook was designed for landscape-use on 12.9” iPads. Textbook pages were shown in a two-column layout. Widgets were programmed to run in full-screen mode and could be started by touching them on a textbook page. To give a brief insight in the learning environment, examples are given in Fig. 2, where a typical two-column textbook page and three typical widgets (distributing pizza, part-whole concept in symbolical representation, and expanding fractions in iconic representation) are shown.

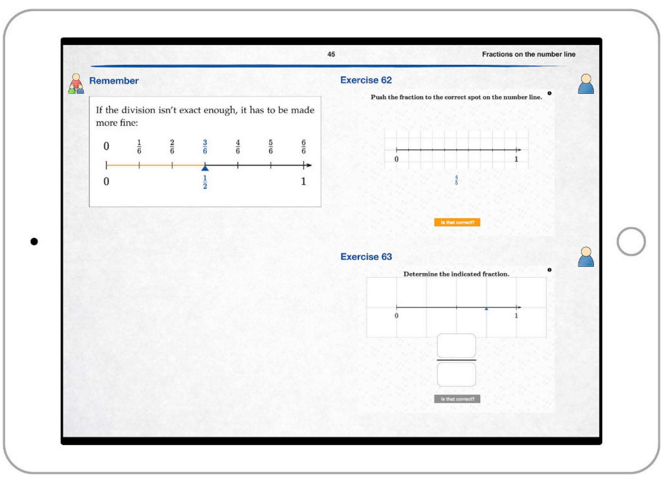

Fig. 2. Sample views of the interactive learning environment “ALICE:fractions” to teach basic fraction concepts for use on iPads (Hoch, Reinhold, Werner, Reiss, & Richter-Gebert, 2018).

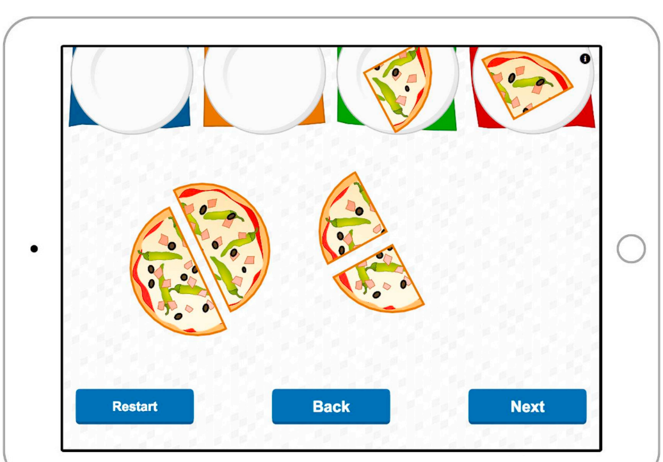

Fig. 2. (*continued*)

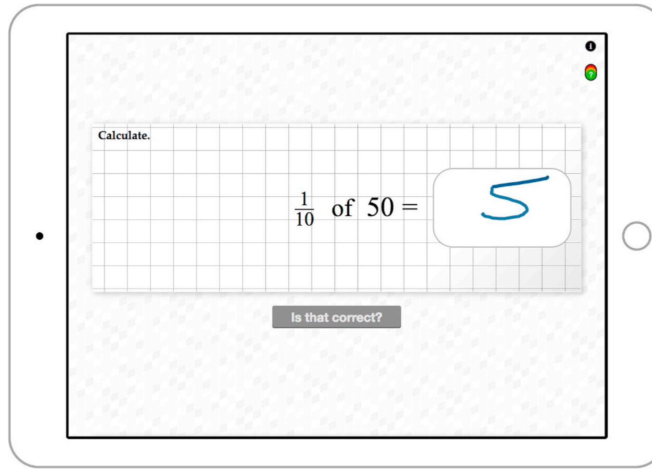

Fig. 2. (*continued*)

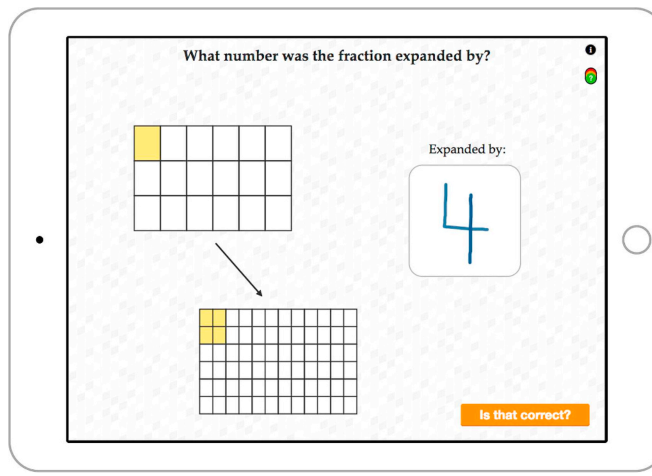

Fig. 2. (*continued*)

2.2.4. Paper-based version

After developing the eBook, a paper-based version of the learning environment was created as a near-to-exact copy of the digital textbook. For that, we printed 5–6 items from each of the 90 widgets into the two-column layout of the textbook and kept instructions and the sequence of the exercises identical: The content consists of short book texts as well as summarizations in formulas giving generic examples. Thus, we assumed that students working with the paper copy version could still benefit from the expertise we put into the content and the instructional design of our material (i.e, elaborated fractions curriculum), but would not be able to additionally benefit from the technological features implemented (i.e., interactive and adaptive scaffolds). The digital version is available in English (Hoch et al., 2018), the paper copy version is availabe in the German original (Hoch, Reinhold, Werner, Reiss, & Richter-Gebert, 2018).

2.3. Instruments

In order to control for effects of prior knowledge of fractions, a pretest (10 items, McDonald's $\omega = 0.827$, 95%CI [0.812, 0.843]; inter-rater reliability $0.885 \leq \kappa \leq 0.986$) was conducted before the intervention. Since the Bavarian curriculum covers fractions in grade six, students in both studies did not learn about fractions before the intervention. For this reason, this pretest focused on every-day fractions (e.g., 1/2, or 3/4) and contained items adapted from an existing German fraction test (Padberg, 2002). Two generic items from the pretest can be seen in Fig. 3. Here, students had to name the fraction 3/4 and had to color 1/3 of a continuous tape diagram.

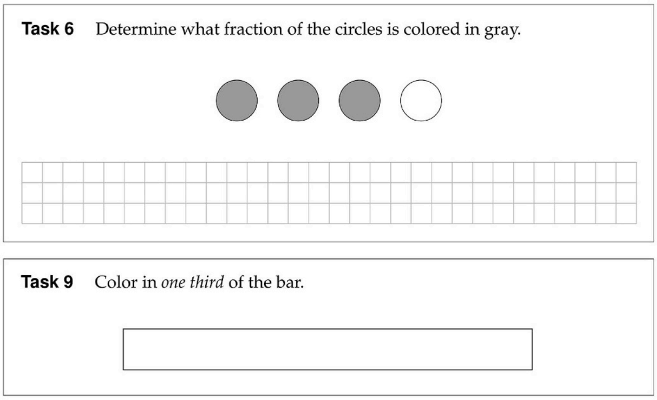

Fig. 3. Sample items from the pretest with focus on the part-whole concept of fractions on a discrete set of items (top) and fraction magnitude on a continuous tape diagram (bottom).

To measure knowledge of fraction concepts after the intervention, a two-dimensional posttest (38 items total; inter-rater reliability $0.812 \le \kappa \le 1.00$) was developed as a second test instrument. Here, items focused on the part-whole concept, expanding and simplifying fractions, and comparing fractions and were constructed based on the following rationale: On the *Procedural Knowledge* scale (18 items; $\omega = 0.879$, 95%CI [0.868, 0.888]) students had to operate with fractions in symbolical representation and no transition between representations was necessary (e.g.: "Write the missing numbers into the boxes."; see Fig. 4, Task 7). In items within the *Conceptual Knowledge* scale (20 items; $\omega = 0.867$, 95%CI [0.856, 0.877]) operating with

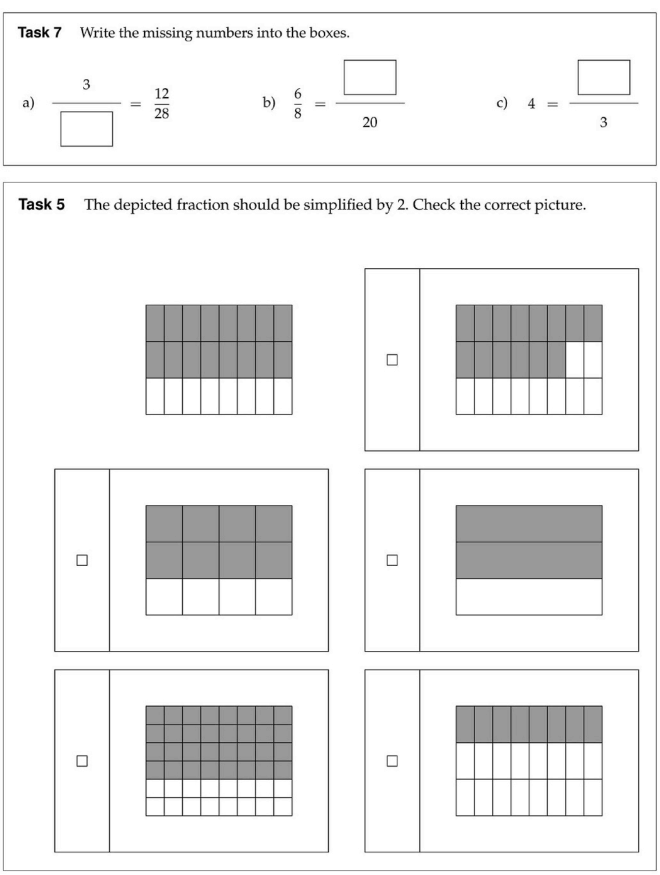

Fig. 4. Sample items from the posttest with focus on expanding and simplifying fractions assessed as procedural knowledge (top) and conceptual knowledge (bottom).

iconic representations of fractions, or transitions between non-symbolic and symbolic representations were necessary (e.g.: "The depicted fraction should be simplified by 2. Check the correct picture."; see Fig. 4, Task 5).

