Design and research potential of interactive textbooks: the case of fractions

Stefan Hoch¹ · Frank Reinhold¹ · Bernhard Werner¹ · Jürgen Richter-Gebert¹ · Kristina Reiss¹

Accepted: 10 July 2018 / Published online: 19 July 2018
© The Author(s) 2018

Abstract

The transition from classical to electronic textbooks seems to be a logical step in the digitization advancing worldwide. However, developing an e-book ought to be more than digitizing text: key features of computer-based learning environments such as interactive exercises, adaptive demands, or automatic feedback should be integrated to take advantage of the digitization. The “ALICE:fractions” project aims at designing and evaluating an interactive mathematics textbook for introducing fractions in the classroom. It is based on research in mathematics education and includes the elements just mentioned. This paper describes the electronic textbook’s implementation and its theoretical background. Moreover, it introduces aspects that allow further information on learning processes to be gleaned. As an example, time on task is regarded in a study with 6th graders who used the electronic textbook in the classroom. Linear mixed models revealed a negative effect of time on task on task success. The effect was moderated by exercise difficulty and slightly by student competence: the effect was less pronounced in difficult exercises and for low-achieving students, whereas for high-achieving students or in easier exercises, the effect intensified.

Keywords Electronic textbooks · Interactive textbooks · Process data · Time on task · Fractions · Rational numbers

Mathematics subject classification 97U20 · 97U50

1 Introduction

Using digital devices in a classroom context has been intensively discussed during the past few years. However, the digitization of schools is a multi-faceted topic. It includes, e.g., the appropriate hardware as well as the provision of adequate software tools. The most prominent aim is to support students’ learning, but digital devices may also be used to better understand their learning processes.

In the following, we concentrate on these two aspects. With respect to the support of students learning, we describe an electronic textbook on fractions and discuss its features. These features exceed the mere digitizing of text. They include, for example, ways of manipulating content, adapting it to students’ needs, and providing immediate feedback. With respect to the understanding of students’ learning, we

introduce process data, which give insight into students’ activities. We consider how much time students spend doing exercises and take into consideration the relation between their time on task and their achievement while acquiring new concepts during regular classroom instruction.

2 Supporting students’ learning: realizing the potential of electronic textbooks to teach fractions

2.1 Teaching and learning fractions: a research review

Teaching and learning in the classroom has the purpose of enhancing students’ knowledge and competencies whether or not digital devices are used. Accordingly, classroom instruction with the help of computers or mobile devices has to be based on knowledge about learning in general and—in the case of teaching mathematics—learning mathematics in particular. Concentrating on fractions accounts for a

✉ Stefan Hoch
stefan.hoch@tum.de

¹ Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany

broad range of research, notably dating back to the 1980s and 1990s.

A crucial result of that research confirms what teachers know from their daily practice: The concept of fractions is complex and difficult for students to acquire (Behr et al. 1983). Several studies across countries show how students struggle to understand and apply rational-number concepts (e.g., Behr et al. 1992; Padberg and Wartha 2017). One main difficulty is that fractions or rational numbers can be interpreted in multiple ways. Depending on the author, the literature mentions between six (Behr et al. 1983) and eight (Padberg and Wartha 2017) different subconstructs. Although named differently, some aspects are described by most authors. The fundamental subconstruct is the *part-whole* interpretation of fractions (Charalambous and Pitta-Pantazi 2005). Here, a rational number is seen as a part of a bigger whole. Other interpretations depend on the part-whole subconstruct. These include rational numbers as measurements, quotients, ratios, operators, solutions to linear equations, and so forth (Behr et al. 1983; Malle 2004; Padberg and Wartha 2017). In school contexts, the focus often lies on the part-whole subconstruct in a formalistic-algorithmic way, neglecting the other subconstructs (e.g., Behr et al. 1983). However, a holistic understanding of fractions encompasses not only the understanding of each subconstruct on its own but also addresses their interrelations (Kieren 1976). This aspect creates an additional difficulty factor.

Another piece of evidence for the difficulty of the topic can be found in typical errors of students. In their review, Eichelmann et al. (2012) analyzed 33 international studies on fraction knowledge and named 58 typical errors discovered empirically. These errors were often systematic in nature, i.e., applying a faulty system, indicating a lack of conceptual understanding. Indeed, when confronted with their errors, students often justified their way of thinking (Padberg 1996). For example, one systematic error committed when adding two fractions was to calculate the result by separately adding the fractions' numerators and denominators without finding a common denominator first. In this case, the algorithm for multiplication was overgeneralized to addition, incorrectly producing sums such as $1/2 + 1/4 = 2/6$ instead of the correct $3/4$.

