Learning from errors:

Combining correct with incorrect worked modelling examples

MASTER THESIS Educational Science and Technology

RESEARCHER: Marion Krooshoop

FIRST SUPERVISOR: Dr. Hans van der Meij
SECOND SUPERVISOR: Dr. Alieke van Dijk

DATE: 25th of August 2019

UNIVERSITY OF TWENTE.

Acknowledgement

This master thesis provided a great opportunity for me to combine two fields that grabbed my interest: instructional design for mathematics, and performing an experimental study. I thank my supervisor Hans van der Meij for making this possible, for thinking along, and for sharing an extensive expertise on e.g. instructional design. You made time whenever possible, and your feedback was supportive and encouraging. I would also like to thank my second reader Alieke van Dijk for providing valuable feedback.

Special thanks go to Marc van Zanten, who made time to share his expertise on fraction didactics. I thank the principals, teachers, and students of the schools (Mariaschool, De Triangel, De Rietslenke, and De Talenter) for their cooperation and effort. In particular, I thank the students who helped me with the pilot tests and with recording the videos. Moreover, in regards to my job, I thank my principal and employer for their support concerning this master study.

At last, I thank my boyfriend, family, and friends for their support at any time, for believing in me, for thinking along, and for their encouragement. I am grateful to have you around me and I am happy to present you my master thesis.

Marion Krooshoop,
August, 2019

Abstract

This experimental study investigates the incorporation of errors into example-based learning, which is promising, yet insufficiently established. The fields of worked examples and modelling examples were combined, which resulted in *worked modelling examples*. The examples instructed on adding fractions. Two conditions were compared, one with correct and incorrect worked modelling examples (the C-I condition), and one with correct worked modelling examples (the C-C condition). 82 Fifth grade participants (mean age 11.2) started with a self-efficacy and self-regulation questionnaire, followed by a pre-test to measure knowledge on fractions. Next, three pairs of examples were provided in the form of instructional videos, which were alternated with practice. Video logs recorded how much of the videos was played (i.e., engagement), and practice was used as a measurement. Next, the self-efficacy questionnaire was administered again. To assess knowledge on adding fractions, an immediate post-test was administered. This test was repeated a week later (delayed post-test), followed by a transfer test to assess more complex knowledge. For both conditions, log data revealed high engagement. The C-I condition had significantly higher play rates on several comparisons. Self-efficacy increased considerably, especially in the C-C condition. Performance outcomes showed substantial increases in both conditions from pre-test to practice, and to the immediate and delayed post-test. Self-regulated learning was positively related to performance in the C-C condition, but this was not substantial in the C-I condition. This study contributes to the field of example-based learning and learning from errors, by revealing the positive effects of the combination of correct and incorrect worked modelling examples.

Key words: learning from errors, worked examples, modelling, engagement, fractions

Table of Contents

Introduction

Example-based learning is highly effective and efficient for novices learning initial problem-solving skills, which is demonstrated by a vast amount of studies (see Atkinson, Derry, Renkl, & Wortham, 2000; Sweller & Cooper, 1985; van Gog & Rummel, 2010; Wittwer & Renkl, 2010). Providing learners with examples has several prominent benefits. First, worked-out step-wise examples cost less time and effort, which is referred to as the *worked example effect* (see Renkl, 2014a). Second, learners become focused on the provided steps, which supports them to generalize rules which can be applied in other situations and contexts (Sweller, van Merriënboer, & Paas, 1998). And third, the observer builds a cognitive schema by observing the model, which he can use in other situations (Bandura, 1977).

A distinction can be made between *modelling examples* and *worked examples* (see Renkl, 2014b; van Gog & Rummel, 2010). Worked examples can be defined as step-wise expert examples that show how to find a solution for a problem statement, setting the example for similar problems to be solved (Atkinson et al., 2000). Key components are that they are textually displayed and constructed by an expert. Modelling examples can be defined as examples where a model shows his way to accomplish an exercise, and often provides explanation (Hoogerheide, Loyens, & van Gog, 2014). This *mastery model* shows competence while demonstrating how to perform an exercise (e.g., see Schunk, Hanson, & Cox, 1987). Key is that the problem is solved by the model's approach, and is communicated in a spoken form. The model can be visible or non-visible (see Hoogerheide et al., 2014).

A development that is gaining attention is the inclusion of errors in example-based learning. Errors are a fruitful learning source. They provide the opportunity to deepen understanding (see Tulis, Steuer, & Dresel, 2016). With the incorporation of errors into *example-based learning*, an opportunity emerges to stimulate learning from errors. However, this aggregation has not univocally demonstrated superior learning benefits over learning from correct examples: experimental research on examples with errors has given somewhat mixed results (e.g., see McLaren, van Gog, Ganoe, Karabinos, & Yaron, 2016). Therefore, further investigation on optimizing example-based learning with errors is of interest, and is the focus of this study.

Furthermore, learning from errors requires more than just encountering an error; *motivational factors*, such as *self-regulation* skills (like monitoring, persistence, and dealing with difficulties) and motivational beliefs (like *self-efficacy*) also play a role in order to learn from errors (Tulis et al., 2016). Motivational factors have been ignored in worked example research (van Gog & Rummel, 2010). An element that is associated with motivation is *engagement*. It represents how much time learners spend on examples, which might indicate motivation and involvement. It is of interest because being engaged is essential for learning. Measuring time during training was not common in incorrect example research, and results were mixed.

Motivational beliefs, in particular self-efficacy perceptions, have gained attention in research on modelling examples. Various sorts of models might have a different impact on self-efficacy, and in line with that, on learning outcomes (see van Gog & Rummel, 2010). Therefore, self-efficacy is a relevant factor in the present study.

Self-regulation has a positive relation with motivation (Schunk, 2005), and has been a component of research on modelling examples (e.g., Kitsantas, Zimmerman, & Cleary, 2000; Zimmerman & Kitsantas, 2002). Several empirical studies have looked at how self-regulation was influenced by modelling examples. However, the present study focuses on how self-regulation relates to performance outcomes in example-based learning.

To conclude, the present study incorporates powerful features of worked examples into modelling examples. That is, written worked-out stepwise procedures are implemented in modelling examples, i.e., auditory comments are provided on textually displayed steps. These optimized modelling examples are from now on referred to as *worked modelling examples*. This aggregation fits the advice of the review on example-based learning of van Gog and Rummel (2010).

To optimize example-based learning with errors, the present study uses a combination of correct and incorrect worked modelling examples, rather than incorrect examples without correct examples, in order to foster cognitive factors. Having correct understanding is essential for learning from errors (e.g., Dunning, Johnson, Ehrlinger, & Kruger, 2003).

Furthermore, a relevant factor that concerns the design of examples is the presence of profound explanation, and how this is provided. Pictorial explanation in addition to textual explanation was not always present in incorrect example research. In addition, learners often needed to employ extra skills, for example to find an error, or to self-explain the error in order to provoke deeper understanding. The requirement to self-explain might impede learners who lack this meta-cognitive skill, especially novices (Berthold & Renkl, 2009). Overall, providing explanatory instruction with depictive representations is paramount in the present study.

Hence, the current study investigates the influence of the combination of correct and incorrect worked modelling examples on cognitive factors (i.e., practice, immediate, delayed, and transfer performance). The examples are about mathematics, in particular, video examples instruct on adding fractions at the primary education level. Moreover, motivational factors are investigated, i.e., the influence of the combination of correct and incorrect worked modelling examples on engagement and self-efficacy, and the relation between self-regulation and performance when learning from such examples.

Theoretical Framework

Learning from Errors

Using errors has great potential for education. Understanding of errors, in addition to having correct knowledge, can enrich the mental model (Heemsoth & Heinze, 2014). Information that is inconsistent (in the present study: correct vs. incorrect information), makes differences stand out, which fosters learning (Bransford & Schwarz, 1999). By becoming aware of errors, knowledge is deepened, and choosing the correct step becomes self-evident, especially when errors are illuminated (Große and Renkl, 2007). Regarding mathematics, errors could foster understanding (Borasi, 1987). *Errors* can be viewed as a general term which includes mistakes due to misconceptions or to other factors. Misconceptions are repetitive regular errors (Smith, diSessa, & Roschelle, 1993), due to deficits in a cognitive framework (Hadjidemetriou & Williams, 2002). Other factors causing errors might be, for example, flawed remembrance (Hadjidemetriou & Williams, 2002), reading mistakes, and negligence (Confrey, 1990).

Although example-based learning has a solid basis of learning benefits, the promising approach of incorporating errors requires more investigation. In worked examples research this has been investigated by using *erroneous examples*, which can be defined as worked examples containing at least one incorrect step (McLaren et al., 2012; Tsovaltzi, McLaren, Melis, & Meyer, 2012). In modelling examples, errors have been incorporated for quite some time by using *coping models*, who can be defined as models who struggle and make errors on their way to the correct solution (van Gog & Rummel, 2010).

Previous incorrect example studies showed different ways to present and design incorrect examples. Two categories could be distinguished. In the first, learners were provided with the correct solution *after* they *fixed* or *explained* errors in an example, or were *not* provided with a correct example at all. In the second category, learners *received* a correct solution *together with* or *prior to* an incorrect solution.

In the first category, empirical findings were inconclusive. Some found positive results (Adams et al., 2014; Tsovaltzi et al., 2012), especially when the errors were indicated (Barbieri & Booth, 2016). Others found benefits for learners with high prior knowledge (Heemsoth & Heinze, 2014). Some research found no differences compared to correct examples (Wang, Yang, Liu, Cheng, & Liu, 2015), and in the study of Große (2018), correct examples outperformed incorrect examples.