We assumed that procedural knowledge items can be solved using basic arithmetics with natural numbers and without an elaborated conceptual understanding of fractions (e.g., dividing 28 by 4 in Task 7a; Fig. 4). In contrast, we think that solving conceptual knowledge items requires such elaborated conceptual understanding of fractions (here: simplifying fractions as the given division becoming coarser), especially when common misconceptions of simplifying fractions (here: subtracting two parts, or dividing by 2) are given as distractors. We hereby follow commonly accepted argumentations regarding the use of representations in the field of mathematics education (e.g., Behr et al., 1983; Bruner, 1960; Duval, 2006; Lesh, 1981; Padberg & Wartha, 2017).

2.4. Procedure

Study 1 was conducted at the beginning of the school year 2016/2017 and study 2 was conducted at the beginning of the school year 2017/2018. Schools and teachers could decide whether they participated in the intervention study or not on a voluntary basis before they were randomly assigned to the three different intervention groups.

The classroom intervention studies covered the first four weeks of the school year. Due to several school-dependent reasons, the total number of mathematics lessons during these four weeks differed slightly between the classrooms ($M = 15.48$ lessons of 45 min each, $SD = 2.03$). Only topics concerning fractions were taught during these lessons. As the variance in total instruction time can assumed to be relevant for students' outcome, we incorporated instruction time on fractions in the following analyses (see Appendix A).

The pretest was conducted in the first 15 minutes of the first lesson, the posttest (55 minutes) was conducted after the four-week intervention. Students took both tests anonymously on a voluntary basis and with the informed consent of their parents.

During the intervention, students from the Scaffolded Curriculum group worked with the eBook on iPads. Teachers in this group were advised to use the eBook for at least half of their instruction time, to make use of the given introductions for each topic and to let their students study with the interactive exercises. Students from the Curriculum group worked with the paper copy version during classroom instruction. Teachers in this group were also advised to make use of the given introductions for each topic and to use the exercises in the paper copy version of the learning environment. Because of this set-up, we assume that students within the Curriculum group did not have access to the scaffolds and features implemented into our interactive learning environment—e.g., feedback and adaptivity in the Curriculum group *may* be provided by the teachers on a *classroom-level*, yet feedback and adaptivity in the Scaffolded Curriculum group *may also* be provided by teachers on a classroom-level, but *will* be provided by the eBook on a *student-level*. Students and teachers from the Traditional group did not receive any of our material before and during the intervention. Teachers in this group were only advised with educational objectives for the four-week classroom instruction, to ensure the teaching of the same content during the intervention across all three groups. They chose their material and their method of introducing fractions on their own.

Both the Bavarian Ministry of Education and the responsible local education authority approved the studies.

2.5. Implementation check

In both studies, we conducted the same 90-minutes teacher training before the intervention to advise the teachers with the educational objectives for the four weeks of instruction. Furthermore, each teacher was handed an 18-page booklet with detailed information on the aim of the research project, the educational objectives (i.e., the part-whole concept, expanding and simplifying fractions, fractions on the number line, fractions representing more than one whole, mixed numbers; and comparing fractions), and ideas on how to use the developed material (eBook and paper copy) in mathematics classrooms.

Teachers were interviewed after the intervention to check whether the teachers gave their lessons within their classes as intended (e.g., covering all topics, total instruction time):

For the Traditional group, we consider the implementation to be adequate, when the educational objectives given were all addressed during the intervention. Teachers within the Traditional group stated that this was achieved.

For the Curriculum group, we consider the implementation to be adequate, when the educational objectives given were all addressed during the intervention, when our exploratory introductions were used to start with new topics, and when all 90 tasks (i.e., the printed version of the widgets in the eBook) were worked on at least once. Teachers within the Curriculum group stated that this was realized.

For the Scaffolded Curriculum group, we consider the implementation to be adequate, when the educational objectives given were all addressed during the intervention, when our exploratory introductions were used to start with new topics while using the eBooks' interactive content, and when all 90 widgets were worked on at least once with recourse to the eBooks' scaffolds. Teachers within the Scaffolded Curriculum group stated that this was realized.

For students from the Scaffolded Curriculum group, we have information about their use of the eBook during the intervention in form of process data (e.g., total number of items worked on, solution rates for each widget, total time on task), additionally. For the purpose of this paper, we used this rather large pool of additional information solely to verify *whether* the eBook was used during classroom instruction. Since an adequate analysis of this process data would exceed the framework of this article and would focus on only one of the three treatment groups (i.e., the Scaffolded Curriculum group), we will not refer to information about *how* the eBook was used during the intervention within this article. For a description of potential benefits of process data for research purposes, see Hoch et al. (2018) or Goldhammer, Naumann, Rölke, Stelter, and Tóth (2017).

We interviewed teachers from both experimental groups about the learning environment (e.g., applicability in real classroom contexts, curricular validity, etc.), and asked teachers from the Scaffolded Curriculum group about the integration of tablet PCs into their classroom practice (e.g., easiness of use, additional preparation time, student-device-interaction, etc.). The teachers' statements were positive in general, and no salience that would affect the results of the intervention from our point of view were reported. Thus, considering the teacher training, the detailed booklet, the teacher interviews, and the process data, we concluded that the teachers in all three intervention groups implemented the treatment as intended.

2.6. Data and statistical analyses

Generalized linear mixed models (GLMM) were used to estimate differences between the three intervention groups in answering the test items, as they have several advantages over other statistical methods (e.g., handling of unbalanced designs and nested structures of the sample, see Brauer & Curtin, 2018, and handling of dichotomous data, see Anderson, Verkuilen, & Johnson, 2010). Four separate analyses were conducted for study 1 (high-achieving students, both procedural and conceptual knowledge) and study 2 (low-achieving students, again both procedural and conceptual knowledge). The full models contained fixed effects for the predictor variable *Group* (factor, using the Traditional group as the baseline for the model) and for the control variables *Prior knowledge* (number of items solved in the pretest, centered at grand mean), *Gender* (-0.5 = female; 0.5 = male), and *Instruction time* (number of fraction lessons, centered at grand mean). The models allowed for random intercepts for *Students*, *Classrooms* and *Items*. Likelihood-ratio tests for different models (i.e., from baseline models to full models, see Table 3) were conducted to determine appropriate models for the different analyses while considering the nested structure of the data (students within classrooms). We chose the final models regarding both those likelihood-ratio tests and post-hoc power analyses (see Appendix A). All analyses were conducted in *R* (R Development Core Team, 2008) using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015).

Estimates are given as log-odds which can be transformed into probabilities of obtaining a correct answer. In consequence of the data preparation described above, the *Intercept* describes the estimated probability of getting a right answer on an item of average difficulty from an average student within the Traditional group, i.e., a student who had no access to interactive and adaptive scaffolds and was not taught with regards to our elaborated fractions curriculum. Differences between the intervention groups were investigated via post hoc analyses using Tukey contrasts. For that, the *multcomp* package (Hothorn, Bretz, & Westfall, 2008) was used.

From our point of view there is no canonical way to calculate or report standardized effect sizes for group differences in GLMMs.

Table 3
Overview of the models used for the prediction of the effects of the intervention.

Fixed effects	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	yes	yes	yes	yes	yes	yes
Predictor variables						
Intervention groups	no	yes	yes	yes	yes	yes
Control variables						
Prior knowledge	no	no	yes	yes	yes	yes
Gender	no	no	no	no	yes	yes
Instruction time	no	no	no	no	no	yes
Random effects	**Model 0**	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Student	yes	yes	yes	yes	yes	yes
Item	yes	yes	yes	yes	yes	yes
Classroom	no	no	no	yes	yes	yes

Note. See Appendix A for a complete overview of the model selection procedure for all four analyses (high-achieving students and low-achieving students, both for procedural knowledge and conceptual knowledge of fractions).

Therefore, we report unstandardized within-study effect sizes (Pek & Flora, 2018; Wilkinson & Task Force on Statistical Inference, 1999) for all comparisons of the intervention groups: We compare the likelihood of students from one group to solve an item in the test correctly to the likelihood of students in another group as a quotient, relative to the less-likely-to-solve group. We further report post-hoc power analyses for those effects between intervention groups, calculated via Monte Carlo Simulation with the *simr* package (Green & MacLeod, 2016).

As can be seen in Tables 2 and 1% of the high-achieving students and 16% of the low-achieving students did not take part in the pretest. Their missing values for the control variable *Prior knowledge* were estimated using multivariate imputation by chained equations. Here, a random sample from the observed values was imputed using the *mice* package (Buuren & Groothuis-Oudshoorn, 2011). It should be noted that dropping the results of these students completely from the analyses would not alter the results, but would reduce the power of the analyses substantially.