Many of these systematic errors can be explained by a *natural number bias* (e.g., Ni and Zhou 2005): Properties known to be true for natural numbers are overgeneralized to rational numbers (e.g., "Multiplication always enlarges."). It has been shown that even expert mathematicians could not overcome this bias completely when comparing rational numbers (Obersteiner et al. 2013). When advancing from natural numbers to rational numbers, concepts known to be true for natural numbers have to be reexamined and adapted to work in the context of fractions. How such changes can be achieved fruitfully is

the content of the *conceptual change theory* (cf. Vamvakoussi and Vosniadou 2004). Four conditions are beneficial (Posner et al. 1982) for achieving conceptual change in students: (1) there needs to be dissatisfaction with existing conceptions, i.e., students need to be confronted with problems that they cannot use their existing knowledge to solve, (2) the new conception must be intelligible, (3) it has to be plausible from the beginning, and (4) there should be potential for the new conception to be used in a broader context. A textbook introducing rational number concepts should therefore deliberately confront students with examples in which rational numbers differ from natural numbers.

Following Bruner (1977), information can be portrayed and acquired in three modes of representation. In enactive representations, content can be explored throughout actions. Iconic representations are image-based, whereas symbolic representations work in a more or less abstract mode. These representations must not necessarily be regarded as hierarchical and they are not necessarily connected to a child's age. Particularly in mathematics education, all three aspects are regarded as important for acquiring a deep understanding of concepts. Lesh (1979) reconceptualized and adapted Bruner's modes of representation to the field of mathematics and problem solving. Now consisting of five different modes, the theory differentiates among real scripts (i.e., real-world events, corresponding to Bruner's enactive mode), manipulatives and static images (corresponding to Bruner's iconic mode), and spoken and written symbolic representation (corresponding to Bruner's symbolic mode). His model was then adapted to the field of fractions within the *Rational Numbers Project* (Behr et al. 1983). Lesh et al. (1987a) also stressed the importance of translations between the five modes for developing rational-number concepts. Moreover, they particularly pointed out that using concrete manipulations in the classroom when teaching fractions plays an important role in the learning process (Lesh et al. 1987b).

The designers of a textbook on fractions should obviously take these findings into account. Considering students' individual understanding and their correct as well as their erroneous strategies is an important facet. Moreover, the use of manipulations, images, and symbolic representations should foster their grasp and their mastery of the concept of fractions. Regular paper-based textbooks offer scant possibilities for implementing these features; however, electronic textbooks may be better equipped. In particular, tablet use may confer a further benefit. Findings of Black et al. (2012) showed that students added better on an iPad than those adding with the same software on a conventional computer. They argued that the tablet affords a more natural way of input, explaining the difference in learning outcome.

2.2 Developing an interactive textbook: ALICE:fractions

Although a digital learning environment should incorporate the ideas mentioned in the last paragraph, it also has to accommodate the technical features. The development of ALICE:fractions (Hoch et al. 2018) as an interactive textbook for supporting students' understanding of fractions tried to combine these aspects.

ALICE:fractions provides learning material on the iPad. The learning environment's contents foster a holistic understanding of fractions. In particular, different subconstructs of rational numbers are introduced. An important characteristic is the use of various representations and the translations between them, especially the translation between iconic and symbolic representations. ALICE:fractions furthermore allows hands-on activities on the iPad in the form of manipulatives. Topics are introduced with the aim of encouraging students to revise their established conceptions about natural numbers.

At its current state of development, the interactive textbook encompasses seven sections, each intended for about 90 min of instruction time. The first section introduces fractions at a visual and symbolic level and fosters various representations. The following two sections deal with initial calculations in terms of an operator aspect of fractions ("How much is $\frac{3}{4}$ of 8?", etc.). Section four introduces equal fractions and expanding and reducing of fractions with special focus on the visual meaning of these operations. The whole of section five covers fractions on the number line. The next section introduces the concept of mixed numbers. The last

fully developed section, at the time of writing, approaches the topic of size comparisons using various feature-based strategies (see Reinhold et al. 2017a).

Each section provides teachers and students with introductory examples in the form of digital manipulatives. The content is then formalized in short book texts and precisely summarized in formulas giving generic examples. A variety of interactive exercises closes each section (see Fig. 1 for an example page).

Since the interactive textbook is viewed on a tablet, it may integrate features of computer-based learning environments, e.g., adaptivity or automatic feedback. ALICE:fractions uses iBooks Author (Apple Inc. 2017) as a framework, which enables the user to create an interactive textbook. Interactivity is put into effect using encapsulated web pages, so-called widgets, that can be accessed from the book and run in full-screen mode or even run in the page itself.