In the second category, learners could either be presented with a problem with as well an incorrect solution procedure, as a correct solution in the same example (e.g., Große & Renkl, 2007; Schunk et al., 1987), or they could be presented with two problems, one providing a correct solution procedure, and the other providing an incorrect solution procedure (e.g., Booth, Lange, Koedinger, & Newton, 2013). Empirical findings in this category were also inconclusive. Durkin and Rittle-Johnson (2012) found that the combination of correct and incorrect examples was beneficial for learning. Große and Renkl (2007) gained positive results for learners with high prior knowledge, and in case the error was highlighted, the learners with low prior knowledge benefitted. Zhao and Acosta-Tello (2016) found only benefits for learners with high prior knowledge. Isotani et al. (2011) did not find differences compared to correct examples. In the study of Booth et al. (2013), superiority of either the combined condition or the correct condition depended on what task was measured. Baldwin (1992) demonstrated that showing the combination of a correct an incorrect model was superior over only a correct model. Schunk et al. (1987) found superiority in performance of coping models over mastery models, whereas Schunk and Hanson (1985) found no differences. Braaksma, Rijlaarsdam, and van den Bergh (2002) discovered that weak observers benefitted from coping models, whereas good observers benefitted from mastery models. Other empirical research on mastery vs. coping models gave inconclusive results about learning outcomes (Lauzier & Haccoun, 2014).

Hence, both categories yielded mixed results. In line with the second category, the focus of the present study is on the combination of correct and incorrect examples. This approach is believed to be beneficial, because providing a correct example prior to or together with an incorrect example can be essential. Namely, correct information should serve as a foundational framework to enable learners to comprehend errors, especially when learners do not know a lot about the content at hand (Dunning et al., 2003; van Gog, 2015).

Design of the Examples

According to the *multimedia principle*, explanatory depictive representations (e.g., pictorial and graphical) integrated with descriptive representations (textual and verbal), has shown to support deeper understanding and to enrich mental models (e.g., see Butcher, 2014). This integration (i.e., the split-attention principle) does not seem common in empirical studies on the combination of correct and incorrect examples. Durkin and Rittle-Johnson (2012) did provide a picture and text, however, those were presented apart from each other, and were provided after instructional explanation was given. Zhao and Acosta-Tello (2016) provided textual expert explanation together with the example, yet did not include depictive representations. Booth et al. (2013), Große and Renkl (2007), and Isotani et al. (2011) prompted self-explanations (either with menu options or without) to analyze the example, yet did not include depictive representations either. Concluding, an opportunity lies in providing a depictive representation integrated with a descriptive representation. In mathematics, science, and technology, this can be done by supporting an abstract (descriptive) representation, with a concrete, meaningful (depictive) representation. This combination has repeatedly shown to benefit learning, at least if the concrete representations are gradually replaced by more abstract representations, and if connections are provided between those types of representations (e.g., see Pashler et al., 2007). Hence, in order to improve benefits of incorrect examples, the design of examples can be optimized following principles of multimedia learning.

Process and Personal Factors

Apart from cognitive outcomes, the learning process and motivational factors are important in learning from errors (Tulis et al., 2016). The measurement engagement, i.e., the time the students spend on the examples, can provide information about the learning process. It could reveal possible involvement of the learners, which might indicate how motivated or interested they are. Regarding example videos, engagement can refer to *absolute time* (i.e., how much playing time the examples consume), and *relative time* (i.e, how much of the video is being played). Measuring time during the learning process is not common in incorrect example-based learning, even though time is an important factor of the worked example effect. This effect has to do with absolute time: how much time does the learning process take? Whether incorrect example-based learning could emulate this effect is unclear.

Findings on absolute time measurements by Tsovaltzi et al. (2012) gave inconsistent results. Kopp, Stark, and Fischer (2008) found that erroneous examples and worked examples required a similar training time, however, McLaren et al. (2016) found that erroneous examples demanded more time, which might be due to the need to find and fix errors. Isotani et al. (2011) performed a study that matches the present study design (i.e., they used a combination of correct and incorrect examples), and found that erroneous examples required more time than problem solving. This finding might be related to requirements to self-explain. All in all, the effect of incorrect examples with instructional explanation (i.e., without the need to find, explain, and fix errors) on absolute time demands remains unknown. To our knowledge, there is no previous research on relative time in incorrect example-based learning.

Tulis et al. (2016) propose a model, in which motivational beliefs (e.g., self-efficacy beliefs) impact the reactions on errors, and where management skills (e.g., self-regulation skills) guide learning from errors. Motivation is an important factor in whether or not the learner will actualize what was learned by the example (van Gog & Rummel, 2010). While the motivation element self-efficacy has not gained much attention in worked example research, it has been an important topic in modelling example research.

Self-efficacy beliefs (i.e., perceived capability in particular domains based on a person's own criterion) have demonstrated to influence cognitive performance (see Zimmerman, 1996). It is believed that observing another person does raise self-efficacy, because observing someone who accomplishes a task increases one's competence belief (Bandura, 2012). Schunk (1981) demonstrated that self-efficacy and performance were increased by modelling examples.

Incorporating errors can especially have a positive influence on self-efficacy. Observing someone dealing with arduous problems might increase self-efficacy belief, because the observer believes he will be able to manage as well (Bandura, 1977). This is a rationale for learning from coping models. Coping models may especially enhance self-efficacy of observers who question their competence, probably because the level of the coping model is in line with the observer's level (Schunk, 1987). Empirical research demonstrated benefits, but also showed equivalent results on self-efficacy (see Schunk, 1987). The study of Huang (2017) demonstrated superiority of coping model examples regarding self-efficacy development, but not regarding performance. A study on incorrect examples (Tsovaltzi et al., 2012), showed that self-efficacy reports were inconsistent and not in line with performance outcomes. All in all, research remains inconclusive on the influence of incorrect examples on self-efficacy.

Using self-regulation strategies (e.g., planning, monitoring, concentrating) has impact on performance (see Bandura, 2006; Pintrich, 2000; Schunk, 2005; Zimmerman, 1990). Self-regulation is an important factor in learning from errors, because errors need to be acted on in order to be beneficial (Tulis et al., 2016). Self-regulation has not gained much attention in worked example research (Tulis et al, 2016). In modelling examples, it was part of several studies, particularly regarding the impact of (coping) models on self-regulation (e.g., see Schunk & Zimmerman, 2003). In contrast, the present study examines the relation between self-regulation and learning outcomes. To our knowledge, no research on incorrect examples has examined this.

In sum, the question remains whether and how the combination of correct and incorrect worked modelling examples affects engagement and self-efficacy, and how self-regulation relates to cognitive performance in the field of learning from incorrect examples.

Research Design and Questions

This study investigated the effects of the combination of correct and incorrect worked modelling examples in the form of videos about adding fractions. It had an experimental design with a control condition and an experimental condition, respectively: a correct-correct condition (C-C condition) and a correct-incorrect condition (C-I condition). In the C-C condition, two correct worked modelling examples of a similar problem type were presented. In the C-I condition, one correct worked modelling example was followed by an incorrect worked modelling example of a similar problem type. In total, three example pairs were provided. This study examined four research questions.

Research question 1: What is the effect of a combination of correct and incorrect examples on engagement?

As described above, not much previous research with incorrect examples measured engagement, and the research that did, gave inconclusive results. Because this study equalized time demands by providing expert explanation in both conditions, it was expected that there were no differences in absolute time. Regarding relative time, it could be speculated that including incorrect examples might be more engaging than only playing correct examples. However, since research on this topic is absent, no particular outcomes were predicted.

Research question 2: What is the effect of a combination of correct and incorrect examples on self-efficacy?

There is insufficient evidence on the increase of self-efficacy by one condition over the other. Therefore, there were no specific predictions.

Research question 3: What is the effect of a combination of correct and incorrect examples on task performance (i.e., practice, immediate, delayed, and transfer performance)?

Because of the partly positive, and partly inconclusive results on learning from incorrect examples, and in particular the combination of correct and incorrect examples, it could be expected that there is either no effect, or a positive effect for the C-I condition. However, the latter expectation seemed most likely, because this research integrated several design features which could in particular improve learning from incorrect examples.

Research question 4: What is the relation between self-regulation and cognitive performance when learning from a combination of correct and incorrect examples?

Since self-regulation has shown to impact learning, it could be expected that higher self-regulation is related to higher task performance. Earlier empirical research did not provide answers to make assumptions on whether self-regulation benefits learning from correct examples or learning from incorrect examples.

Method

Participants

Three primary schools in the east of the Netherlands were selected via convenience sampling. The schools had four 5th grade classes in total, resulting in a total number of 82 participants, with a mean age of 11.2 years. Within the classes, students were randomly assigned to one of the two conditions. A check on the random distribution showed no significant differences between conditions regarding age (11.1 years in the C-C condition and 11.2 years in the C-I condition). Gender was equally distributed over conditions, through gender stratification. Table 1 shows the distribution among conditions for gender and for all students. One male student was removed because he accidently started with videos of the wrong condition, resulting in 40 students in the C-C condition.

The Ethical Committee of the University gave approval for the study. Parents gave active consent in advance, in order for the students to be included into the research. Each teacher will receive a report of the outcomes of each student of their own class.

Table 1

Distribution of gender among conditions

	Gender		
Condition	Male	Female	All
Correct-Correct	23	18	41
Correct-Incorrect	23	18	41
Total	46	36	82

Instructional Materials

The instructional materials were designed specifically for this study. They covered the domain of adding fractions with unequal denominators. Videos were designed to instruct the content; booklets provided procedural instructions, questionnaires, practice and pre- and post-tests. The training consisted of three pairs of videos, each followed by a practice section. The content and design were enhanced through consultation with a math expert and a design expert, and through performing pilot tests for usability with 5th grade learners at several points in time during the design. The design guidelines that were applied are summarized after the sections Videos and Booklets.

Videos. The videos explained how to solve an operation by changing one fraction so that both denominators become the same, hence the fractions can be added easily. For example, in $^1/_2 + {}^1/_4$, the first fraction can be changed into $^2/_4$ leading to the new operation $^2/_4 + {}^1/_4$ with the solution $^3/_4$. Each condition contained six videos, each video presented one operation task. The videos provide a solution procedure, which appeared on screen step-by-step, and was narrated by a model, who was not visible. A link to the videos can be found in Appendix A.