3. Results

We report analyses of the pretest results for both studies at first. Results of the posttest are reported for study 1 and study 2 separately before differences between high-achieving and low-achieving students are given.

3.1. Prior knowledge of fractions

First, we checked whether high-achieving and low-achieving students, girls and boys, and students from the three different intervention groups differed in terms of prior knowledge of fractions before the intervention. We used the pretest as a measure of this prior knowledge. Table 4 summarizes the results from the pretest for high-achieving students (left) and low-achieving students (right), respectively. Here, the mean number of items solved in the pretest are reported separately for the total sample, female and male students, and different intervention groups.

For our a priori appraisal of high-achieving and low-achieving students to be eligible, we expected the high-achieving students to have more prior knowledge of fractions than the low-achieving students *before* the intervention. As can be seen in Table 4, the data is in line with our expectation: high-achieving students solved on average 4.84 of the 10 items in the pretest, whereas low-achieving students only got a mean number of 2.28 items correct. Thus, there was a significant and large effect of the a priori achievement grouping—operationalized through different school tracks of the German three-track public school system—on the pretest scores, $t(645.45) = 15.66$, $p < .001$, $d = 0.95$, 95%CI [0.81, 1.10], in favor of the high-achieving students. Therefore, we concluded that our operationalization of high-achieving and low-achieving students is appropriate.

Table 4
Results from the pretest for high-achieving students (study 1, left) and low-achieving students (study 2, right); reported separately for the total sample, female and male students, and the different intervention groups.

	Study 1: High-achieving students			Study 2: Low-achieving students		
	M	*SD*	*N*	*M*	*SD*	*N*
Total	4.84	2.88	745	2.28	2.02	260
Gender[a]						
Female	4.36	2.78	342	2.21	2.04	110
Male	5.27	2.91	393	2.33	2.01	150
Intervention groups						
Scaffolded Curriculum	4.92	2.91	159	2.55	2.18	107
Curriculum	4.96	2.92	320	2.17	1.96	71
Traditional	4.65	2.82	266	2.02	1.81	82

Note. Scale of Pretest: 0–10; *M* = Mean number of items solved, *SD* = Standard deviation. [a]10 high-achieving students did not report their gender.

To interpret possible gender differences in the effects of interactive and adaptive scaffolds, we investigated whether girls and boys differed in the pretest. High-achieving boys solved on average one item more than high-achieving girls, whereas differences for low-achieving students seem negligible (see Table 4). In fact, there was a significant but small effect of gender on the pretest score for high-achieving students, $t(726.24) = -4.32$, $p < .001$, $d = 0.32$, 95%CI [0.17, 0.47], while the effect was not significant for low-achieving students, $t(232.95) = -0.46$, $p = 0.64$. These differences between girls and boys should be kept in mind when interpreting the results for high-achieving students.

Most important for the interpretation of these cluster randomized classroom studies, the three different intervention groups should be comparable in terms of prior knowledge of fractions for both high-achieving students and low-achieving students. While the Curriculum group reached the highest pretest scores in the high-achieving sample, the Scaffolded Curriculum group showed the highest pretest results in the low-achieving sample (see Table 4). However, these differences in pretest scores were not significant for both high-achieving students, $F(2, 742) = 0.90$, $p = .41$, and low-achieving students, $F(2, 257) = 1.74$, $p = .18$. Hence, we concluded that the Scaffolded Curriculum group, the Curriculum group and the Traditional group had comparable prior knowledge of fractions before the intervention in both the high-achieving and the low-achieving sample.

3.2. Study 1: Effects on high-achieving students

We asked whether the scaffolds and/or the curriculum had a positive effect on students' development of fraction knowledge. In particular, we checked whether students who had access to interactive and adaptive scaffolds (students from the Scaffolded Curriculum group) or only to our elaborated fractions curriculum (students from the Curriculum group) showed a higher probability of getting correct answers in the posttest on fractions than students taught with traditional material (students from the Traditional group). Parameter estimates based on the GLMMs for both procedural and conceptual knowledge of fractions in the high-achieving sample are given in Table 5 (left). Estimates are given for the predictor variables (intervention groups) as well as control variables (prior knowledge, and gender). They are reported as log-odds in Table 5 and are transformed into estimated probabilities in the following paragraphs; levels of significance are indicated with asterisks. Results differed for procedural and conceptual knowledge in the high-achieving sample.

3.2.1. Procedural knowledge of fractions for high-achieving students

The estimated probability for getting a correct answer on an average item in the procedural knowledge scale of the posttest for an average student from our high-achieving sample was 70%, 95%CI [57.8, 79.4]. This baseline represents the results for students in the Traditional group. The log-odds for getting a correct answer varied between students ($Var = 0.74$), classrooms ($Var = 0.07$), and items ($Var = 1.06$). Here, no significant differences between the three intervention groups were found, with students from the Scaffolded Curriculum condition showing an estimated probability of 70%, 95%CI [57.9, 80.4], and students from the Curriculum condition showing an estimated probability of 75%, 95%CI [63.8, 83.1], for getting a correct answer in the procedural knowledge scale of the posttest, all $ps > .05$. Thus, we conclude that neither our elaborated fractions curriculum nor interactive and adaptive scaffolds did lead to better procedural knowledge of fractions in the high-achieving sample.

There was a significant effect of prior knowledge of fractions on solving procedural knowledge items in the posttest, $p < .001$, with

Table 5
Parameter estimates based on the generalized linear mixed models for posttest data (both procedural and conceptual knowledge); reported separately for high-achieving students (study 1, left) and low-achieving students (study 2, right).

	Study 1: High-achieving students						Study 2: Low-achieving students					
	Procedural knowledge			Conceptual knowledge			Procedural knowledge.			Conceptual knowledge		
Fixed effects	Est.	*SE*	Eff.	Est.	*SE*	Eff.	Est.	*SE*	Eff.	Est.	*SE*	Eff.
Intercept	0.84**	0.26		0.75***	0.18		−1.91***	0.33		−1.20***	0.24	
Predictor variables												
Traditional →Scaffolded Curriculum	0.03	0.17	–	0.62***	0.15	18%	0.50**	0.15	54%	0.55***	0.12	48%
Traditional →Curriculum	0.24	0.14	–	0.68***	0.13	19%	0.01	0.17	–	0.24	0.13	–
Curriculum →Scaffolded Curriculum	−0.21	0.17	–	−0.07	0.15	–	0.49**	0.16	54%	0.31*	0.13	21%
Control variables												
Prior knowledge	0.22***	0.01		0.22***	0.01		0.10***	0.03		0.11***	0.03	
Gender	–			−0.21**	0.07		–			–		
Random effects	Var.	*SD*		Var.	*SD*		Var.	*SD*		Var.	*SD*	
Student	0.74	0.86		0.44	0.66		0.59	0.77		0.39	0.62	
Item	1.06	1.03		0.46	0.68		1.73	1.31		0.94	0.97	
Classroom	0.07	0.26		0.06	0.25		–			–		

Note. Study 1: High-achieving students (Procedural knowledge: 13410 observations, 745 students, 29 classrooms, 18 items; Conceptual knowledge: 14900 observations, 745 students; 29 classrooms; 20 items); Study 2: Low-achieving students, i.e., students in lowest school track (Procedural knowledge: 4680 observations, 260 students, 18 items; Conceptual knowledge: 5200 observations, 260 students; 20 items). Predictor variables: differences between groups are given as post hoc Tukey contrasts. Control variables are centered at grand mean of the sample. Est. = Estimates are given as log-odds. Eff. = Within-study effect sizes are given as relative differences in estimated probabilities for solving an item correctly between intervention groups. Var = Variance, *SE* = Standard error, *SD* = Standard deviation. Levels of significance: *$p < .05$, **$p < .01$, ***$p < .001$.

solving one additional item in the pretest resulting in a 6.3% more likely correct answer on a procedural knowledge item in the posttest. Gender and instruction time had neglectable effects on procedural knowledge for the high-achieving sample (Appendix A).

3.2.2. Conceptual knowledge of fractions for high-achieving students

The estimated probability for getting a correct answer on an average item in the conceptual knowledge scale of the posttest for an average student from our high-achieving sample was 68%, 95%CI [59.8, 74.9]. Again, this baseline represents the results for students in the Traditional group. The log-odds for getting a correct answer varied between students ($Var = 0.58$), classrooms ($Var = 0.06$), and items ($Var = 0.74$). We expected that students from the Scaffolded Curriculum group would show the highest probability in solving items from the posttest since they could benefit from both the interactive and adaptive scaffolds on tablet PCs and our elaborated fractions curriculum to develop fraction concepts. In fact, students from the Curriculum group had an estimated probability of 81%, 95%CI [74.7, 85.4], to answer conceptual knowledge items from the posttest correctly, whereas students from the Scaffolded Curriculum group reached a slightly lower estimated probability of 80%, 95%CI [72.6, 85.1]. Post-hoc Tukey tests showed significant differences between the Scaffolded Curriculum group and the Traditional group, $p < .001$, post-hoc power of 98%, 95%CI [97.0, 98.9], and between the Curriculum group and the Traditional group, $p < .001$, post-hoc power of 100%, 95%CI [99.6, 100.0]: Students from the Scaffolded Curriculum group were 18% more likely and students from the Curriculum group were 19% more likely to answer a conceptual knowledge item correctly than students from the Traditional group. Yet, no significant difference between the treatment groups could be found in the high-achieving sample, $p = .90$. These results suggest that our holistic material to teach initial fraction concepts was helpful to teach high-achieving students both in the scaffolded version on tablet PCs and the paper-based version. Yet, no additional benefit from interactive and adaptive scaffolds on tablet PCs was found.