Each section of the interactive textbook makes use of widgets and offers digital manipulatives, for example cutting a pizza into pieces and fairly distributing it to plates (see Fig. 2). Furthermore, at the end of each section, many interactive exercises are provided in the form of widgets (cf. Fig. 2). Each widget is self-developed by the ALICE:fractions team using HTML5 techniques, JavaScript, and CindyJS, a framework to create interactive (mathematical) content for the web (von Gagern et al. 2016). Cindy allows content from the interactive geometry software Cinderella to be displayed in web browsers.

Each exercise consists of series of tasks, which are randomly generated by the system as long as the user asks for them. Each exercise focusses on one single aspect. If, for

The image shows a screenshot of an interactive textbook page titled "What fraction is bigger?" with the page number 49 and the section "Fractions".

Remember

If one fraction is **greater than one half** (e.g. $\frac{4}{7}$) and another fraction is **less than one half** (e.g. $\frac{2}{5}$), you can easily compare the two:

Two circles are shown. The first circle is divided into 7 equal sectors, with 4 sectors shaded purple, representing the fraction $\frac{4}{7}$. The second circle is divided into 5 equal sectors, with 2 sectors shaded teal, representing the fraction $\frac{2}{5}$. Below the circles, the fractions $\frac{4}{7}$ and $\frac{2}{5}$ are written, separated by a greater-than sign ($>$).

Exercise 78

Use this strategy to solve the problems.

Pick the **bigger** fraction.

Two boxes are shown, each containing a fraction. The first box contains $\frac{4}{6}$ and the second box contains $\frac{1}{4}$. Below the boxes is an orange button labeled "Correct?".

Fig. 1 Example page of the interactive school book. Content summary on the left, interactive exercise on the right. To start the exercise, students need to tap on the preview, causing the exercise to expand to full screen and start

Fig. 2 Digital manipulative (distributing pizza, left) and interactive exercise (right) in ALICE:fractions

example, the first task of an exercise is “Find $\frac{1}{2}$ of the circle”, all tasks in this exercise are “Find x of the circle.”

The algorithm generating the tasks is designed to adapt to the users and to take their knowledge into consideration, which has been shown in computer games to be more effective than a non-adaptive approach (Sampayo-Vargas et al. 2013). The *adaptivity* is restricted to each exercise and does not comprise the whole environment. The reason is that all chapters are introductory in nature and only partly build on each other. The adaptivity algorithm is simple. Different levels of increasing task difficulty were designed into each exercise. The exercises randomly generate sets of tasks governed by the given difficulty level’s parameters. Upon completion of this generated set (incorrectly completed tasks have to be reworked at the end of the set), the algorithm decides the difficulty level at which the next set is to be generated. If the number of incorrect answers lies below a threshold of 30–40% (depending on the exercise) of the number of tasks in the set, the next greater difficulty level is selected. Otherwise, another set is generated at the same difficulty level.

Moreover, the learning environment provides feedback (cf. Fig. 3). In their meta-analysis, Hattie and Timperley

(2007) summarized what effective *feedback* looks like. The interactive exercises are programmed to take those authors’ findings into account as far as possible and to give automatic, immediate task-level feedback that is adapted to the students’ answers.

ALICE:fractions exercises are designed to offer more possibilities for self-regulated learning than those in a normal environment. Since students are able to choose the order in which to work with the exercises and the number of tasks they complete in each widget, the interactive textbook fosters independent use of learning time. Students’ self-regulation is furthermore supported by the *graded assistance* (Reiss and Hammer 2013) that the exercises provide. This is suitably generated for each task and can be accessed at any time while working on the exercise (see Table 1 for an example).

Overall, the seven sections of ALICE:fractions contain 88 interactive widgets, 59 of which are interactive exercises like those described above. The remaining 27 widgets are introductory manipulatives or single-use exercises used in the formalizing part of the sections.

ALICE:fractions material was used in a 4-week intervention in classrooms (for details see Reinhold et al. 2017b).

Fig. 3 Different kinds of feedback in ALICE:fractions: simple feedback for a wrong response to “Find $\frac{3}{5}$ of the apples.” (left) and detailed explanation of a possible comparison between $\frac{1}{4}$ and $\frac{4}{6}$ (right)

Table 1 Graded assistance for the task “Fill in the missing number: $4/9 = 32/x$.”

Assistance level	Assistance text
Low	Try to think about whether you need to simplify or expand the fraction and by what number
Medium	This fraction was expanded. Think about what number was used
High	This fraction was expanded by 8

The learning material yielded significant learning improvements in comparison to conventional instruction. The results emphasize the usefulness of textbooks with a scientifically founded design. In particular, they provide evidence that using different representations of fractions can benefit students.