Problem types. Every two videos had a different problem type (see Table 2). The difficulty of the problem types increases, namely, in the beginning, the numerator of the fraction that needed to be converted was 1, which was not the case in the second problem type. In addition, the denominators became more complex in the third problem type: for example, converting a fraction of thirds to twelfths can be viewed as more complex than from thirds to sixths.

Table 2

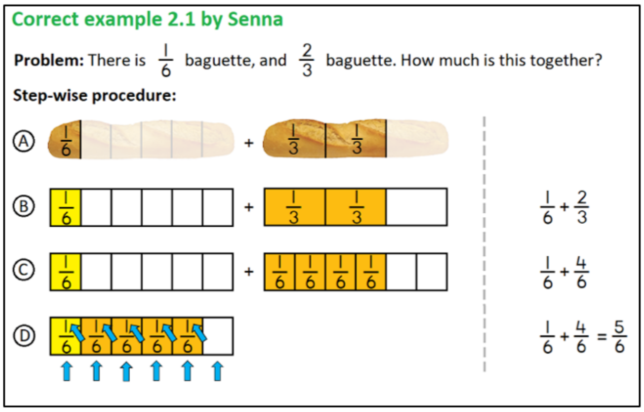
Problem types and corresponding operation task of each video

Problem type	Video number	Operation task
1. Simple numerator, simple denominators	1.1	$\frac{1}{2}+\frac{1}{4}$
	1.2	$\frac{1}{4}+\frac{3}{8}$
2. Complex numerator, simple denominators, different sequence	2.1	$\frac{1}{6}+\frac{2}{3}$
	2.2	$\frac{5}{10}+\frac{2}{5}$
3. Simple numerator, complex denominators	3.1	$\frac{1}{3}+\frac{3}{12}$
	3.2	$\frac{1}{2}+\frac{2}{6}$

Example design. The videos presented both the symbolic representation of the operation, as well as a visual representation that supported conceptual understanding. Figure 1 shows a screenshot of correct video 2.1. The narration of the fictive learner and representation changes of step D are included on the right side. Video 2.1 is the same for both conditions.

The following description of video 2.1 demonstrates the structure and design of all correct videos. A blanc screen was filled step-by-step with the following parts. A green heading showed that the example provided a correct procedure. Next, there was a problem statement. The solution path began with a realistic problem representation including formal symbols (i.e., “There is $^{1}/_{6}$ baguette and $^{2}/_{3}$ baguette, how much is this together?”). Step A shows a real context (photos of baguettes). In step B, the baguette was represented by bars (in two different colours), and the operation was presented on the right side. In Step C, the bar of $^{2}/_{3}$ was converted into $^{4}/_{6}$, after which the converted operation appeared on the right side. In step D, a bar of 6 pieces was presented, and an animation merged the coloured pieces of step C into that bar. After counting the coloured pieces, the numerator 5 appeared at the final answer on the right; and after counting the total amount of pieces, the denominator 6 appeared.

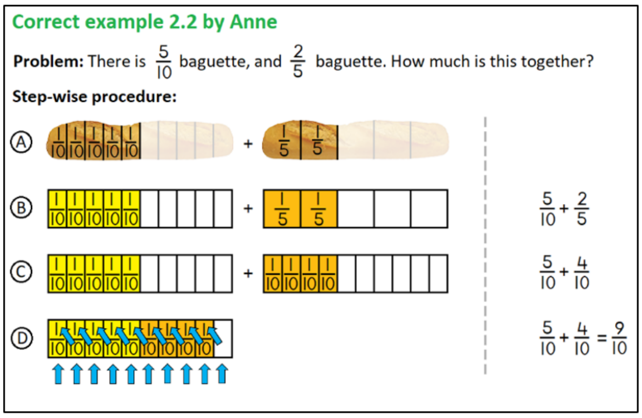
All correct videos had the same underlying structure, yet the surface features were different (i.e., different fractions were used), and the amount of explanation declined. The narration was provided by an expert model (adult) and a peer model (student). The expert introduces the examples, after which the peer reads the problem and explains his steps towards the correct solution.



Narration during step D:	Corresponding representation:
I count and add.	Arrows point to numerators.
1, 2, 3, 4, 5.	
The numerator is 5.	5 appears at symbolic solution.
The whole baguette is divided into 6 pieces.	Arrows point to 6 pieces.
Hence, the denominator is 6.	6 appears at symbolic solution.
My answer is $^{5}/_{6}$.	

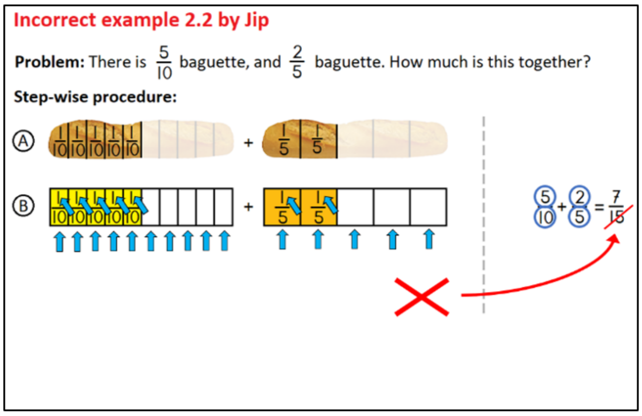
Figure 1. The left side presents the final screen image of video 2.1, including the temporary signalling arrows. The right side presents the narration during step D and the corresponding representation changes.

Incorrect videos. Both conditions contained six videos. The C-C condition consisted of only correct videos, whereas in the C-I condition, three correct videos were replaced by three incorrect videos, using the same operations, yet showing an error in the process. Figure 2 shows correct video 2.2, and figure 3 shows incorrect video 2.2.



Narration during step D:	Corresponding representation:
I count and add. 1, 2, (...), 9.	Arrows point to numerators.
The numerator is 9.	9 appears at symbolic solution.
The whole baguette is divided into 10 pieces.	Arrows point to 10 pieces.
Hence, the denominator is 10.	10 appears at symbolic solution.
My answer is $^{9}/_{10}$.	

Figure 2. The left side presents the final screen image of correct video 2.2, including the temporary signalling arrows. The right side presents the narration during step D and the corresponding representation changes.



Narration during step D:	Corresponding representation:
I count and add. 1, 2, 3, 4, 5.	Arrows point to yellow numerators.
The numerator is 5.	Circle appears on 5 in the operation.
Plus 1, 2.	Arrows point to orange numerators.
The numerator is 2.	Circle appears on 2 in the operation.
Together makes 7.	7 appears at symbolic solution.
Of the 1, 2, (...), 10.	Arrows point to 10 pieces.
10.	Circle appears on 10 in the operation.
Plus 1, 2, 3, 4, 5.	Arrows point to 5 pieces.
5.	Circle appears on 5 in the operation.
Together makes 15.	15 appears at symbolic solution.
My answer is $^{7}/_{15}$.	Red signals appear.

This answer is wrong. You have not divided the baguette into equal pieces, and have not taken them together. And in the end, you counted all pieces of two baguettes, whereas a fraction is a part of one whole, hence a part of one baguette and not of two.

Figure 3. The left side presents the final screen image of incorrect video 2.2, including the temporary signalling arrows and circles. The right side presents the narration during step D and the corresponding representation changes. The italic text is narrated by the expert.

The structure and design of the incorrect videos was the same as of the correct videos. Only the heading was coloured red, and the peer explanation stopped after the error was made. The expert detected the error and explained what was done wrong and why (see italics in Figure 3).

Screenshots of all videos can be found in Appendix B. Table 3 shows which correct videos were replaced by incorrect videos, and it shows the corresponding incorrect solution and to what type of error this solution belonged. The solution in incorrect video 1.2 was wrong, because the fractions were not equalized, and of the two denominators, the highest was chosen. In incorrect video 2.2, again the denominators were not equalized, and the denominators were added. In incorrect video 3.2, the first steps were performed correctly, however, the denominators were added. The common errors were selected based on research about errors in fraction operations performed by Aksoy and Yazlik (2017), Borasi (1987), Eichelmann, Narciss, Schnaubert, and Melis (2012), and Ni and Zhou (2005).

Table 3

Video sequence, replacement of correct with incorrect videos and corresponding types of errors.

Correct-Correct condition	Correct-Incorrect condition	Incorrect solution	Type of error
Video 1.1 correct	Video 1.1 correct		
Video 1.2 correct	Video 1.2 incorrect	$\frac{1}{4}+\frac{3}{8} = \frac{(1+3)}{8} = \frac{4}{8}$	Did not equalize, added numerators, picked highest denominator
Video 2.1 correct	Video 2.1 correct		
Video 2.2 correct	Video 2.2 incorrect	$\frac{5}{10}+\frac{2}{5} = \frac{(5+2)}{(10+5)} = \frac{7}{15}$	Did not equalize, added numerators, added denominators
Video 3.1 correct	Video 3.1 correct		
Video 3.2 correct	Video 3.2 incorrect	$\frac{1}{2}+\frac{2}{6} = \frac{3}{6}+\frac{2}{6} = \frac{(3+2)}{(6+6)} = \frac{5}{12}$	Added numerators, added denominators

Video construction and presentation. The videos had a duration that varied between 2 min and 18 s (incorrect video 1.2) and 3 min and 43 s (correct video 1.1). Appendix C shows the video lengths of all videos. The lengths fit the guideline of Brar and van der Meij (2017): videos should have a maximum *length of 3 to 5 minutes*. Incorrect video 1.2 and incorrect video 2.2 were shorter than their correct equivalents, because the errors occurred halfway through the process. The total duration of all videos in the correct condition was 19 min and 43 s, and of the incorrect condition this was 17 min and 51 s.

Students had access to the videos via a website (there were two different websites, one for each condition). The website consisted of three tabs, each tab consisted of one pair of videos and was labelled (e.g., “Videos 1.1 and 1.2”). Above the pair of videos, a short text instructed the students that they were able to replay, pause, fast-forward, rewind, and watch the videos as often as they wanted to. Below the pair of videos, an instructional text directed the students back to the booklet which contained practice tasks. The tabs were distinguished by the use of three different colours. These colours linked the videos to the corresponding practice, which is referred to as *colour-coding* (Berthold & Renkl, 2009).