As expected, higher prior knowledge of fractions resulted in significantly better conceptual knowledge for high-achieving students in the posttest, $p < .001$. Being able to solve one additional item in the pretest before the intervention led to a 6.7% higher estimated probability to give a correct answer in a conceptual knowledge item within the posttest.

Bearing in mind that high-achieving boys turned out to be better than girls in terms of prior knowledge of fractions before the intervention, it seems noteworthy that high-achieving girls outperformed high-achieving boys after the intervention significantly, $p < .01$. Girls were 7.5% more likely to solve a conceptual knowledge item in the posttest than boys. Hence, the results suggest that female high-achieving students did not only close up to male high-achieving students, but developed better initial fraction concepts than male students. Instruction time had a neglectable effect on conceptual knowledge for the high-achieving students (Appendix A).

3.3. Study 2: Effects on low-achieving students

Analyses were likewise conducted for the low-achieving sample. Parameter estimates based on the GLMMs for the low-achieving students (again, both procedural knowledge and conceptual knowledge) are given in Table 5 (right) for posttest data. Estimates are also reported as log-odds. Here, we equivalently investigated whether the use of scaffolds and/or the curriculum lead to better fraction knowledge when compared to traditional classroom instruction. For the low-achieving sample, results for procedural and conceptual knowledge are similar. Here, gender and instruction time had neglectable effects on both scales, and classroom-level random intercepts are not considered (Appendix A).

3.3.1. Procedural knowledge of fractions for low-achieving students

For the low-achieving sample, the estimated probability for getting a correct answer on an average procedural knowledge item for an average student from the Traditional group was 13%, 95%CI [7.2, 22.1]. The log-odds varied between students ($Var = 0.59$) and items ($Var = 1.73$).

Being taught with our elaborated fractions curriculum for initial fraction concepts (i.e., taught in the Curriculum condition) resulted in an estimated probability for getting a correct answer of, again, 13%, 95%CI [7.2, 22.3]. Yet, working with our interactive and adaptive learning environment on tablet PCs (i.e., taught in the Scaffolded Curriculum condition) increased the probability of solving a procedural knowledge item to 20%, 95%CI [11.4, 31.5]. Post-hoc Tukey tests showed that the differences between the Scaffolded Curriculum group and the Traditional group was significant, $p < .01$, resulting in a 54% higher chance of solving a procedural knowledge item in the posttest, post-hoc power of 90%, 95%CI [87.9, 91.7]. Yet, no significant difference between the Curriculum group and the Traditional group could be found in the low-achieving sample, $p = .99$. Students from the Scaffolded Curriculum group outperformed students from the Curriculum group, yielding a 54% higher chance of solving a procedural knowledge item in the low-achieving sample, $p < .01$, post-hoc power of 87%, 95%CI [84.9, 89.1]. These results suggest that scaffolds during working with digital devices in mathematics classrooms can increase short-term learning outcomes on a procedural knowledge level for low-achieving students.

Higher prior knowledge of fractions resulted in significantly better procedural knowledge achievement in this low-achieving sample, $p < .001$. Here, students who solved one more item in the pretest were 9.0% more likely to solve procedural knowledge items in the posttest. This improved performance seems noteworthy since the estimated probability for solving posttest items is altogether low for low-achieving students.

3.3.2. Conceptual knowledge of fractions for low-achieving students

For low-achieving students, the estimated probability for getting a correct answer on an average conceptual knowledge item for an average student from the Traditional group was 23%, 95%CI [16.0, 32.4]. The log-odds varied between students ($Var = 0.39$) and items ($Var = 0.94$).

Students from the Curriculum group had a slightly higher estimated probability for getting a correct answer of 28%, 95%CI [19.4, 37.9]. Again, students in the Scaffolded Curriculum group had the highest estimated probability for solving a conceptual knowledge item of 34%, 95%CI [25.0, 45.1]. Post-hoc Tukey tests showed that the difference between the Scaffolded Curriculum group and the Traditional group was significant, $p < .001$, resulting in a 48% higher chance of solving a conceptual knowledge item, post-hoc power of 99%, 95%CI [98.6, 99.7]. Again, no significant difference between the Curriculum group and the Traditional group could be found in the low-achieving sample, $p = .18$. Students from the Scaffolded Curriculum group outperformed students from the Curriculum group, yielding a 21% higher chance of solving a conceptual knowledge item, $p < .05$, post-hoc power of 73%, 95%CI [70.0, 75.6]. In contrast to the high-achieving sample, these results suggest that interactive and adaptive scaffolds used in mathematics classrooms combined with our elaborated fractions curriculum can lead to better conceptual knowledge of fractions in low-achieving students, while the curriculum alone seems insufficient.

Higher prior knowledge of fractions also resulted in significantly better conceptual knowledge achievement of low-achieving students, $p < .001$, with students who solved one more item in the pretest being 8.7% more likely to solve a conceptual knowledge item in the posttest.

3.4. Differences between low-achieving and high-achieving students

We asked whether the effect of interactive and adaptive scaffolds and/or an elaborated fractions curriculum on students learning fractions in grade 6 differed between high-achieving and low-achieving students. In regard to cognitive load theory and cognitive theory of multimedia learning we assumed that the use of scaffolds while learning fractions could be beneficial for both high-achieving and low-achieving students, but the impact for low-achieving students could be higher due to a possible expertise-reversal effect. We combined insights from both samples to answer this question. The results from the posttest for high-achieving and low-achieving students—and both procedural

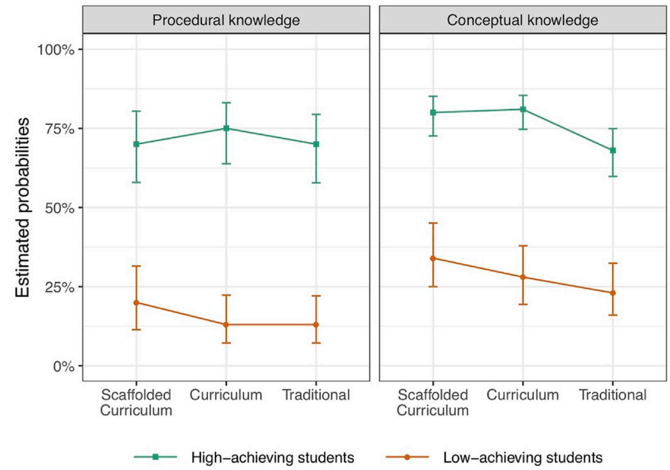

Fig. 5. Estimated probabilities for solving procedural knowledge items (left) and conceptual knowledge items (right) in the posttest compared for the different intervention groups, represented for high-achieving students and low-achieving students separately. Estimates and 95%CIs are given.

knowledge and conceptual knowledge—are summarized in Fig. 5. Here, the estimated probabilities to give a correct answer in the posttest are represented as points in the diagram for students from the different intervention groups, with 95% confidence intervals also given. Procedural knowledge is shown on the left, and conceptual knowledge is shown on the right part of the figure. The colored lines represent the high-achieving and the low-achieving sample.

As can be seen in Fig. 5, the results were mostly in line with our assumptions, yet a differentiated look on the effects seems insightful: For high-achieving students, both experimental groups showed comparable and higher conceptual knowledge of fractions after the intervention compared to the Traditional group. Using our elaborated fraction curriculum to develop initial fraction concepts both with and without interactive and adaptive scaffolds on tablet PCs seemed to have a comparable positive effect on student outcomes. Yet, the results suggest that there is no difference in high-achieving students' procedural knowledge of fractions between the three intervention groups. However, in the low-achieving sample the Scaffolded Curriculum group proved significantly better in both procedural and conceptual knowledge of fractions than the Curriculum group and the Traditional group after the intervention.

4. Discussion

The goal of the present studies was to identify effects of an elaborated fractions curriculum and interactive and adaptive scaffolds—technologically implemented into an eBook—on students learning of fraction concepts in real classroom situations. For that, we proposed a framework for the development of interactive and adaptive digital learning environments exemplified for the topic of fractions (see Fig. 1). This framework suggests grounding the developmental process of adequate learning environments on three pillars: *Content* (i.e., the subject matter that is addressed, here: basic fraction concepts; see Section 1.1), *Instructional Design* (i.e., psychological theories about learning with multimedia, here: cognitive load theory, cognitive theory of multimedia learning, integrated model of text and picture comprehension; see Section 1.2), and *Technological Implementation* (i.e., features educational technology can provide, here: interactivity and explicit use of touchscreen technology for building up embodied tasks, adaptivity, feedback; see Section 1.3). Following this framework, the designed learning environment can enable the distinction between effects of theory-grounded material design and its technological implementation in research contexts. We will discuss the results from both studies, limitations of the studies, and implications for classroom instruction as well as further research.

4.1. Effects of educational technologies' features on students' learning of fraction concepts

The results suggest that an interactive and digital learning environment with adaptive scaffolds can be used to convey concepts of fractions. Yet, we found evidence that effects of features that educational technologies can provide on students' development of fraction concepts are complex and, therefore, no general answer seems appropriate.