3 Understanding students’ learning: realizing the potential of electronic textbooks in research on time on task

3.1 Time on task as a possible prerequisite for successful learning processes

Studies of time on task have a long history in educational research (Bloom 1974; Kovanović et al. 2015). How time is used during classroom instruction is part of theories for effective instruction (e.g., Hattie 2009). There are important differences between time allocated for instruction, the amount really used for instruction (instructional time), and the engaged time during which students actually pay attention to tasks—often also referred to as *time on task* (Hattie and Yates 2013).

In a similar way and with reference to Weinert (1996), the work of Helmke (2009) addresses *usable instruction time* and *active learning time*. The latter thinks of classroom instruction as an offer, through which the teacher creates an *opportunity to learn* for the students. How the students use such an offer is described by their active learning time or their *time on task*. According to Winfield (1987), opportunity to learn may be measured by “time spent in reviewing, practicing, or applying a particular concept [...] with particular groups of students” (p. 439). Chickering and Gamson (1989) considered increased time on task to be one of the key principles of an effective education. Recent research, however, shows that the relationship between time on task and learning success is not as simple or direct as that (Goldhammer et al. 2017; Hattie 2009; Hattie and Yates 2013).

Time on task has been measured by observation and coding of, e.g., video recordings or rather rough indicators such as the numbers of lessons attended (cf. Kovanović et al. 2015). This kind of measurement has disadvantages and above all it is time-consuming. The use of computer-based environments in education and testing can offer alternative

measurement methods, as it allows data collection while students work with the system. These data are commonly referred to as trace data or *process data*, as they allow the students’ processes to be traced. Process data usually take the form of log files, in which different interactions of the student with the system are recorded.

One of the easiest measures to derive from such logs are count measures, counting how many actions students take within the system (e.g., how many tasks they complete) and as a further step, the frequency of certain actions. Time measures can be obtained thanks to the additional logging of timestamps. By calculating time from task start to task completion, process data allow time on task to be measured. The underlying assumption is that the whole time is spent on doing the task (Goldhammer et al. 2014). Therefore, all recorded time-on-task values have to be seen as estimations only. Since only activities within the system can be recorded, any off-task behavior outside of the system cannot be detected and time spent on it will be counted as time on task. In that sense, process data provide an upper bound to actual time on task.

Long off-task activities or an end of a task that cannot be determined by the system create outliers inducing estimation problems. Therefore, time-on-task values are preprocessed to counteract such outliers. Throughout the literature, different preprocessing methods are used (cf. Kovanović et al. 2015). Given a threshold, either chosen heuristically or obtained statistically, times exceeding that limit are either replaced by the threshold itself or the mean time value for such an action, or removed altogether from the analysis. This preprocessing (or trimming) of time data is an important step and the choice of strategy influences the results of the analysis. However, no strategy has been proven to be superior to the others (Kovanović et al. 2015).

According to van der Linden (2007, 2009), time on task can be used in two different modeling approaches. On the one hand, it can be seen as indicating a latent construct (e.g., reasoning speed; see Goldhammer and Klein Entink 2011). On the other hand, one can examine its relation to task success, using it as a predictor for differences between subjects and items, as done by Goldhammer et al. (2017).

For a given person, an inverse relation between task success and speed is expected for any task (e.g., Wickelgren 1977): the faster a person works, the more accuracy decreases (speed-accuracy trade-off). However, at a

population level, the observed effect of time on task on success may be heterogeneous or even negative (cf. van der Linden 2007).

Comparing several studies, Goldhammer et al. (2014) found that “in difficult items, task success is positively related to time on task, whereas in easy items [...], the relation is negative.” (p. 6)—and replicated that result for reading tasks and problem-solving tasks. Furthermore, in the field of problem solving, the positive effect was found to be especially strong for low-achieving adults. Conversely, the negative effect in easy tasks was more pronounced for high-achieving adults. They further argued that one reason for this moderation by task difficulty and competence can be found in the dual processing theory (Shiffrin and Schneider 1977). This theory distinguishes between mental processes that run without active control of the thinker (automatic processes) and mental process that require active control and are therefore slower (controlled processes). Hence, in tasks that allow for automatic processing—these are usually easy tasks—a negative effect of time on task on task success can be expected.

3.2 Research aims

How process data gathered during students’ work with an interactive textbook may be used in educational research is examined below. In particular, the following research questions are addressed:

1. Do students’ process data from the ALICE:fractions interactive textbook reveal an overall effect of time on task on task success in beginning instruction on fractions?
2. If so, does the effect vary between the individual students and the different exercises in the interactive textbook?

For other content domains, these questions have already been examined using different measurement tools and adult samples in testing situations (see above). Following these findings, a total effect was expected. It was also expected that the effect varied across both students and exercises.