Booklets. There were four booklets for each student, containing the questionnaires, practice, tests, and instruction on the online environment. Both conditions received the same booklets. All booklets started with an introduction page, on which the icons that appeared in the booklet were explained (for an example, see Appendix D). The booklets also instructed what was expected of the students, e.g., “This test contains tasks which could be new for you. Do not worry about not understanding these or making mistakes. Try to answer them. We would like to see what you can do”. Instructions on what to do after the tasks were also provided.

The last part of the first booklet consisted of *pre-training*. To prepare for practice, this pre-training taught the students how to divide bars into a certain number of parts (e.g., divide the bar into 5 parts), see Appendix E. Pre-training was not part of the measurements. The second booklet consisted of instructional guidance on how and when to go to the online environment, and when to attend to practice in the booklet (namely, each video pair was followed by paper-and-pencil practice). Screenshots of the online environment were inserted, and through *signalling* (i.e., hairlines), the learners were guided to the correct elements. As an example, Appendix F demonstrates the booklet instructions on entering video 1.1 and 1.2. The second booklet also contained the practice tasks belonging to each video pair. The tasks were preceded by the instruction that the learners were no longer allowed to go back to the corresponding video pair. Booklet 3 and 4 contained a questionnaire and tests.

Learning fractions. The following didactical background served as an essential foundation for the design of the examples and practice, and explains why providing depictive representations was paramount in the present study. The domain adding fractions with unequal denominators is difficult for learners, however, it is an essential foundation for understanding algebra (Wu, 2001). Ni and Zhou (2005) reviewed the complexity of learning fractions and pointed out that learners have trouble with performing symbolic operations with fractions, due to deficits in their conceptual representation (i.e., realistic representation, like a bar), rather than due to difficulty of the symbolic representation. They emphasized that giving meaning to symbolic operations is essential. This can be illustrated with an error often made by learners, as demonstrated by Ball and Wilson (1996) and Mack (1995), namely that verbal questions like "how much is one fifth plus one fifth" often result in the correct answer "two fifths", whereas symbolically, $^1/_5 + ^1/_5$ often leads to the incorrect answer $^2/_{10}$. The importance of giving meaning to symbolic representations is in line with the theory *realistic mathematics education* (RME) (e.g., Van den Heuvel-Panhuizen, & Drijvers, 2014), which describes the value of presenting real and depictive representations in addition to symbolic operations, to establish conceptual understanding. For example, imaginable contexts and representations (like a bar representing a baguette), serve as a foundation for symbolic mathematics (like $^1/_6 + ^2/_3$).

The videos and practice were developed to fit the curriculum of the schools. The 5th graders had already learned to add fractions with equal denominators, and separately, they learned about finding equivalent fractions. Adding fractions with unequal denominators was new to them. All other content was kept simple: there were only operations with a solution under 1; simple fractions were used (hence with denominators up to 12); and only one of the fractions needed to be adjusted in order to obtain equal denominators.

Design guidelines. Several guidelines were used for the design of the videos and practice, an overview is provided in Table 4. A part of the guidelines was already mentioned in the previous sections of Instructional Materials. The other guidelines are elaborated on right now. Videos showed text merged with visualizations, this is in line with the *split-attention principle* (Ayres & Sweller, 2014). In accordance with the *modality principle* (Low & Sweller, 2014), information which did not require visual presentation, was provided orally. For example, the principle "fractions can be added when denominators are equal" was narrated by the student model. To foster profound understanding of the problem, explanation was provided, this design choice is supported by the *explanation-help principle* (Renkl, 2014a). To focus attention on important components, the *signalling principle* was used (van Gog, 2014), e.g., by pointing arrows towards bar parts that needed to be counted. The steps were *numbered* (A to D) to emphasize that the steps were sequential (van der Meij & Gellevij, 2004).

Table 4

Guidelines for the design of the videos

Guideline	Description To advance learning:	Reference
Video length	limit the length of the videos to a maximum of 3-5 minutes	Brar and van der Meij (2017)
Colour coding	connect related elements in separate representations	Berthold and Renkl (2009)
Pre-training	instruct essential characteristics in advance	Mayer and Pilegard (2014)
Split-attention principle	integrate text and visualizations	Ayres and Sweller (2014)
Modality principle	distribute information over visual and auditory channels	Low and Sweller (2014)
Explanation-help principle	provide explanations when self-explaining is difficult	Renkl (2014a)
Signalling principle	use cues to highlight important parts	van Gog (2014)
Numbered steps	number the steps to emphasize succession	van der Meij and Gellevij (2004)

Measurement Instruments

User logs. To gather information about engagement, activity data on the videos was recorded (i.e., playing, pausing, replaying) through a logging program which was connected to the online environment. From the moment the video was set in motion, every second was logged. Two types of measures were computed: relative time and absolute time.

Relative time. This measurement presented percentages of the total number of seconds of a video, with the length of the video serving as the baseline. For example, when 172 s of video 2.1 (215 s) were played, this resulted in a score of 80% (172/215). There were three distinct relative measures. *Play* consisted of the number of seconds that the video was played and replayed. For example, a student could play 80% for the first time and then replay 40% of the video, resulting in a play score of 120%. *Unique play* showed how many seconds of the video were set in motion, without replay, expressed in percentages with a maximum of 100%. E.g., when 129 s were watched of video 2.1 (215 s), and a part of that was replayed, only 129 s of the video was played uniquely, giving a score of 60% (129/215). *Replay* was the number of seconds that were played again, converted to a percentage. Since replay was low, i.e., it had a total mean percentage of 1.6% (*SD* = 5.2), and there was no difference between conditions, replay measurement was not used for further analyses.

Absolute time. This measurement presented the total number of seconds that a video was played. For *total play time*, this meant all played and replayed seconds. For *unique play time*, this meant all played seconds uniquely. To illustrate, when a student played all 19 min and 43 s of all videos together, and he replayed 1 minute, his total play time score would be 20 min and 43 seconds, whereas his unique play time score would be 19 min and 43 seconds. For the same reason as described above, no *replay time* measures were taken into account.

Questionnaires. The questionnaires were administered on paper and are displayed in Appendix G (self-efficacy) and Appendix H (self-regulation).

Self-efficacy. The self-efficacy questionnaire was constructed according to the guidelines of Bandura (2006), and was specifically focussed on the domain, as was argued by Bandura (2006) and Zimmerman (1996). In total, there were 9 items about the leaners' perceived competence regarding the learning domain (e.g., "How good are you at adding fractions?", and "How good are you at computing $^3/_4 - ^1/_4$?"), which were rated on a 7-point Likert scale, ranging from 1 (very good) to 7 (very poor). The questions did not include operations which were used in training and tests. The scores were reversed during analysis to make them easier to read. The minimum test score was 1, and the maximum 9. The scores were converted into percentages. Reliability analysis using Cronbach's alpha led to excellent results for the self-efficacy before test (α = 0.94) and the self-efficacy after test (α = 0.94).

Self-regulation. The self-regulation questionnaire was constructed according to the guidelines of Bandura (2006), and included statements about e.g., planning, concentrating, monitoring, and dealing with difficulties (see Bandura, 2006; Tulis et al., 2016; Zimmerman, 1990). The questionnaire consisted of 7 items about the learners' perceived ability to regulate their learning (e.g., “How good are you at planning your work?”, and “How good are you at recognizing whether something goes right or wrong?”). The questions were scored and converted to percentages, in the same manner as the self-efficacy questions, with a maximum score of 7 (100%). Reliability analysis using Cronbach’s alpha showed good results (α= 0.86).

Performance tests. All performance tests were paper-and-pencil tests. A codebook is included in Appendix I, where all items, answers and coding are displayed.

Pre-test. This test was based on Dutch curriculum guidelines described by Noteboom, Aartsen, & Lit (2017) and Centrum Educatieve Dienstverlening-Groep (n.d.). The test contained 10 items. There were 8 general items which matched the content that was taught in school (e.g., “Which of these fractions are equal to $^3/_{12}$? There is more than one correct answer.”, after which the learners could choose from $^1/_4$ $^2/_3$ $^9/_{48}$ $^6/_{24}$ $^1/_3$). There were 2 items about the content that would be instructed in the videos (e.g., “There is $^1/_3$ cake and $^5/_{12}$ cake. How much is this together? You could draw it.”), without any stepwise or depictive support. Correct items yielded 1 point, incorrect items 0 points. Items with subitems could yield a maximum score of 1 point. The scores were converted to percentages, a maximum score of 10 points resulted in a score of 100%. Reliability was analysed with Cronbach’s alpha and showed a satisfactory score (α = 0.66). A repetition of the 8 general items (containing different fractions) served as a start-up for the immediate post-test, and was used to review whether there were changes between the general fraction knowledge before and after training. There was only a small improvement, and no significant differences were found between conditions. This repetition was not used for further analysis.

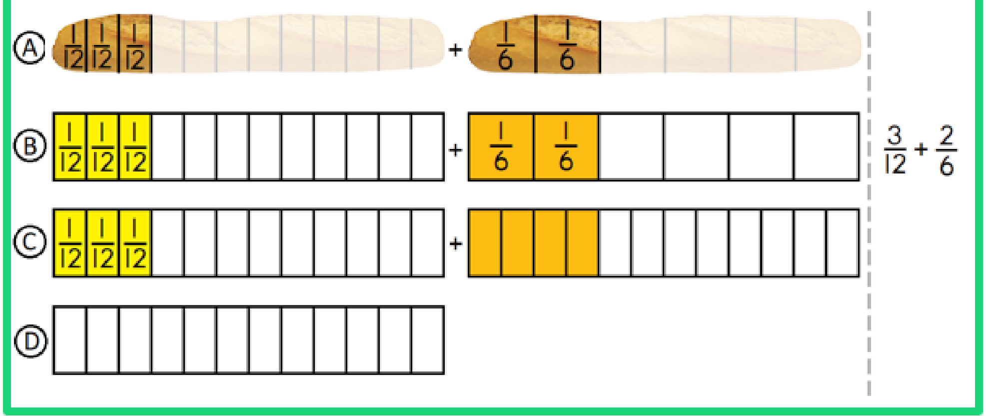
Practice. There were three practice sections. Each section revolved around one practice problem, which had the same operation type as the preceding videos. The first operation was $^1/_5 + ^3/_{10}$, the second $^3/_{12} + ^2/_6$, and the third $^1/_3 + ^2/_9$.