The findings from the present studies can be considered reliable empirical evidence that for low-achieving students interactive and adaptive scaffolds framed within a theory-grounded material can lead to better procedural and conceptual fraction knowledge during regular classroom instruction than traditional teaching and learning scenarios: Students working with our newly-developed interactive learning environment on iPads showed significantly higher procedural and conceptual knowledge of fractions than the students working with traditional material. Yet, an elaborated fraction curriculum implemented in a paper-based book with regards to the multimedia principle (Butcher, 2014) does probably not lower cognitive load sufficiently in the complex learning of fractions to be beneficial for low-achieving students: No positive effect of the use of only the elaborated fractions curriculum in a paper-based book was found for low-achieving students, as no significant differences between the Curriculum group and the Traditional group were found after the intervention. It seems that low-achieving students needed to work with interactive and adaptive content together with automatic feedback—presented on tablet PCs—to be able to benefit from the material. Hence, interactive hands-on activities, adaptive task difficulty and individual feedback have proven to successfully support low-achieving students in developing fraction concepts.

In contrast, the findings of study 1 reveal that high-achieving students did in fact profit from our elaborated fractions curriculum in a paper-based book, but did not benefit additionally from features that educational technology can provide: Both experimental groups outperformed the Traditional group in conceptual knowledge items in the posttest, while no significant difference between the two treatment groups and for procedural knowledge were found.

These differences between high-achieving and low-achieving students are in line with findings from cognitive load theory and cognitive theory of multimedia learning: Researchers report different effects between students with different prior knowledge when using instructional material that is intended to lower cognitive load in complex learning tasks (Blayney, Kalyuga, & Sweller, 2009; Kalyuga, 2007; Mayer & Pilegard, 2014; Sweller et al., 2011). However, these studies mostly made use of relatively short interventions (de Jong, 2010). Our interpretation of the results of the present studies reflects these findings: We conclude that using iconic representations of fractions to address the necessity to revise established conceptions about natural numbers did lower cognitive load. While this was beneficial for high-achieving students with higher prior knowledge, intrinsic load in these complex learning tasks was probably still too high for low-achieving students. Hence, cognitive load had to be lowered additionally to support low-achieving students in building up fraction concepts. Here, interactive material, hands-on activities, adaptive task difficulty, and individual feedback have obviously provided lower cognitive load, as also suggested by works from other authors (Arroyo et al., 2014; Black et al., 2012; Cheung & Slavin, 2013; Ma et al., 2014; Mayer & Moreno, 2003; Moreno, 2004; Moreno et al., 2006; Steenbergen-Hu & Cooper, 2014). This additional decrease in cognitive load seemed not to be necessary for high-achieving students, but was essential for students from the low-achieving sample. This suggests that specific features educational technologies can provide when designed adequately can in fact lead to better fraction knowledge within real classrooms scenarios when compared to traditional classroom instruction, especially for low-achieving students.

4.2. Limitations of these studies

Our studies show differences in the outcomes of low-achieving students and high-achieving students. Yet, our operationalization of high-achieving and low-achieving students—as students from the highest and lowest school track of the German (and with more detail: the Bavarian) public-school system—may be considered a limitation, since a priori no additional cognitive covariates were assessed to classify the students before the intervention. With regard to differences between these school tracks in standardized assessment of mathematics (Götz et al., 2013; Sälzer, Reiss, Schiepe-Tiska, Prenzel, & Heinze, 2013) and the large effects found in the pretest as well as both scales of the posttest within our studies, we assume this operationalization is appropriate. However, future research might seek to replicate the findings using different operationalizations of high-achieving and low-achieving students, e.g., standardized cognitive covariates.

In addition, these differences between low-achieving and high-achieving students lead to new questions that need further investigation and cannot be answered based on our studies: as appropriate feedback as well as adaptive task difficulty have shown to be beneficial for the learning process in different studies (Arroyo et al., 2014; Cheung & Slavin, 2013; Ma et al., 2014; Steenbergen-Hu & Cooper, 2014), one must ask why these aspects had no effect on high-achieving students in the experiment—and *how* they worked for low-achieving students. Regarding this, it can be asked whether students use feedback or whether they chose the next task immediately. Here, process data (Goldhammer et al., 2014) may be a reasonable way to address this: by looking at the data captured by the devices—e.g., time on task or the period of time that feedback was displayed (see Hoch et al., 2018)—insights into students' working routines with the devices can be investigated. For the present studies, it seems possible that high-achieving students did not use feedback, as the tasks had been easy for them. In addition, these routines—i.e., feedback, adaptivity—seemed to work well for low-achieving students. This opens up new questions about whether this only is a matter of prior knowledge or if other moderators for the effectiveness of interactive and adaptive scaffolds can be found. Regarding this, additional research is needed to identify possible learning types that can benefit more from learning with interactive and adaptive scaffolds than others.

Another aspect that might have influenced the outcome is the lack of experience both teachers and students had with using educational technology as an instrument for learning. Since none of the participating schools had their own devices, teachers as well as students worked with educational technology—providing interactive and adaptive scaffolds for each student—in the classroom for the very first time. Regarding the outcomes for high-achieving students, 90 minutes of teacher training for teachers who never used educational technology on a student level in their classrooms before may be considered insufficient. However, since there was a positive effect for low-achieving students, we conclude that the training prepared teachers for utilizing the features implemented into the eBook in their classroom beneficially and that the eBook was developed in a way that students could use it during their lessons.

As both studies took place in real classroom scenarios, it could be argued that instruction cannot be considered identical between different classrooms. On the other hand, it seems necessary to conduct studies using interactive and adaptive scaffolds *within* schools to provide ecologically valid evidence for possible benefits (Kucirkova, 2014; see also; de Jong, 2010) since they are where mathematics lessons take place. We argue that the benefits for educational research on interactive and adaptive scaffolds within real classroom settings and the relatively high sample size of both studies overcome the potential shortcomings of different instruction in different classrooms, that are taught by different teachers.

Regarding study 2, 260 students did not allow us to consider the nesting within classrooms to full extent within our model due to power restrictions. As we were not able to provide the hardware necessary to implement the Scaffolded Curriculum condition in more than six classrooms per study, this was a restriction we needed to accept. Yet, the comparisons of different models within Appendix A—allowing for a classroom random intercept in study 2—may put this potential shortcoming into perspective: directions of the effects are not altered when controlling for nestedness within classrooms.

4.3. Research on digital media in real classroom scenarios

We focus on the development of fraction concepts within this article. As this is reflected in the suggested framework for developing interactive and adaptive learning environments (Fig. 1) within the *Content* dimension, and both *Instructional Design* and *Technologically Implementation* do not focus on fractions—or mathematics—specifically, we think that our results may be generalized to the acquisition of mathematical knowledge—and knowledge developed in school contexts—in general.

Regarding this, we would like to propose the application of the suggested framework for future research in various contents: as portrayed in this article, we suggest that the development of interactive and digital learning environments should focus on both the content and instructional design, producing high-quality material explicitly developed to support the curriculum, as requested as well by other authors (e.g., Kucirkova, 2014; Pepin et al., 2017). With regard to the results of our studies, research on the effect of specific features of educational technology should address both, effects of the developed material and its technological implementation, distinctively. These methodological approaches can help gaining insights into the subtle effects of interactive and adaptive scaffolds in students' learning within school contexts, as can be seen in the different results for high-achieving and low-achieving students in our studies. In addition, such methodological approaches may underpin key aspects of the media debate (R.E. Clark, 1994; Kozma, 1994): implementing educational technology should not focus on the mere digital *device*, but on *how* specific features (e.g., embodied hands-on activities, adaptivity, feedback) are *integrated* in interactive learning environments. However, this approach is not common in educational research, and only few other articles can be found using similar research designs (e.g., Lichti & Roth, 2018; Starbek, Starčič Erjavec, & Peklaj, 2010; Tatli & Ayas, 2013).

4.4. Implications for classroom instruction

Our results suggest that students can benefit from an elaborated fractions curriculum (i.e., using non-symbolic fraction representations to inform both concepts and procedures; tasks explicitly designed to yield appropriate learning opportunities to initiate conceptual change; focusing on *representation*, *density*, and *size* before focusing on *operations*) while acquiring fraction concepts. Here, we want to emphasize the use of iconic representations of fractions—both prepartitioned and continuous, i.e., allowing for counting strategies and suppressing them—to teach basic concepts of rational numbers. This demand has a long tradition in mathematics education (e.g., Behr et al., 1983; Hart, 1989; Lamon, 2012; Post, Cramer, Behr, Lesh, & Harel, 1993). Yet, fractions instruction in curriculums and in schools still focuses largely on arithmetic procedures but not on conceptual understanding of fractions and rational numbers (Vamvakoussi et al., 2011), which can lead to common errors in handling fractions (Resnick et al., 1989). Our results suggest that a teaching approach using a more elaborated fractions curriculum—as described in this article—can support students, especially in developing conceptual knowledge of fractions.

Furthermore, we support the use of well-informed interactive and adaptive scaffolds built in appropriately designed digital learning environments to teach fraction concepts—and other mathematical concepts—in schools, regarding the results of our studies. We found evidence that especially low-achieving students benefit from interactive learning environments that work adaptive and give feedback, while there was no negative effect for high-achieving students. Moreover, non-cognitive measures were not investigated within both studies. Here, research has shown that motivation, self-efficacy and anxiety towards mathematics can be altered positively using appropriately designed features of educational technology (Hillmayr et al., submitted for publication; Hilton, 2018; Hung, Sun, & Yu, 2015; Moyer-Packenham et al., 2015; Riconscente, 2013; Schuetz, Biancarosa, & Goode, 2018). Regarding the results, well-informed and adequately developed digital material should be considered a fruitful supplement for mathematics classrooms, with tablet PCs offering a suitable integration into school contexts.