3.3 Method and sample

155 students from six German grade 6 classrooms (65 females, 90 males) participated in a four-week intervention study. Their teachers were asked to introduce the contents of the seven chapters implemented in ALICE:fractions using the ebook on iPads. No further recommendations or restrictions were specified. The instruction encompassed 15 h of classroom teaching; the actual instruction time varied ($M = 15.17$, $SD = 1.21$) due to outside influences. Process

data were collected from 51 interactive exercises in the e-book. To get insights into students’ working processes, the interactive exercises in ALICE:fractions record the current timestamp on certain actions of the user. Informed consent was obtained from all individual participants included in the study.

Most interactive exercises in ALICE:fractions consisted of two phases: a *working phase* and a *feedback phase*. In working phases, students were confronted with a specific task to complete. Depending on the nature of the exercise, this included the entering of handwritten digits or manipulating visualizations and could even include requesting additional help from the e-book, if implemented for the given exercise. When the student requested feedback on an answer by pressing a button, a working phase ended and a feedback phase started. In feedback phases, the system evaluated the students’ input and gave task-level feedback based on the correctness and the possibly detected typical error. In most cases, feedback was provided in the form of text as well as in the form of visualizations. In some cases, detailed explanations of the correct solution could be requested. Pressing a button ended the phase, and a new working phase began with another task of the same kind. On each start and end of a working phase, the current timestamp was recorded and saved in addition to task specifics and the students’ answers. These data were saved on the iPad as a stringified JSON object using Web Storage (Hickson 2015) and could later be used to calculate time on task as the difference between the two timestamps. Students could choose to close an exercise at any point. In this case, no time was recorded.

It is important to note that time on task recorded by ALICE:fractions differed from other time-on-task measures used, say, in web-based learning. By design, tasks in the e-book were short and rarely took more than a minute to complete. Another point to note is the fact that the data were gathered during regular classroom instruction and not in a testing situation, removing some of the contextual effects a testing situation may create.

For time on task analyses, all solutions with a response time smaller than 500 ms were thought of as being randomly induced and were omitted. Afterwards, following Goldhammer et al. (2014), times were log-transformed, pulling outliers toward the middle of the distribution. Then, mean time on task and standard deviations were calculated for each exercise, as the distribution varied greatly between the widgets. To remove noise generated from unfinished exercises left open over longer periods of time (e.g., due to other classroom activities, breaks, or overnight), all times greater than the mean plus three standard deviations were replaced by this bound. Less than 1% of data points were replaced. Time values were then re-transformed into minutes for the analysis. To ease interpretation of the analysis (see Raudenbush and Bryk 2002), time-on-task values were centered

around the mean within each student and exercise (cluster-mean centering). To exclude task-level effects generated by the forced repetition of tasks, only the first occurrences were used in the analysis. A total of 48,003 log entries were included in the analysis.

To account for the multiple sources of non-independence (the same students work on the same exercises), a generalized linear mixed model (GLMM) was used for analysis. When modeling mixed effects, it is assumed that effects vary between units, in the case on hand exercises or students (random effects), while constant effects for the whole population and all exercises (fixed effects) are still estimated.

Since the effect of time on task on task success was to be examined, the way recorded time on task affects correctness was examined as a fixed effect. It describes how the probability of a correct answer varied on average when the time spent doing the task changed. Following the advice of Barr et al. (2013) as stated by Brauer and Curtin (2017), both a random intercept and a random slope for both students and exercises were included, yielding the following model:

$$\begin{aligned}\eta = & (\text{intercept } \beta_0) + \beta_1 \cdot (\text{time on task } t_{se}) \\ & + (\text{random intercept by student } b_{0s}) \\ & + b_{1s} \cdot (\text{time on task } t_{se}) \\ & + (\text{random intercept by exercise } b_{0e}) \\ & + b_{1e} \cdot (\text{time on task } t_{se}).\end{aligned}$$

The generated prediction, η , is a continuous quantity linked to the dichotomous observation (correct/incorrect) by the log-odds of obtaining a correct response: If p is the probability of a correct answer, then $\eta = \text{logit}(p) = \ln(p/(1-p))$.

The random effects, b_{0s} , b_{1s} , b_{0e} , and b_{1e} , were modeled as a normal distribution around 0, with their variances being estimated. Furthermore, the correlations between random intercept and random slope (random time-on-task effect) were estimated.

The random by-student intercept, b_{0s} , can be seen as the individual competence, whereas the random by-exercise intercept, b_{0e} , can be seen as the relative easiness of the exercise. The random slopes, b_{1s} and b_{1e} , describe the variation of the way time on task affects task success with regard to the individual student and exercise. The estimated correlations between the random effects can be used to examine the effect of item difficulty and student competence on the time-on-task effect.

The statistical analyses were conducted using R (R Development Core Team 2008). In particular, the lme4 package (Bates et al. 2015) was used to estimate the mixed models.