As an example, practice after videos 2.1 and 2.2 is shown in Figure 4. A step-wise procedure with the same underlying structure as the video examples was given. Final steps of the procedure were left out, which is referred to as *incomplete examples*; this gradual transition from example to problem solving is effective for performance (see Renkl, Atkinson, Maier, & Staley, 2002). The learners were guided to fill in the missing steps by answering the questions below the incomplete example.

You will need the green box below to complete the tasks. Now start with task 1 beneath the green box.

Problem: There is $\frac{3}{12}$ baguette, and $\frac{2}{6}$ baguette. How much is this together?

Step-wise procedure:



1. Take a look at step C. What fraction needs to be written in the orange pieces? *Write it down in the orange pieces.*

2. *Write down the new operation of step C next to the dotted line.*

3.1 In how many pieces is the bar of step D divided? *ANSWER: …*
3.2 *Colour the yellow baguette and the orange baguette together in that bar.*
3.3 *Write down the fractions in those pieces.*

4. Which fraction belongs to the coloured pieces of step D? *ANSWER: …*

5. How can you add fractions easily? *Choose 1 answer.*
   - o When the numerators are equal
   - o When the denominators are equal
   - o When the numerator is equal to the denominator

GO TO THE NEXT PAGE OF THIS BOOKLET 9

Figure 4. Practice after video 2.1 and 2.2.

The questions were coded into a total of 26 items, each with a score of 1 for a correct answer and 0 for an incorrect answer. Items with subitems could yield a maximum score of 1 point. The scores were converted to percentages, where the maximum score of 26 was equal to 100%. Reliability was analysed with Cronbach's alpha and showed a good to excellent score (α = 0.89).

Immediate post-test. This test consisted of six problems. An example can be found in Appendix J. The first three problems were completion exercises which resembled the practice during training (containing questions like "Divide the bar into the same number of pieces."). The last three problems were symbolic problems without depictive support (e.g., "Calculate. $^{2}/_{4}$ + $^{1}/_{8}$ = … You could draw it."). The supported problems contained several subitems, which were merged into a smaller set of items. In total, there were 29 items. Correct items yielded 1 point, and incorrect items 0 points. Scores were converted to percentages, with a maximum score of 100% (29 points). Reliability analysis gave an excellent Cronbach's alpha (α= 0.94).

Delayed post-test. The surface features of the delayed post-test differed from the immediate post-test (i.e., other fractions were used), but they had the same underlying structure. The difficulty of the fractions was comparable. The number of items and scoring was identical, and again Cronbach's alpha showed an excellent reliability score (α= 0.93).

Transfer test. The items of the transfer test related to the instructed content, yet they were more complex. They were all fraction arithmetic problems. For example, adding fractions based on a circle representation instead of a bar representation, subtracting unequal fractions (e.g., "Calculate. You could draw it. $^{11}/_{12}$ – $^{2}/_{3}$ = …"), and adding fractions while both fractions needed to be changed. Correct answers yielded 1 point, incorrect answers yielded 0 points. Scores were converted to percentages, with a maximum score of 100% (7 points). Reliability analysis showed that for the transfer test Cronbach's alpha was satisfactory (α = 0.70).

Procedure

The study took place during regular school hours in the students' own classroom, in which both conditions were mixed. Students sat at their own table with a laptop, earplugs, a grey, yellow, and orange pencil, an eraser, and their own reading book. The students were informed to perform tasks individually and they were instructed by the researcher, who is a primary school teacher. Students could only ask for help when a technical problem occurred. The study consisted of two sessions.

In the first session, the students received three numbered booklets. The booklets consisted of several phases, which might only be initiated when indicated by the researcher. There was a time slot for every phase, students were instructed to read their own book when they had time left. Students who were not ready in time, had to stop when the time slot was over. The first booklet started with a practice item of self-efficacy, which was instructed by the researcher, followed by the self-efficacy and self-regulation questionnaire (5 minutes). Next, the prerequisite and prior knowledge test was administered (12 minutes). Then, the students practiced on paper with dividing bars, which was guided step-by-step by the researcher with the use of the interactive whiteboard (5 minutes). Next, the students had to follow the instruction in the second booklet in order to watch the videos and make practice tasks in the booklet. The time slot of all videos plus practice was 45 minutes. A 5-minute break was provided. Then, in the third booklet, the self-efficacy questionnaire was administered again (3 minutes), followed by the prerequisite knowledge test (8 minutes), and the immediate post-test (25 minutes).

One week later, the second session took place. This started with the delayed post-test (25 minutes), and was followed by the transfer test (30 minutes). In between, a short break (3 minutes) was given.

Data Analysis

Assumptions on normality of distribution and homogeneity of variance were tested, which revealed violations for the engagement measures, and all performance measures. Therefore, non-parametric tests are reported for these measurements (i.e., Mann-Whitney *U* test, and Wilcoxon Signed Rank test). Means and standard deviations were computed with the independent samples *t*-test. The self-efficacy and self-regulations had no violations of normality and homogeneity. Therefore, the independent samples *t*-test and paired samples *t*-test were used for these variables. For correlation, non-parametric tests were conducted (i.e., Spearman Rank correlation). All comparisons used two-sided tests with alpha set at 0.05 for significance.

Results

Engagement

Relative measures. The relative measures present what percentage of the videos was played without replay (i.e., unique play) or with replay (i.e., play). In this section, first, analyses are presented of the total amount of unique play for all videos together. Second, it was analysed which condition played (including replay) most of each separate video. And third, a deeper look was taken into each pair of videos (e.g., video 1.1 and 1.2). Per condition, it was analysed what percentage was played uniquely of each first video in a pair, compared to each second video.

Table 5 presents the data for unique play, i.e., the percentage of unique seconds that has been played. Unique play was high in both conditions, in total, 81.6% of the videos was played. Mann-Whitney showed that unique play of three videos was significantly higher in the C-I condition. This was the case for video 1.2 (U = 969.5, p = .042), video 2.1 (U = 981.5, p = .044), and video 2.2 (U = 1067.0, p = .008).

A table with the results on play, i.e., the total percentage of how much of the video was played and replayed, can be found in appendix K. Results of the Mann-Whitney test on play showed a significant difference between conditions for video 2.2. For the C-I condition the play rate was 89.3% (SD = 28.2), and for the C-C condition this was 68.3% (SD = 37.8) (U = 1080.5, p = .005).

Comparisons of unique play between every first and second video in a pair (e.g., the difference between video 1.1 and video 1.2), showed that in the C-I condition, the unique play percentage did not significantly decline between the first videos in each pair (i.e., the correct videos), and the second videos in each pair (i.e., the incorrect videos). Namely, the Wilcoxon Signed Rank test showed no significant differences between video 1.1 and 1.2 (Z = -0.8, p = .398), video 2.1 and 2.2 (Z = -0.9, p = .362), and video 3.1 and 3.2 (Z = -1.8, p = .070). On the contrary, in the C-C condition, every second video was played less than the first, i.e., a significant difference was found between video 1.1 and 1.2 (Z = -2.5, p = .013), video 2.1 and 2.2 (Z = -3.1, p = .002), and video 3.1 and 3.2 (Z = -2.6, p = .010).

Table 5

Mean percentages of correct scores and standard deviations on unique play

	Video 1.1		Video 1.2		Video 2.1		Video 2.2		Video 3.1		Video 3.2		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct-Correct (n = 40)	97.9	15.7	84.6	31.2	83.7	27.0	67.2	38.0	71.8	35.3	57.8	40.1	77.2	21.5
Correct-Incorrect (n = 41)	93.5	22.6	90.6	29.1	92.6	20.1	88.8	27.9	78.5	33.4	71.1	41.2	85.8	18.1
Total (N = 81)	95.7	19.5	87.6	30.1	88.2	24.0	78.1	34.8	75.2	34.3	64.5	41.0	81.6	20.2

Absolute measures. Analyses of absolute measures, i.e., the absolute number of seconds that was played (and replayed), gave no significant differences between conditions. Namely, the Mann Whitney test for absolute unique playtime scores (i.e., seconds played uniquely) gave U = 742.0, p = .459, and for absolute total playtime scores (i.e., seconds played and replayed) gave U = 777.5, p = .688. Table 6 shows the absolute total playtime scores. In appendix L, absolute unique playtime scores can be found.

Table 6

Absolute total playtime scores and standard deviations in seconds

	Video 1.1		Video 1.2[a]		Video 2.1		Video 2.2[a]		Video 3.1		Video 3.2[a]		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct-Correct (n = 40)	233	44	190	70	194	73	143	79	124	59	87	59	972	233
Correct-Incorrect (n = 41)	248	104	128	43	203	47	130	41	135	57	129	75	973	199
Total (N = 81)	240	80	159	65	199	61	136	63	130	58	109	70	972	215

[a]Videos 1.2, 2.2, and 3.2 have different lengths in each condition. For an overview of all video lengths, see Appendix B.

Self-Efficacy

Data on self-efficacy (see Table 7) revealed a high mean score of 4.94 (*SD* = 1.11) before training and an even higher score of 5.19 (*SD* = 1.07) after training. Although self-efficacy before was higher for the C-I condition, this difference was not significant. Self-efficacy after was almost identical between the conditions, no significant differences were found. However, a paired samples t-test revealed that the *increase* from self-efficacy before to self-efficacy after was significant for the C-C condition ($t(39) = -3.95$, $p < .001$.), but not for the C-I condition ($t(40) = -1.67$, $p = .103$).