5. Conclusion

The findings from the present studies can be considered reliable empirical evidence that interactive and adaptive scaffolds implemented as features of educational technology can lead to higher-developed fraction concepts in students during regular classroom instruction than traditional teaching and learning scenarios, especially for low-achieving students. Regarding the content, we suggest that a holistic curriculum for basic fraction concepts should consist of much more than just the presentation of arithmetical procedures: it should not focus solely on symbolic fraction arithmetic that might appear meaningless to many students, but yield appropriate learning opportunities for initiating necessary conceptual change. The results suggest that an interactive and adaptive learning environment that demands constant transition between different representations of fractions can be used to convey an elaborated concept of fractions. Moreover, research on education technologies should consider the differentiation between learning material and its realization with technology, as different users may react differently to both aspects—what is an important feature for classroom work.

Declarations of interest

None.

Funding

This work is part of the "ALICE:fractions" project, which was supported by the Heinz Nixdorf Foundation, Germany, under the name "Lernen mit dem Tablet-PC: Eine Einführung in das Bruchrechnen für Klasse 6" (reference: 12502). The Heinz Nixdorf Foundation did not provide support for the preparation of the article, in study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

Approval of the study

This study was approved by both the Bavarian ministry of education, Germany (reference: X.7-BO5106/141/8) and the responsible local education authority, Munich, Germany (reference: SchR III/Erh106/1). We want to thank all students and teachers for taking part in the study.

Acknowledgements

We want to thank the editor and all reviewers for their detailed and supportive feedback that led to the final article.

Appendix A. Model Selection

We compared six different models for each of the four analyses, as depicted in Table 3 within the main article. Model 0 is the baseline model, containing a fixed intercept and random effects for students and items. In Model 1, intervention group is added as a fixed factor, with the Traditional group as the baseline, giving estimates of the treatment effect with no statistical controls. In Model 2, prior knowledge is added as a fixed effect. In Model 3, a classroom random intercept is added. In Model 4, gender is added as a fixed effect. In Model 5, instruction time is added as a fixed effect, as classes varied slightly in the number of lessons during the four weeks' intervention.

We conducted likelihood-ratio tests for those models to determine appropriate models for the different analyses. Results are given for high-achieving students and procedural knowledge (Table A1) and conceptual knowledge (Table A2), as well as for low-achieving students and procedural knowledge (Table A3) and conceptual knowledge (Table A4), separately, where the models selected for analysis are highlighted in gray.

Table A.1
Parameter estimates based on different generalized linear mixed models for procedural knowledge within the high-achieving group (study 1), given from a baseline model to a full model.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)
Intercept	0.95 (0.25)	0.79 (0.25)	0.84 (0.25)	0.84 (0.26)	0.83 (0.26)	0.82 (0.26)
Predictor variables						
Traditional → Scaffolded Curriculum	–	0.07 (0.12)	0.02 (0.11)	0.03 (0.17)	0.03 (0.17)	0.06 (0.17)
Traditional → Curriculum	–	0.32 (0.10)	0.25 (0.09)	0.24 (0.14)	0.25 (0.14)	0.26 (0.14)
Curriculum → Scaffolded Curriculum	–	−0.25 (0.12)	−0.24 (0.10)	−0.21 (0.17)	−0.22 (0.17)	−0.20 (0.16)
Control variables						
Prior knowledge	–	–	0.22 (0.01)	0.22 (0.01)	0.22 (0.01)	0.22 (0.01)
Gender	–	–	–	–	−0.05 (0.08)	−0.05 (0.08)
Instruction time	–	–	–	–	–	0.03 (0.03)
Random effects	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)
Student	1.23 (1.11)	1.20 (1.10)	0.81 (0.90)	0.74 (0.86)	0.74 (0.86)	0.74 (0.86)
Item	1.06 (1.03)	1.06 (1.03)	1.06 (1.03)	1.06 (1.03)	1.06 (1.03)	1.06 (1.03)
Classroom	–	–	–	0.07 (0.26)	0.07 (0.26)	0.06 (0.25)
Likelihood-ratio test	X^2	X^2	X^2	X^2	X^2	X^2
	–	10.28**	221.50***	17.07***	0.31	0.92

Note. 13410 observations, 745 students, 29 classrooms, 18 items. Predictor variables: differences between groups are given as post hoc Tukey contrasts. Control variables are centered at grand mean of the sample. Est. = Estimates are given as log-odds. Var = Variance, *SE* = Standard error, *SD* = Standard deviation. X^2 = Likelihood ratio test used for model comparison. Levels of significance: *$p < .05$, **$p < .01$, ***$p < .001$. We chose model 3.

Table A.2
Parameter estimates based on different generalized linear mixed models for conceptual knowledge within the high-achieving group (study 1), given from a baseline model to a full model.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)
Intercept	1.17 (0.16)	0.74 (0.16)	0.78 (0.16)	0.76 (0.18)	0.75 (0.18)	0.75 (0.18)
Predictor variables						
Traditional → Scaffolded Curriculum	–	0.63 (0.11)	0.57 (0.09)	0.59 (0.16)	0.62 (0.15)	0.62 (0.16)
Traditional → Curriculum	–	0.70 (0.09)	0.63 (0.08)	0.65 (0.13)	0.68 (0.13)	0.68 (0.13)
Curriculum → Scaffolded Curriculum	–	−0.07 (0.11)	−0.06 (0.09)	−0.06 (0.15)	−0.07 (0.15)	−0.06 (0.15)
Control variables						
Prior knowledge	–	–	0.21 (0.01)	0.21 (0.01)	0.22 (0.01)	0.22 (0.01)
Gender	–	–	–	–	−0.21 (0.07)	−0.21 (0.07)
Instruction time	–	–	–	–	–	0.00 (0.03)
Random effects	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)
Student	0.97 (0.99)	0.86 (0.93)	0.52 (0.72)	0.45 (0.67)	0.44 (0.66)	0.44 (0.66)
Item	0.45 (0.67)	0.45 (0.67)	0.46 (0.67)	0.46 (0.68)	0.46 (0.68)	0.46 (0.68)
Classroom	–	–	–	0.07 (0.26)	0.06 (0.25)	0.06 (0.25)
Likelihood-ratio test	X^2	X^2	X^2	X^2	X^2	X^2

(*continued on next page*)

Table A.2 (*continued*)

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)
	–	63.85***	265.64***	26.30***	10.09**	0.02

Note. 14900 observations, 745 students; 29 classrooms; 20 items. Predictor variables: differences between groups are given as post hoc Tukey contrasts. Control variables are centered at grand mean of the sample. Est. = Estimates are given as log-odds. Var = Variance, *SE* = Standard error, *SD* = Standard deviation. X^2 = Likelihood ratio test used for model comparison. Levels of significance: *$p < .05$, **$p < .01$, ***$p < .001$. We chose model 4.

Table A.3
Parameter estimates based on different generalized linear mixed models for procedural knowledge within the low-achieving group (study 2), given from a baseline model to a full model.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)
Intercept	−1.70 (0.32)	−1.94 (0.33)	−1.91 (0.33)	−1.91 (0.36)	−1.93 (0.36)	−1.96 (0.36)
Predictor variables						
Traditional → Scaffolded Curriculum	–	0.55 (0.15)	0.50 (0.15)	0.51 (0.23)	0.52 (0.23)	0.55 (0.24)
Traditional → Curriculum	–	0.01 (0.17)	0.01 (0.17)	0.05 (0.25)	0.03 (0.26)	0.07 (0.29)
Curriculum → Scaffolded Curriculum	–	0.54 (0.16)	0.49 (0.16)	0.46 (0.24)	0.50 (0.24)	0.48 (0.25)
Control variables						
Prior knowledge	–	–	0.10 (0.03)	0.10 (0.03)	0.09 (0.03)	0.09 (0.03)
Gender	–	–	–	–	0.28 (0.12)	0.28 (0.12)
Instruction time	–	–	–	–	–	−0.03 (0.09)
Random effects	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)
Student	0.70 (0.84)	0.63 (0.79)	0.59 (0.77)	0.52 (0.72)	0.49 (0.70)	0.49 (0.70)
Item	1.72 (1.31)	1.73 (1.31)	1.73 (1.31)	1.74 (1.32)	1.74 (1.32)	1.74 (1.32)
Classroom	–	–	–	0.08 (0.29)	0.09 (0.30)	0.09 (0.30)
Likelihood-ratio test	X^2	X^2	X^2	X^2	X^2	X^2
	–	17.27***	10.57**	7.24**	5.09*	0.12

Note. 4680 observations, 260 students, 16 classrooms, 18 items. Predictor variables: differences between groups are given as post hoc Tukey contrasts. Control variables are centered at grand mean of the sample. Est. = Estimates are given as log-odds. Var = Variance, *SE* = Standard error, *SD* = Standard deviation. X^2 = Likelihood ratio test used for model comparison. Levels of significance: *$p < .05$, **$p < .01$, ***$p < .001$. We chose model 2.