Table 2 Overview of model estimations

Variable	Estimate	SE
Fixed effects		
Intercept	1.1314***	0.12
Time on task	− 0.7236***	0.17
Random effects		
By student ($N=155$)		
Variance of intercept	0.23	
Variance of time on task	0.03	
Correlation between random effects	− 0.82**	
By exercise ($N=51$)		
Variance of intercept	0.63	
Variance of time on task	1.10	
Correlation between random effects	− 0.68***	

48,003 Observations; significance of correlations between random effects obtained by a likelihood-ratio test in comparison to a model with correlation forced to 0

Levels of significance: *** $p < .001$, ** $p < .01$, * $p < .05$, *ns* not significant

3.4 Results

Table 2 gives an overview of the results.

Overall, a significant fixed intercept of $\beta_0 = 1.1314$ ($z = 9.540$, $p < .001$) was estimated. It signifies an average student's log-odds of a correct response on an average exercise when spending a mean time on the item. The corresponding probability was 75.61%. In general, ALICE:fractions exercises can be thought of as easy, fitting the introductory theme of the content.

A significant negative time-on-task effect of $\beta_1 = -0.7236$ ($z = -4.163$, $p < .001$) was found over all exercises in ALICE:fractions. Thus, on average exercises, spending more time on an exercise decreased the probability of obtaining a correct response. This is in line with the findings of Goldhammer et al. (2014) for adults on easier tasks, like those in our case.

Further in line with Goldhammer et al. (2014), the effect varied across exercise and subject level. The random effect of the exercises on the time-on-task effect had a variance of 1.10—meaning that the effect varied among the different exercises. Further, a negative correlation of -0.82 was found between the random slope and the random intercept by exercise. Thus, the negative effect increased in easy exercises, whereas it declined in harder items to the point where it was even positive for the hardest five items. To check whether the correlation was significant, a restricted model was run forcing the correlation of the by-exercise random intercept and by-exercise random slope to 0. Both models were compared using the likelihood-ratio test, as appropriate for random effects (Bolker et al. 2009). The test

suggested that the unrestricted model fits the data better, $\chi^2(1) = 15.175$, $p < .001$. The correlation between intercept and slope was thus significant.

On the student level, only a little variation was found. Across students, the time-on-task effect had a variance of 0.03. The correlation between the by-student time-on-task effect and the by-student intercept was estimated to be -0.68 . Thus, the negative time-on-task effect increased slightly for higher performing students but decreased slightly for lower performing students. To check whether the correlation was significant, a restricted model was run forcing the correlation of the by-student random intercept and by-student random slope to 0. The model comparison test suggested that the unrestricted model fits the data better, $\chi^2(1) = 7.1193$, $p < .01$. The correlation between intercept and slope was thus significant.

4 Discussion

The process data from ALICE:fractions were used to examine the way time on task affects the successful solving of items during initial instruction on fractions. An overall negative time-on-task effect was found: longer time spent on the tasks apparently did not yield better results, whereas shorter response times were associated with correct answers. Moreover, exercise difficulty moderated the effect: the negative effect was weaker in harder exercises and stronger in easier exercises. Furthermore, the effect varied slightly across student competence: for high-achieving students the time-on-task effect increased slightly and for low-achieving students the effect decreased slightly.

The results reveal that—even when students acquire knowledge of fractions—the effect of the time spent on tasks is not a universal one but depends on the task itself and the person doing the task. It emphasizes that when recording time on task in a heterogeneous setting of either items or subjects, time may play different roles and its interpretation may be different.

The overall negative effect is somewhat surprising, as data were gathered during the acquisition of new conceptions, a stage when task performance is very lengthy and susceptible to errors (Anderson 1992). It is, however, possible that the effect is biased as the sample was selected from an education track consisting of those students who performed well in grade four. These students are supposed to learn new concepts relatively easily. Another explanation may be found in the fact that students used the exercises not only to acquire but also to practice rational-number concepts. The latter may overshadow effects of the former.

Furthermore, the result of observing wrong answers linked to longer times joins similar results across various domains (e.g., Beckmann 2000; Hornke 2000; Goldhammer

et al. 2014). Hornke (2005) adds a valuable interpretation: a wrong answer may be preceded by lingering pondering and finally end in random guessing.

It has to be stated that—as in the research done by Goldhammer et al. (2014)—a causal effect of time on task on task success cannot be inferred, since the time allotted to students to complete the exercises in the ALICE:fractions e-book was not manipulated. Further, it could also be the case that the difference in the effect of time on task across exercises did not come from the variation in difficulty but was rather due to individual difference in choosing varying speed-accuracy compromises. However, this possibility is unlikely, as there is empirical evidence that people do not tend to do so (cf. Goldhammer and Klein Entink 2011).