Table 7

Mean scores and standard deviations for self-efficacy before and after training and for self-regulation

Condition	Self-efficacy before		Self-efficacy after		Self-regulation	
	M	*SD*	*M*	*SD*	*M*	*SD*
Correct-Correct (*n* = 40)	4.85	1.15	5.20	1.01	5.06	0.88
Correct-Incorrect (*n* = 41)	5.02	1.07	5.18	1.13	5.42	0.80
Total (*N* = 81)	4.94	1.11	5.19	1.07	5.24	0.85

Note. Mean scores range from 1 to 7; a higher score means a higher perception.

Task Performance

Performance outcomes are displayed in Table 8. The Wilcoxon Signed Rank test revealed significant increases between the total means of pre-test and practice in the C-C condition ($Z = 4.74$, $p < .001$), as well as in the C-I condition ($Z = 3.51$, $p < .001$). This was also the case between the pre-test and immediate post-test in the C-C condition ($Z = 4.34$, $p < .001$) and the C-I Condition ($Z = 2.92$, $p = .003$), and between the pre-test and delayed post-test in the C-C condition ($Z = 4.78$, $p < .001$) and C-I condition ($Z = 4.35$, $p < .001$).

The increase between the immediate post-test and delayed post-test was significant in the C-I condition ($Z = 2.35$, $p = .019$), yet not in the C-C condition ($Z = 1.15$, $p = .248$).

Table 8

Mean percentages of correct scores and standard deviations on the performance measures

Condition	Pre-test		Practice		Immediate post-test		Delayed post-test		Transfer test	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct-Correct (*n* = 40)	59.3	22.1	79.9	18.2	80.3	22.3	82.7	20.2	33.6	27.0
Correct-Incorrect (*n* = 41)	61.7	16.7	75.4	22.3	73.7	26.7	79.9	22.6	29.6	21.8
Total (*N* = 81)	60.5	19.4	77.6	20.4	77.0	24.7	81.3	21.3	31.6	24.4

Self-Regulation and Correlations

Results of self-regulation are shown in Table 7. Self-regulation gave high scores (M = 5.24, SD = 0.85). Table 9 shows the Spearman Rank correlation coefficients for all possible pairs of the measurement variables of both conditions together. Tables of the two separate conditions can be found in Appendix M. A significant moderate positive correlation was found for the C-C condition between self-regulation and performance tests before and after training (e.g., for the delayed test $r_s(39) = .580$, $p < .001$), however, there was no significant correlation for the C-I condition (for the delayed test $r_s(40) = .237$, $p = .135$). In addition, self-regulation did not significantly correlate with practice in either one of the conditions.

Table 9

Summary of Spearman Rank correlations for all possible pairs

Measure	1 SEb	2 SR	3 Pre	4 Play	5 Uplay	6 A. Pl	7 A. Upl	8 Prac	9 SEa	10 Imm	11 Delay	12 Trans
1. Self-efficacy before	—	.52**	.40**	-.14	-.16	-.19	-.21	.15	.83**	.27**	.38**	.41**
2. Self-regulation		—	.34**	-.21	-.19	-.25*	-.24*	.20	.57**	.25*	.37**	.43**
3. Pre-test			—	-.06	.00	-.09	-.03	.48*	.42**	.45**	.41**	.48**
4. Play				—	.91**	.94**	.87**	-.01	-.12	.06	.01	-.05
5. Unique play					—	.83**	.93**	.12	-.10	.11	.04	-.03
6. Absolute playtime						—	.90**	-.04	-.15	.06	-.04	-.06
7. Absolute unique playtime							—	.13	-.133	.12	.02	-.02
8. Practice								—	.27*	.64**	.67**	.54**
9. Self-efficacy after									—	.37**	.47**	.49**
10. Immediate post-test										—	.73**	.60**
11 Delayed post-test											—	.62**
12 Transfer post-test												—

Note. N = 81. In the column heads, abbreviations are used. They correspond to the terms in the leftmost column.
*p < .05. **p < .01.

Discussion and Conclusion

The present study investigated how the combination of correct and incorrect worked modelling examples would affect performance (on practice, immediate post-test, delayed post-test, and transfer test), engagement, and self-efficacy, and how self-regulation related to performance outcomes when learning from such examples. The examples were compared to correct worked modelling examples. All examples were carefully designed in the form of instructional videos in the domain of adding fractions.

Learning fractions arithmetic is essential for later understanding of mathematics (Lortie-Forgues, Tian, & Siegler, 2015; Wu, 2001). Learners often make errors while learning fractions (Lortie-Forgues et al., 2015). The design of the present study aimed to contribute to learning by making use of errors, specifically, in combination with correct examples. The design of the incorrect examples was intended to amplify their effect, by making use of worked-out steps in combination with modelling examples; by using depictive representations in addition to descriptive representations; by providing instructional explanations; and by making use of design guidelines of worked examples and learning from videos. Key findings are high engagement, especially in the C-I condition, and performance gains in both conditions.

Effects on Engagement

In both conditions, engagement was high, which suggests that the videos supported the attention of the students. Several design features might have contributed to this finding. First, the use of a *human voice* in the examples might have influenced engagement. According to the *voice principle*, a human voice is more effective than a machine voice or no voice (Mayer, 2014). Social cues (such as voice) can elicit a social reaction, i.e., they can activate the learner to attend to the content. Second, important aspects of attention are *space* and *time* (Smith & Kosslyn, 2009). The amount of information a person can attend to at a moment is limited: space concerns the amount information one can absorb at once. The same holds for time, there are limitations to the speed the information can be presented, in order for the learner to absorb the information. In the examples in the present study, the steps were not all presented at once, but they gradually appeared on screen; and the pace of providing information was tailored to the learners with the use of pilot testing, and was gradually increased during the learning process. Third, the use of depictive representations rather than only descriptive representations might have attracted the students. This is in line with Sung and Mayer (2012), who found that students liked learning material better with pictures than without. And fourth, mentioning the problem issue (i.e., "Fractions can be added when denominators are equal. Right now, they are not") might have aroused learners to discover how this issue is solved.

Although these design features were the same for both conditions, the C-I condition showed more engagement on several analyses. It was found that the percentage that was played of three videos was substantially higher in the C-I condition. Moreover, in the C-I condition, the percentage that was played of each second video in a pair did not decline much, whereas this decline was substantial in the C-C condition.

A possible explanation for this might be that viewing an incorrect example could provoke interest in discovering what goes wrong. This is in line with the argumentation of VanLehn, Siler, Murray, Yamauchi, & Bagget (2003), who emphasize that incorporating errors is believed to motivate and activate the learner. Another possible explanation is that it might be more involving to watch a different type of video (i.e., an incorrect video) instead of another correct video. This variation might arouse curiosity. Variation is an essential aspect of the design of the present study, namely, incorrect examples are combined with correct examples, rather than one of these alone.

Regarding absolute time spent on the examples (i.e., the total number of minutes and seconds played and replayed of the video), the outcomes in both conditions were similar. The design of the incorrect examples in the present study might have contributed to restraining the amount of time that incorrect examples consume. Namely, incorrect examples could demand more time of the learners, as was found by McLaren et al. (2016) and Isotani et al. (2011). However, they required the learners to find, fix or explain the errors. The present study located and explained the error, which presumably decreased necessary learning time.

Effects on Self-Efficacy

Self-efficacy was high in both conditions before training, and was even higher after training. This is an important finding because self-efficacy is important for academic motivation and learning outcomes (see Zimmerman, 1996).

The increase was substantial in the C-C condition, but not in the C-I condition. This contrasts with theories of Bandura (1977) and Schunk (1987), and findings of Huang (2017). They suggested that self-efficacy increases more from observing coping models than observing mastery models. This incongruity might be explained by differences between coping modelling examples and the incorrect examples in the present study. Coping models show signs of insecurity, and eventually attain the correct solution of the particular problem. That was not the case in the present study, where no correct solution was provided in the incorrect examples; correct solutions only appeared in the correct examples. Because observing a model accomplishing a task increases one's competence belief (Bandura, 2012), it might be that final accomplishment in incorrect examples is influential in the stimulation of self-efficacy.

Effects on Task Performance

The data could not find support for performance benefits of the combination of correct and incorrect worked modelling examples over correct worked modelling examples. Nevertheless, it was found that both conditions substantially increased performance from pre-test to practice and post-tests. The results indicate that both types of worked modelling examples are effective for task performance. The equal findings between conditions matches part of the previous research, which neither found differences between a combination of correct and incorrect examples and only correct examples (e.g., Isotani et al., 2011; Schunk & Hanson, 1985).

It was however expected that support in understanding how to solve abstract equations by using more explanatory, depictive representations would benefit learning from errors. That is, with realistic support, it would be better understood why an error is wrong and how a correct procedure is performed. Combining realistic and more abstract (symbolic) representations has repeatedly shown to advance learning, specifically when realistic representations are gradually faded (e.g., see Pashler et al., 2007). The use of errors in mathematics could foster understanding (Borasi, 1987), which could lead to better performance.

However, since the C-C condition also increased their performance, it might be assumed that students with correct worked modelling examples also benefitted from this support in understanding. Probably, learners who kept making errors were not able to utilize the opportunity to learn from errors. This is in line with the argument of VanLehn et al. (2003), who describe that a possibility to learn something does not necessarily mean that this has been taken advantage of. To make use of the learning opportunity (i.e., an error), metacognitive skills and motivational aspects are important (Tulis et al., 2016). Perhaps, if learners would have been provided with correctness feedback, they might have benefitted more from errors. Presumably, feedback makes learners aware of their errors and supports them in learning from errors in the examples. Feedback can be helpful for learning (Hattie & Timperley, 2007). Empirical studies on the combination of correct and incorrect examples which used feedback found positive results of this combination on performance (Booth et al., 2013; Durkin & Rittle-Johnson, 2012). Whether feedback is essential, requires further investigation.