Table A.4
Parameter estimates based on different generalized linear mixed models for conceptual knowledge within the low-achieving group (study 2), given from a baseline model to a full model.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)
Intercept	−0.91 (0.22)	−1.23 (0.24)	−1.20 (0.24)	−1.19 (0.25)	−1.19 (0.25)	−1.24 (0.25)
Predictor variables						
Traditional → Scaffolded Curriculum	–	0.61 (0.13)	0.55 (0.12)	0.55 (0.16)	0.55 (0.16)	0.60 (0.16)
Traditional → Curriculum	–	0.25 (0.14)	0.24 (0.13)	0.25 (0.18)	0.26 (0.18)	0.35 (0.19)
Curriculum → Scaffolded Curriculum	–	0.36 (0.13)	0.31 (0.13)	0.30 (0.17)	0.29 (0.17)	0.25 (0.16)
Control variables						
Prior knowledge	–	–	0.11 (0.03)	0.10 (0.03)	0.11 (0.03)	0.11 (0.03)
Gender	–	–	–	–	−0.08 (0.10)	−0.09 (0.10)
Instruction time	–	–	–	–	–	−0.06 (0.06)
Random effects	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)	Var. (*SD*)
Student	0.51 (0.72)	0.44 (0.66)	0.39 (0.62)	0.36 (0.60)	0.36 (0.60)	0.36 (0.60)
Item	0.94 (0.97)	0.94 (0.97)	0.94 (0.97)	0.94 (0.97)	0.94 (0.97)	0.94 (0.97)
Classroom	–	–	–	0.03 (0.18)	0.03 (0.17)	0.02 (0.15)
Likelihood-ratio test	X^2	X^2	X^2	X^2	X^2	X^2

(*continued on next page*)

Table A.4 (*continued*)

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)	Est. (*SE*)
	–	23.04***	18.55***	3.01	0.62	1.05

Note. 5200 observations, 260 students, 16 classrooms, 20 items. Predictor variables: differences between groups are given as post hoc Tukey contrasts. Control variables are centered at grand mean of the sample. Est. = Estimates are given as log-odds. Var = Variance, *SE* = Standard error, *SD* = Standard deviation. X^2 = Likelihood ratio test used for model comparison. Levels of significance: $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$. We chose model 2.

Model selection was performed with regards to a tradeoff between the nested structure of the data due to the cluster random sampled design (students within classrooms), and the potential power of the analysis (as sample size was restricted due to available hardware, described in detail within Section 2.1): considering a classroom level random intercept will lead to severe loss of power in the analyses.

First, we excluded *Instruction time* as a fixed effect with regards to the results: neither did the inclusion of instruction time lead to significantly more suitable models in any of the four analyses, nor was the effect size (i.e., the log-odd) large enough to be interpretable (see Tables A.1, A.2, A.3 and A.4). Therefore, it seems reasonable to conclude that slightly more or less lessons during the intervention did not alter the students' outcomes.

For study 1 (high-achieving students), post-hoc power analyses suggested that the sample size was large enough to control for classroom level random intercepts within the analyses, appropriately. We relied on the likelihood ratio tests—suggesting gender to be of specific interest for conceptual knowledge only—and thus, we chose model 3 for the analysis of procedural knowledge in the high-achieving sample (see Section 3.2.1) and model 4 for the analysis of conceptual knowledge in the high-achieving sample (see Section 3.2.2).

Yet, for study 2 (low-achieving students), post-hoc power analyses suggested the sample size to be not large enough to consider classroom level random intercepts *and* yield enough power to interpret effects between the intervention groups. Considering the consistent direction of the effects throughout models 2 to 5, the comparable effect sizes (i.e., log odds), and the rather small classroom variance compared to student and item variance for both procedural and conceptual knowledge (see Tables A.3 and A.4), as well as the non-significant likelihood ratio test when including classroom level nestedness for conceptual knowledge items (see Table A.4) , we see the exclusion of a classroom level random intercept as an appropriate decision for low-achieving students. Thus, we chose model 2 for the analysis of both procedural knowledge (see Section 3.3.1) and conceptual knowledge (see Section 3.3.2) in the LST sample.

References

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *The Journal of the Learning Sciences, 21*(2), 247–286. https://doi.org/10.1080/10508406.2011.611446.

Anderson, C. J., Verkuilen, J., & Johnson, T. R. (2010). *Applied generalized linear mixed models: Continuous and discrete data for the social and behavioral sciences.* New York: Springer.

Apple Inc (2017). iBooks Author. Create and publish amazing books for iPad and Mac. Retrieved from https://www.apple.com/ibooks-author.

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education, 24*(4), 387–426. https://doi.org/10.1007/s40593-014-0023-y.

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113*(3), 447–455. https://doi.org/10.1016/j.jecp.2012.06.004.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*(4), 723–742.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), https://doi.org/10.18637/jss.v067.i01.

Behr, M., Lesh, R. A., Post, T. R., & Silver, E. (1983). Rational number concepts. In R. Lesh, & M. Landau (Eds.). *Acquisition of mathematics concepts and processes* (pp. 91–125). New York: Academic Press.

Behr, M., Wachsmuth, I., & Post, T. R. (1985). Construct a sum: A measure of children's understanding of fraction size. *Journal for Research in Mathematics Education, 16*(2), 120. https://doi.org/10.2307/748369.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. L. (2012). Embodied cognition and learning environment design. In D. Jonassen, & S. Land (Eds.). *Theoretical foundations of learning environments* (pp. 198–223). Oxford: Routledge.

Blayney, P., Kalyuga, S., & Sweller, J. (2009). Interactions between the isolated–interactive elements effect and levels of learner expertise: Experimental evidence from an accountancy class. *Instructional Science, 38*(3), 277–287. https://doi.org/10.1007/s11251-009-9105-x.

Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology, 37*(4), 247–253. https://doi.org/10.1016/j.cedpsych.2012.07.001.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23*(3), 389–411. https://doi.org/10.1037/met0000159.

Bruner, J. S. (1960). *The Process of education.* Cambridge: Harvard University Press.

Butcher, K. (2014). The multimedia principle. In R. E. Mayer (Ed.). *The cambridge handbook of multimedia learning* (pp. 174–205). (2nd ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.010.

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), https://doi.org/10.18637/jss.v045.i03.

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research & Development, 42*(2), 21–29. https://doi.org/10.1007/BF02299088.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences, 3*(9), 345–351. https://doi.org/10.1016/s1364-6613(99)01361-3.

Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Educational Studies in Mathematics, 72*(1), 127–138. https://doi.org/10.1007/s10649-009-9198-9.

DeWolf, M., Bassok, M., & Holyoak, K. J. (2015). Conceptual structure and the procedural affordances of rational numbers: Relational reasoning with fractions and decimals. *Journal of Experimental Psychology: General, 144*(1), 127–150. https://doi.org/10.1037/xge0000034.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics, 61*(1–2), 103–131. https://doi.org/10.1007/s10649-006-0400-z.

von Gagern, M., Kortenkamp, U., Richter-Gebert, J., & Strobel, M. (2016). CindyJS: Mathematical visualization on modern devices. In G.-M. Greuel, T. Koch, P. Paule, & A. Sommese (Eds.). *Mathematical software – ICMS 2016: 5th international conference, berlin, Germany, july 11-14, 2016, proceedings* (pp. 319–326). . https://doi.org/10.1007/978-3-319-42432-3_39.

Goldhammer, F., Naumann, J., Rölke, H., Stelter, A., & Tóth, K. (2017). Relating product data to process data from computer-based competency assessment. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.). *Competence assessment in education* (pp. 407–425). . https://doi.org/10.1007/978-3-319-50030-0_24.

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology, 106*(3), 608–626. https://doi.org/10.1037/a0034716.

Götz, L., Lingel, K., & Schneider, W. (2013). *DEMAT 6+. Deutscher Mathematiktest für sechste Klassen [German mathematics test for grade 6].* Göttingen, Germany: Hogrefe.

Green, P., & MacLeod, C. J. (2016). Simr: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504.

Hart, K. (1989). *Children's mathematical frameworks 8-13: A study of classroom teaching.* Windsor: NERF-Nelson.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487.

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S., & Reiss, K. (submitted). The Effectiveness of Learning Mathematics and Sciences with Digital Media in Secondary Schools: A Comprehensive Meta-Analysis. Computers and Education. Submitted for publication.