Finally, it should be noted again that the data underlying the above analyses were not collected in a testing situation, but during working phases in regular classroom instruction. The logs were collected starting with the first lesson of instruction. The next step, therefore, is to examine whether the effect of time on task might depend not only on person and task but also on the time since the beginning of the acquisition of rational number concepts.

The increasing availability of digital media in classrooms offers textbook designers chances to broaden the scope of what textbooks can do. ALICE:fractions shows that hands-on experiences can be incorporated in a digital way and that it is further possible to include other interactive exercises that give feedback to the students and adapt their difficulty themselves. Exercises using iconic representations, which usually ask for larger amounts of time and specific material, could be more easily included in classroom instruction. Further, other media like audio or video could be directly embedded in the textbook. The design of ALICE:fractions took findings from (mathematics) education research and educational-psychology research into account. The results from the work of Reinhold et al. (2017b) gave evidence for the success of the scientific textbook design.

From a research perspective, electronic textbooks offer new possibilities for gathering data during classroom instruction outside of testing situations. Besides solutions to interactive tasks, process data can be obtained, yielding insights into students' learning processes and their textbook use. One such measurement is the time which students spend working on the exercises—time on task. Interactive textbooks may therefore be used to support the assessment of effective-education parameters as formulated by Helmke (2009).

The data gathered during students' work with ALICE:fractions show that new forms of textbooks can be used not only to incorporate hands-on activities more easily, but also as a research tool, offering new avenues of

research in mathematics education and especially, in the field of mathematics textbooks.

Acknowledgements ALICE:fractions is funded by the Heinz Nixdorf Foundation. Many thanks as well to the anonymous reviewers for their constructive comments on an earlier draft of this article.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anderson, J. R. (1992). Automaticity and the ACT theory. *The American Journal of Psychology*, 105(2), 165. <https://doi.org/10.2307/1423026>.
- Apple Inc. (2017). *iBooks Author. Create and publish amazing books for iPad and Mac*. <https://www.apple.com/ibooks-author/>. Accessed 9 October 2017.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 148. <https://doi.org/10.18637/jss.v067.i01>.
- Beckmann, J. F. (2000). Differentielle Latenzzeiteffekte bei der Bearbeitung von Reasoning-Items. *Diagnostica*, 46(3), 124–129. <https://doi.org/10.1026/0012-1924.46.3.124>.
- Behr, M. J., Harel, G., Post, T., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 296–333). New York: Macmillan Publishing.
- Behr, M. J., Lesh, R., Post, T. R., & Silver, E. A. (1983). Rational-number concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 91–126). New York: Academic Press.
- 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* (2nd edn., pp. 198–223). New York: Routledge.
- Bloom, B. S. (1974). Time and learning. *American Psychologist*, 29(9), 682–688. <https://doi.org/10.1037/h0037632>.
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24(3), 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>.
- Brauer, M., & Curtin, J. J. (2017). 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*. <https://doi.org/10.1037/met0000159>.
- Bruner, J. S. (1977). *The process of education*. Cambridge: Harvard University Press.
- Charalambous, C. Y., & Pitta-Pantazi, D. (2005). Revisiting a theoretical model on fractions: Implications for teaching and research. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 233–240). Melbourne: PME.
- Chickering, A. W., & Gamson, Z. F. (1989). Seven principles for good practice in undergraduate education. *Biochemical Education*, 17(3), 140–141. [https://doi.org/10.1016/0307-4412\(89\)90094-0](https://doi.org/10.1016/0307-4412(89)90094-0).
- Eichelmann, A., Narciss, S., Schnaubert, L., & Melis, E. (2012). Typische Fehler bei der Addition und Subtraktion von Brüchen—Ein Review zu empirischen Fehleranalysen. *Journal für Mathematik-Didaktik*, 33(1), 29–57. <https://doi.org/10.1007/s13138-011-0031-5>.
- Goldhammer, F., & Klein Entink, R. H. (2011). Speed of reasoning and its relation to reasoning ability. *Intelligence*, 39(23), 108–119. <https://doi.org/10.1016/j.intell.2011.02.001>.
- 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: Research, models and instruments* (pp. 407–425). Cham: Springer International Publishing. 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>.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London & New York: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Hattie, J., & Yates, G. C. R. (2013). *Visible learning and the science of how we learn*. New York: Routledge. <https://doi.org/10.4324/9781315885025>.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (2nd edn.). Seelze-Velber: Klett/Kallmeyer.
- Hickson, I. (2015). *Web Storage (2nd Edition), W3C Recommendation*. <https://www.w3.org/TR/webstorage/>. Accessed 9 October 2017.
- Hoch, S., Reinhold, F., Werner, B., Reiss, K., & Richter-Gebert, J. (2018). *Bruchrechnen. Bruchzahlen & Bruchteile greifen & begreifen* [Apple iBooks Version] (4th ed.). München, Technische Universität München. <http://go.tum.de/623496>. Accessed 12 May 2018.
- Hornke, L. F. (2000). Item response times in computerized adaptive testing. *Psicológica*, 21(1), 175–189.
- Hornke, L. F. (2005). Response time in computer-aided testing: A “Verbal Memory” test for routes and maps. *Psychology Science*, 47(2), 280–293.
- Kieren, T. E. (1976). On the mathematical, cognitive and instructional foundations of rational numbers. In *Number and measurement. Papers from a research workshop* (Vol. 7418491, pp. 101–144). ERIC.
- Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Does time-on-task estimation matter? Implications for the validity of learning analytics findings. *Journal of Learning Analytics*, 2(3), 81–110. <https://doi.org/10.18608/jla.2015.23.6>.
- Lesh, R. (1979). Mathematical learning disabilities: Considerations for identification, diagnosis, and remediation. In R. Lesh, D. Mierkiewicz & M. G. Kantowski (Eds.), *Applied mathematical problem solving* (pp. 111–180). Columbus: ERIC/SMEAC.
- Lesh, R., Post, T., & Behr, M. (1987a). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representations in the teaching and learning of mathematics* (pp. 33–40). Hillsdale: Lawrence Erlbaum.
- Lesh, R., Post, T., & Behr, M. (1987b). Dienes revisited: Multiple embodiments in computer environments. In I. Wirsup & R. Streit