Because the C-C condition also seemed to have gained enough understanding to be able to perform well, it would also be interesting to see what the effect is of the combination of correct and incorrect worked modelling examples on performance, when content becomes more complex. For example, when besides adding fractions, learners also have gained knowledge on multiplying fractions. Solving these two different operations requires different approaches. The different procedures are confusing for learners, hence, more errors are made (Lortie-Forgues et al., 2015). It could be that learning from the combination of correct and incorrect modelling examples could be even more fruitful when content becomes more complex.

The C-I condition had a slight, yet not substantial, dip after practice (i.e., on the immediate post-test). A possible reason could be that the incorrect videos cost more effort of the learners. It would be interesting to see if learning from incorrect examples really requires more effort, i.e., by measuring *cognitive load*, which is prominent in worked example research (van Gog & Rummel, 2010), and to examine whether and how this affects performance. A related reason might be fatigue. The tests before training, the instructions, and the training took 1 hr and 15 min. Probably all students were tired after the tests and training, and probably were less motivated to perform on the immediate post-test. Because the C-C condition did not show much difference between practice and immediate post-test either, fatigue might have played a role in both conditions. Future research should consider how the amount of content and time could be reduced or segregated in order to be optimal for learning.

Relations between Self-Regulation and Task Performance

It was expected that higher self-regulation was positively related to higher task performance, because using self-regulation strategies has impact on performance (see Bandura, 2006; Pintrich, 2000; Schunk, 2005; Zimmerman, 1990). This was confirmed for the C-C condition, since there was a positive relation between perceived self-regulation and performance before and after training (i.e., pre-test, immediate post-test, delayed post-test, and transfer test). There was no substantial relation between self-regulation and performance in the C-I condition. This was unexpected, as it is believed that self-regulation is an important factor in learning from errors (Tulis et al., 2016). Therefore, more research can be devoted to which circumstances are important for learners to benefit from self-regulation skills, and whether and how performance can benefit from self-regulation while learning from incorrect examples.

Limitations

Besides the limitations already described, some limitations have not been pointed out, or require more attention. First, the incorrect examples did not provide the correct solution of the particular problem. This was done to avoid redundancy. Namely, the preceding example already provided a correct procedure and solution, and concerned the same problem type as the incorrect example. Furthermore, the incorrect examples also provide a correct reasoning, until the error was made. In addition, by avoiding redundancy, absolute time demands were decreased. Excluding a correct solution in incorrect examples was similar to a part of the empirical studies which provided a combination of correct and incorrect worked examples (Booth et al., 2013; Zhao & Acosta-Tello, 2016). Other studies showed both the correct solution as the incorrect solution to a particular problem, and found positive and mixed results regarding incorrect examples (Durkin & Rittle-Johnson, 2012; Große & Renkl, 2007; Isotani et al., 2011). The latter form is similar to coping modelling examples, where the model works towards the correct solution. However, in coping modelling examples, the observer hears the model struggle (saying it is difficult, showing insecurities), but in the end the model manages to get to the right solution and increases confidence. This might have been a stimulator for observers' self-efficacy, and it might have resulted in higher performance, because the correct solution on the same problem is shown. Future research could investigate whether it is more effective (and not less efficient) to include a correct solution at the end of an incorrect example.

However, the fact that the incorrect examples were followed by expert comments on what went wrong presumably did not hamper self-efficacy, since the self-efficacy level of the C-I condition after training was very similar to the self-efficacy in of the C-C condition. It would be interesting to see the effects on self-efficacy if the incorrect solution would end with a correct solution, instead of putting focus on what should not be done.

Second, the explanation of the model in the correct example was not followed by comments of an expert. Consequences of including or excluding such comments can be a topic of investigation. The present study used a peer model and an expert (adult) model together in one example, which might have influenced outcomes. Since studies investigated the impact of these models separately (e.g., see van Gog & Rummel, 2010), the impact of the current combination of an adult and a peer can be interesting for future research.

Third, the somewhat lower scores on the transfer test suggest that items might have been too difficult for a part of the learners. Future research might construct transfer items that measure related content that is not too complex.

Future Directions

There is more experimental investigation required on the combination of correct and incorrect worked modelling examples and the impact on learning, for example on what factors contribute to engagement, what the influence is of the design choices of the present study, and what the influence is of combining correct with incorrect examples rather than incorrect examples alone. It can be examined if the findings of the present study are generalizable to other areas than mathematics, and other populations than primary school students. In addition, more examination on the causal relationships between the measured variables, for example, self-regulation on performance, is of interest. Finally, a deeper look can be taken into the type of errors learners commit. Do they still commit the exemplified errors? Or are there other misconceptions, or random errors? This might illuminate possible effects of the exemplified errors.

To conclude, this study consisted of many novel features (e.g., an aggregation of worked and modelling examples, a combination of correct with incorrect examples, measuring motivational factors), and found benefits especially on engagement. The study was conducted in a real-world environment. It brings inspiration to learning from errors, and hopes to promote inquiry into incorrect worked modelling examples, in order to stimulate education to profit from errors, in an efficient, effective, and involving manner.

References

Adams, D. M., McLaren, B. M., Durkin, K., Mayer, R. E., Rittle-Johnson, B., Isotani, S., & van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior*, *36*, 401-411. doi:10.1016/j.chb.2014.03.053

Aksoy, N. C., & Yazlik, D. O. (2017). Student errors in fractions and possible causes of these errors. *Journal of Education and Training Studies*, *5*(11), 219-233. doi:10.11114/jets.v5i11.2679

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181-214. doi:10.3102/00346543070002181

Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 206-226). New York, NY: Cambridge University Press.

Baldwin, T. T. (1992). Effects of alternative modeling strategies on outcomes of interpersonal-skills training. *Journal of Applied Psychology*, *77*(2), 147-154. doi:10.1037/0021-9010.77.2.147

Ball, D. L., & Wilson, S. M. (1996). Integrity in teaching: Recognizing the fusion of the moral and intellectual. *American Educational Research Journal*, *33*(1), 155-192. doi:10.2307/1163384

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215. Retrieved from https://www.apa.org/

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age Publishing.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9-44. doi:10.1177/0149206311410606

Barbieri, C., & Booth, J. L. (2016). Support for struggling students in algebra: Contributions of incorrect worked examples. *Learning and Individual Differences*, *48*, 36-44. doi:10.1016/j.lindif.2016.04.001

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), *70*. doi:10.1037/a0013247

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

Borasi, R. (1987). Exploring mathematics through the analysis of errors. *For the Learning of Mathematics*, *7*(3), 2-8. Retrieved from https://flm-journal.org/

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, *94*(2), 405-415. doi:10.1037//0022-0663.94.2.405

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education* (pp. 61-100). Retrieved from https://aaalab.stanford.edu

Brar, J., & van der Meij, H. (2017). Complex software training: Harnessing and optimizing video instruction. *Computers in Human Behavior*, *70*, 475-485. doi:10.1016/j.chb.2017.01.014

Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 174-205). New York, NY: Cambridge University Press.

Centrum Educatieve Dienstverlening-Groep. (n.d.). *Basisleerlijn rekenen.* Retrieved from http://www.leerlijnen.cedgroep.nl/po-sbo/leerlijnen-po-sbo.aspx

Confrey, J. (1990). A review of the research on student conceptions in mathematics, science, and programming. *Review of Research in Education, 16*, 3-56. doi:10.2307/1167350

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83-87. doi:10.1111/1467-8721.01235

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, *22*(3), 206-214. doi:10.1016/j.learninstruc.2011.11.001

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. doi:10.1007/s13138-011-0031-5

Große, C. S. (2018). Copying allowed - but be careful, errors included! Effects of copying correct and incorrect solutions on learning outcomes. *Learning and Instruction*, *58*, 173-181. doi:10.1016/j.learninstruc.2018.06.004

Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes?. *Learning and Instruction*, *17*(6), 612-634. doi:10.1016/j.learninstruc.2007.09.008

Hadjidemetriou, C., & Williams, J. (2002). Children's graphical conceptions. *Research in Mathematics Education*, *4*(1), 69-87. doi:10.1080/14794800008520103

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

Heemsoth, T., & Heinze, A. (2014). The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instructional Science*, *42*(4), 639-657. doi:10.1007/s11251-013-9302-5

Hoogerheide, V. (2016). *Effects of observing and creating video modeling examples on cognitive and motivational aspects of learning* (Doctoral dissertation). doi:10.13140/RG.2.2.32650.88004

Hoogerheide, V., Loyens, S. M., & van Gog, T. (2014). Comparing the effects of worked examples and modeling examples on learning. *Computers in Human Behavior*, *41*, 80-91. doi:10.1016/j.chb.2014.09.013

Huang, X. (2017). Example-based learning: Effects of different types of examples on student performance, cognitive load and self-efficacy in a statistical learning task. *Interactive Learning Environments*, *25*(3), 283-294. doi:10.1080/10494820.2015.1121154

Isotani, S., Adams, D., Mayer, R. E., Durkin, K., Rittle-Johnson, B., & McLaren, B. M. (2011). Can erroneous examples help middle-school students learn decimals? In C. D. Kloos, D. Gillet, R. M. Crespo García, F. Wild, & M. Wolpers (Series Eds.), *Lecture Notes in Computer Science*: Vol. 6964. Towards Ubiquitous Learning (pp. 181-195). doi:10.1007/978-3-642-23985-4_15

Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology*, *92*(4), 811. doi:10.1037//0022-0663.92.4.811

Kopp, V., Stark, R., & Fischer, M. R. (2008). Fostering diagnostic knowledge through computer-supported, case-based worked examples: Effects of erroneous examples and feedback. *Medical Education*, *42*(8), 823-829. doi:10.1111/j.1365-2923.2008.03122.x

Lauzier, M., & Haccoun, R. R. (2014). The interactive effect of modeling strategies and goal orientations on affective, motivational, and behavioral training outcomes. *Performance Improvement Quarterly*, *27*(2), 83-102. doi:10.1002/piq.21169

Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review*, *38*, 201-221. doi:10.1016/j.dr.2015.07.008

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 227-246). New York, NY: Cambridge University Press.