Hilton, A. (2018). Engaging primary school students in mathematics: Can iPads make a difference? *International Journal of Science and Mathematics Education, 16*(1), 145–165. https://doi.org/10.1007/s10763-016-9771-5.
Hoch, S., Reinhold, F., Werner, B., Reiss, K., & Richter-Gebert, J. (2018). *Fractions. Getting in Touch with Rational Numbers [English Apple iBooks Version]*. Technical University of Munichhttps://books.apple.com/de/book/fractions/id1459274376.
Hoch, S., Reinhold, F., Werner, B., Richter-Gebert, J., & Reiss, K. (2018). Design and research potential of interactive textbooks: the case of fractions. *ZDM Mathematics Education,* 839–848. https://doi.org/10.1007/s11858-018-0971-z.
Hoch, S., Reinhold, F., Werner, B., Reiss, K., & Richter-Gebert, J. (2018). *Bruchrechnen. Bruchzahlen & Bruchteile greifen & begreifen. (Deutsche Version)*. Technical University of Munichhttps://doi.org/10.14459/2018md1436808.
Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346–363. https://doi.org/10.1002/bimj.200810425.
Hung, C.-Y., Sun, J. C.-Y., & Yu, P.-T. (2015). The benefits of a challenge: Student motivation and flow experience in tablet-PC-game-based learning. *Interactive Learning Environments, 23*(2), 172–190. https://doi.org/10.1080/10494820.2014.997248.
de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science, 38*(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0.
Kainulainen, M., McMullen, J., & Lehtinen, E. (2017). Early developmental trajectories toward concepts of rational numbers. *Cognition and Instruction, 35*(1), 4–19. https://doi.org/10.1080/07370008.2016.1251287.
Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509–539. https://doi.org/10.1007/s10648-007-9054-3.
Kieren, T. E. (1976). On the mathematical, cognitive, and instructional foundations of rational numbers. In R. A. Lesh (Ed.). *Number and measurement: Papers from a research workshop* (pp. 101–144). Columbus: ERIC.
Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research & Development, 42*(2), 7–19. https://doi.org/10.1007/BF02299087.
Kucirkova, N. (2014). iPads in early education: separating assumptions and evidence. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00715.
Lamon, S. J. (2012). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers* (3rd ed.). New York: Routledge.
Lesh, R. (1981). Applied mathematical problem solving. *Educational Studies in Mathematics, 12*(2), 235–264. https://doi.org/10.1007/BF00305624.
Leutner, D. (2004). Instructional design principles for adaptivity in open learning environments. In N. M. Seel, & S. Dijkstra (Eds.). *Curriculum, plans, and processes in instructional design. International perspectives* (pp. 289–308). Mahwah, N.J.: Erlbaum.
Lichti, M., & Roth, J. (2018). How to foster functional thinking in learning environments using computer-based simulations or real materials. *Journal for STEM Education Research, 1*, 148–172. https://doi.org/10.1007/s41979-018-0007-1.
Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review, 38*, 201–221. https://doi.org/10.1016/j.dr.2015.07.008.
Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology, 106*(4), 901–918. https://doi.org/10.1037/a0037123.
Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.). *The cambridge handbook of multimedia learning* (pp. 31–48). (2nd ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005.
Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.). *The cambridge handbook of multimedia learning* (pp. 279–315). (2nd ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015.
Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52. https://doi.org/10.1207/s15326985ep3801_6.
Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.). *The cambridge handbook of multimedia learning* (pp. 316–344). (2nd ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.016.
McMullen, J., Laakkonen, E., Hannula-Sormunen, M., & Lehtinen, E. (2015). Modeling the developmental trajectories of rational number concept(s). *Learning and Instruction, 37*, 14–20. https://doi.org/10.1016/j.learninstruc.2013.12.004.
Meert, G., Grégoire, J., & Noël, M.-P. (2010). Comparing the magnitude of two fractions with common components: Which representations are used by 10- and 12-year-olds? *Journal of Experimental Child Psychology, 107*(3), 244–259. https://doi.org/10.1016/j.jecp.2010.04.008.
Merenluoto, K., & Lehtinen, E. (2004). Number concept and conceptual change: Towards a systemic model of the processes of change. *Learning and Instruction, 14*(5), 519–534. https://doi.org/10.1016/j.learninstruc.2004.06.016.
van Merriënboer, J. J. G., & Kester, L. (2014). The four-component instructional design model: Multimedia principles in environments for complex learning. In R. Mayer, & Hrsg (Eds.). *The cambridge Handbook of multimedia learning (2. Aufl., S. 104–148*New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.007.
Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*(1/2), 99–113. https://doi.org/10.1023/b:truc.0000021811.66966.1d.
Moreno, R., Reisslein, M., & Delgoda, G. (2006). Toward a fundamental understanding of worked example instruction: Impact of means-ends practice, backward/forward fading, and adaptivity. *Proceedings – 36th ASEE/IEEE frontiers in education conference*https://doi.org/10.1109/fie.2006.322285.
Moyer-Packenham, P. S., Shumway, J. F., Bullock, E., Tucker, S. I., Anderson-Pence, K. L., Westenskow, A., et al. (2015). Young children's learning performance and efficiency when using virtual manipulative mathematics iPad apps. *Journal of Computers in Mathematics and Science Teaching, 34*(1), 41–69.
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.
Novillis-Larson, C. (1980). Locating proper fractions on number lines: Effect of length and equivalence. *School Science & Mathematics, 80*(5), 423–428. https://doi.org/10.1111/j.1949-8594.1980.tb09687.x.
Obersteiner, A., Van Hoof, J., Verschaffel, L., & Van Dooren, W. (2015). Who can escape the natural number bias in rational number tasks? A study involving students and experts. *British Journal of Psychology, 107*(3), 537–555. https://doi.org/10.1111/bjop.12161.
Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429.
Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.). *The cambridge handbook of multimedia learning* (pp. 27–42). (2nd ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.004.
Padberg, F. (2002). Anschauliche Vorerfahrungen zum Bruchzahlbegriff und zu einfachen Rechenoperationen mit Brüchen in Modellierungskontexten [Descriptive prior knowledge on fractions in modelling contexts]. *Didaktik der Bruchrechnung. Gemeine brüche – dezimalbrüche* (pp. 289–299). (3rd ed.). Heidelberg, Germany: Springer Spektrum.
Padberg, F., & Wartha, S. (2017). *Didaktik der Bruchrechnung [Didactics of fraction arithmetic]* (5th ed.). Berlin, Germany: Springer.
Pek, J., & Flora, D. B. (2018). Reporting effect sizes in original psychological research: A discussion and tutorial. *Psychological Methods, 23*(2), 208–225. https://doi.org/10.1037/met0000126.
Pepin, B., Choppin, J., Ruthven, K., & Sinclair, N. (2017). Digital curriculum resources in mathematics education: Foundations for change. *ZDM, 49*(5), 645–661. https://doi.org/10.1007/s11858-017-0879-z.
Post, T. R., Cramer, K. A., Behr, M., Lesh, R., & Harel, G. (1993). Curriculum implications of research on the learning, teaching, and assessing of rational number concepts. In T. Carpenter, & E. Fennema (Eds.). *Research on the learning, teaching, and assessing of rational number concepts*. Hillsdale, NJ: Lawrence Erlbaum.
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.
R Development Core Team (2008). R: A language and environment for statistical computing. Retrieved from http://www.R-project.org.
Reinhold, F., Reiss, K., Hoch, S., Werner, B., & Richter-Gebert, J. (2018). *Comparing Fractions: The Enactive Way. Supporting Students' Choice of Appropriate Strategies with iPad-Assisted Instruction.* 2018 annual meeting of the American Educational Research Association, New York, USA. AERA 10.302/1303114.
Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S., & Peled, I. (1989). Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education, 20*(1), 8. https://doi.org/10.2307/749095.
Riconscente, M. M. (2013). Results from a controlled study of the iPad fractions game motion math. *Games and Culture, 8*(4), 186–214. https://doi.org/10.1177/1555412013496894.
Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M., & Heinze, A. (2013). Zwischen grundlagenwissen und anwendungsbezug: Mathematische kompetenz im internationalen vergleich [between basic knowledge and application: Mathematical competency compared to international standards]. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Eds.). *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (pp. 47–98). Münster, Germany: Waxmann.
Schnotz, W. (2014). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.). *The cambridge handbook of multimedia learning* (pp. 49–70). (2nd ed.). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.006.
Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. https://doi.org/10.1016/s0959-4752(02)00017-8.
Schuetz, R. L., Biancarosa, G., & Goode, J. (2018). Is technology the answer? Investigating students' engagement in math. *Journal of Research on Technology in Education,* 1–15. https://doi.org/10.1080/15391523.2018.1490937.
Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14*(5), 503–518. https://doi.org/10.1016/j.learninstruc.2004.06.015.
Starbek, P., Starčič Erjavec, M., & Peklaj, C. (2010). Teaching genetics with multimedia results in better acquisition of knowledge and improvement in comprehension: Multimedia in learning genetics. *Journal of Computer Assisted Learning, 26*(3), 214–224. https://doi.org/10.1111/j.1365-2729.2009.00344.x.
Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology, 106*(2), 331–347. https://doi.org/10.1037/a0034752.
Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5.
Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory.* New York:

Springerhttps://doi.org/10.1007/978-1-4419-8126-4.
Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*(4), 424–436. https://doi.org/10.1037/0022-0663.80.4.424.
Tatli, Z., & Ayas, A. (2013). Effect of a virtual chemistry laboratory on students' achievement. *Journal of Educational Technology & Society, 16*(1), 159–170.
Vamvakoussi, X., Christou, K. P., Mertens, L., & Van Dooren, W. (2011). What fills the gap between discrete and dense? Greek and flemish students' understanding of density. *Learning and Instruction, 21*(5), 676–685. https://doi.org/10.1016/j.learninstruc.2011.03.005.
Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*(5), 453–467. https://doi.org/10.1016/j.learninstruc.2004.06.013.
Van Hoof, J., Degrande, T., Ceulemans, E., Verschaffel, L., & Van Dooren, W. (2018). Towards a mathematically more correct understanding of rational numbers: A longitudinal study with upper elementary school learners. *Learning and Individual Differences, 61*, 99–108. https://doi.org/10.1016/j.lindif.2017.11.010.
Wilkinson, L., & Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*(8), 594–604. https://doi.org/10.1037/0003-066X.54.8.594.
Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636. https://doi.org/10.3758/bf03196322.