- (Eds.), *Developments in school mathematics education around the world* (pp. 647–680). Reston: National Council of Teachers of Mathematics.
- Malle, G. (2004). Grundvorstellungen zu Bruchzahlen. *Mathematik lehren*, 123, 4–8.
- 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), 2752. https://doi.org/10.1207/s15326985ep4001_3.
- Obersteiner, A., Van Hoof, J., & Verschaffel, L. (2013). Expert mathematicians' natural number bias in fraction comparison. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 393–400). Kiel, Germany: PME.
- Padberg, F. (1996). Aus Fehlern lernen: Den Mathematikunterricht durch Fehleranalysen verbessern. *Friedrich-Jahresheft: Prüfen und beurteilen*, XIV, 56–59.
- Padberg, F., & Wartha, S. (2017). *Didaktik der Bruchrechnung* (5th edn.). Heidelberg: Springer Spektrum. <https://doi.org/10.1007/978-3-662-52969-0>.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. <https://doi.org/10.1002/sce.3730660207>.
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks: Sage.
- Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J., & Reiss, K. (2017a). iPads in Grade 6 classrooms: Effects on students' choice of strategy for comparing fractions. In B. Kaur, W. K. Ho, T. L. Toh, & B. H. Choy (Eds.), *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, p. 74). Singapore: PME.
- Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J., & Reiss, K. (2017b). Manipulating fractions: Effects of iPad-assisted instruction in Grade 6 classrooms. In B. Kaur, W. K. Ho, T. L. Toh, & B. H. Choy (Eds.), *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 97104). Singapore: PME.
- Reiss, K., & Hammer, C. (2013). *Grundlagen der Mathematikdidaktik: Eine Einführung für den Unterricht in der Sekundarstufe*. Basel: Birkhäuser.
- Sampayo-Vargas, S., Cope, C. J., He, Z., & Byrne, G. J. (2013). The effectiveness of adaptive difficulty adjustments on students' motivation and learning in an educational computer game. *Computers & Education*, 69, 452–462. <https://doi.org/10.1016/j.compedu.2013.07.004>.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127–190. <https://doi.org/10.1037/0033-295X.84.2.127>.
- 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 der Linden, W. J. (2007). A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika*, 72(3), 287–308. <https://doi.org/10.1007/s11336-006-1478-z>.
- van der Linden, W. J. (2009). Conceptual issues in response-time modeling. *Journal of Educational Measurement*, 46(3), 247–272. <https://doi.org/10.1111/j.1745-3984.2009.00080.x>.
- von Gagern, M., Kortenkamp, U., Richter-Gebert, J., & Strobel, M. (2016). CindyJS. 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. 319326). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-42432-3_39.
- Weinert, F. E. (1996). *Psychologie des Lernens und der Instruktion*. Göttingen: Hogrefe.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41(1), 67–85. [https://doi.org/10.1016/0001-6918\(77\)90012-9](https://doi.org/10.1016/0001-6918(77)90012-9).
- Winfield, L. F. (1987). Teachers' estimates of test content covered in class and first-grade students' reading achievement. *The Elementary School Journal*, 87(4), 437–454.