Mack, N. K. (1995). Confounding whole-number and fraction concepts when building on informal knowledge. *Journal for Research in Mathematics Education*, 422-441. doi:10.2307/749431

Mayer, R. E. (2014). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 345-368). New York, NY: Cambridge University Press.

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* (2nd ed., pp. 316–344). New York, NY: Cambridge University Press.

McLaren, B. M., Adams, D., Durkin, K., Goguadze, G., Mayer, R. E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S., & van Velsen, M. (2012). To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In *European Conference on Technology Enhanced Learning* (pp. 222-235). Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-33263-0_18

McLaren, B. M., van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behavior*, *55*, 87-99. doi:10.1016/j.chb.2015.08.038

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. doi:10.1207/s15326985ep4001

Noteboom, A., Aartsen, A., & Lit, S. (2017). *Tussendoelen rekenen-wiskunde voor het primair onderwijs: Uitwerkingen van rekendoelen voor groep 2 tot en met 8 op weg naar streefniveau 1S.* Retrieved from https://slo.nl/

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning*: *IES practice guide* (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ncer.ed.gov

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). doi:10.1016/B978-012109890-2/50043-3

Renkl, A. (2014a). The worked examples principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 391-412). New York, NY: Cambridge University Press.

Renkl, A. (2014b). Toward an instructionally oriented theory of example-based learning. *Cognitive science*, *38*(1), 1-37. doi:10.1111/cogs.12086

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education*, *70*(4), 293-315. doi:10.1080/00220970209599510

Schunk, D. H. (1981). Modeling and attributional effects on children’s achievement: A self-efficacy analysis. *Journal of Educational Psychology*, *73*(1), 93-105. doi:10.1037/0022-0663.73.1.93

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of educational research*, *57*(2), 149-174. doi:10.3102/00346543057002149

Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist*, *40*(2), 85-94. doi:10.1207/s15326985ep4002_3

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, *77*(3), 313-322. doi:10.1037//0022-0663.77.3.313

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology*, *79*(1), 54-61. doi:10.1037//0022-0663.79.1.54

Schunk, D. H., & Zimmerman, B. J. (2003). Self-regulation and learning. In W. M. Reynolds, G. E. Miller, & I. B. Weiner (Eds.), *Handbook of psychology* (pp. 59-78). doi:10.1002/0471264385.wei0704

Smith, E. E., & Kosslyn, S. M. (2009). *Cognitive psychology: Mind and brain* (International ed.). Upper Saddle River, NJ: Pearson Education.

Smith, J. P., III, diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, *3*(2), 115-163. Retrieved from https://www.jstor.org

Sung, E., & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*, *28*(5), 1618-1625. doi:10.1016/j.chb.2012.03.026

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and instruction*, *2*(1), 59-89. doi:10.1207/s1532690xci0201

Sweller, J., van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296. doi:10.1023/A:1022193728205

Tsovaltzi, D., McLaren, B. M., Melis, E., & Meyer, A. K. (2012). Erroneous examples: effects on learning fractions in a web-based setting. *International Journal of Technology Enhanced Learning*, *4*(3-4), 191-230. doi:10.1504/IJTEL.2012.051583

Tulis, M., Steuer, G., & Dresel, M. (2016). Learning from errors: A model of individual processes. *Frontline Learning Research*, *4*(2), 12-26. doi:10.14786/flr.v4i2.168

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. *Encyclopedia of Mathematics Education*, 521-525. doi:10.1007/978-94-007-4978-8

van der Meij, H., & Gellevij, M. (2004). The four components of a procedure. *IEEE Transactions on Professional Communication*, *47*(1), 5-14. doi:10.1109/TPC.2004.824292

van Gog, T. (2014). The signaling (or Cueing) principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 263-278). New York, NY: Cambridge University Press.

van Gog, T. (2015). Learning from erroneous examples in medical education. *Medical Education*, *49*(2), 142-144. doi:10.1111/medu.12655

van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review*, *22*(2), 155-174. doi:10.1007/s10648-010-9134-7

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction*, *21*(3), 209-249. doi:10.1207/S1532690XCI2103_01

Wang, M., Yang, Z. K., Liu, S. Y., Cheng, H. N., & Liu, Z. (2015). Using feedback to improve learning: Differentiating between correct and erroneous examples. *2015 International Symposium on Educational Technology (ISET)* (pp. 99-103). doi:10.1109/ISET.2015.28

Wittwer, J., & Renkl, A. (2010). How effective are instructional explanations in example-based learning? A meta-analytic review. *Educational Psychology Review*, *22*(4), 393-409. doi:10.1007/s10648-010-9136-5

Wu, H. (2001). How to prepare students for algebra. *American Educator*, *25*(2), 10-17. Retrieved from https://pdfs.semanticscholar.org

Zhao, H., & Acosta-Tello, E. (2016). The impact of erroneous examples on students’ learning of equation solving. *Journal of Mathematics Education*, *9*(1), 57-68. Retrieved from http://educationforatoz.com/home.html

Zimmerman, B. J. (1990). Self-regulating academic learning and achievement: The emergence of a social cognitive perspective. *Educational Psychology Review*, *2*(2), 173-201. doi:10.1007/BF01322178

Zimmerman, B. J. (1996). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). doi:10.1017/CBO9780511527692

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, *94*(4), 660-668. doi:10.1037//0022-0663.94.4.660

Appendix A

Link to the Videos

The videos can be viewed through the webpage https://go-lab.gw.utwente.nl/videos/201903-marion/

Appendix B

Screenshots of all Videos

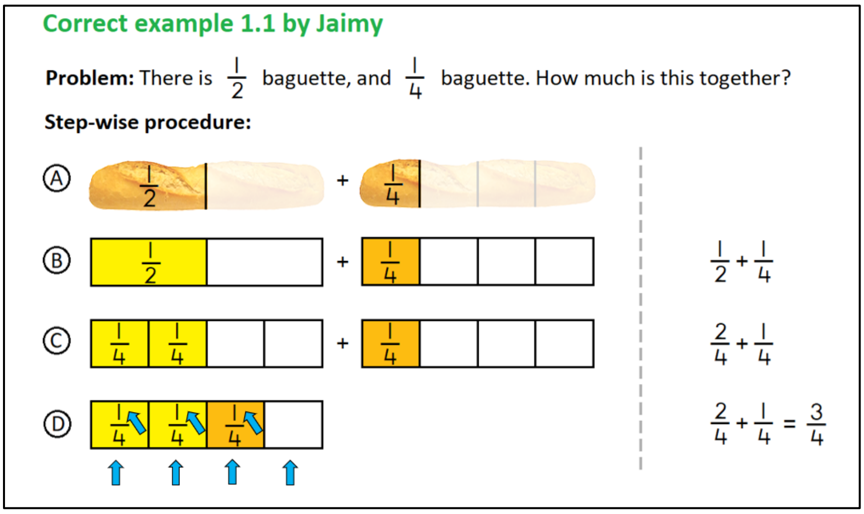


Figure A1. Screenshot of all steps of video 1.1

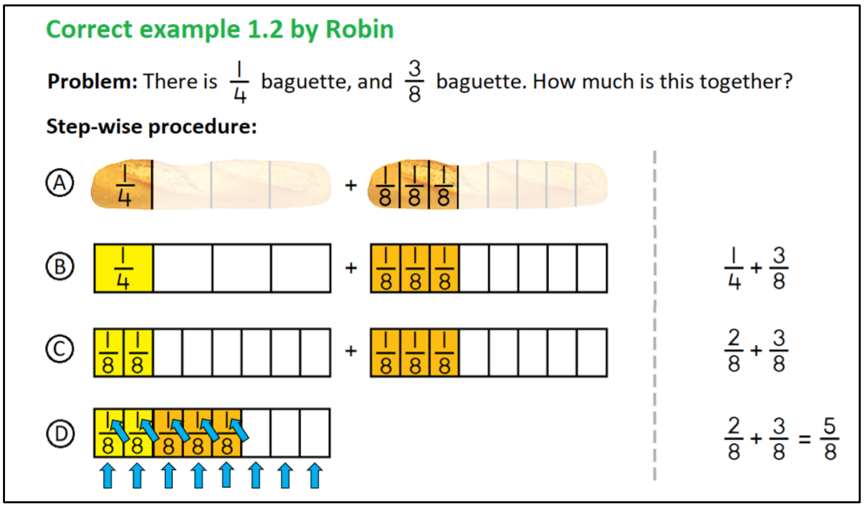


Figure A2. Screenshot of all steps of video 1.2 correct

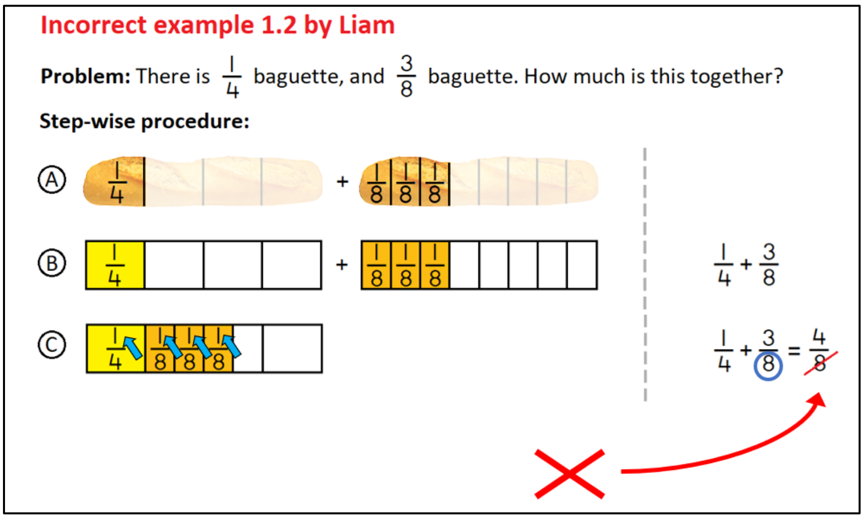


Figure A3. Screenshot of all steps of video 1.2 incorrect

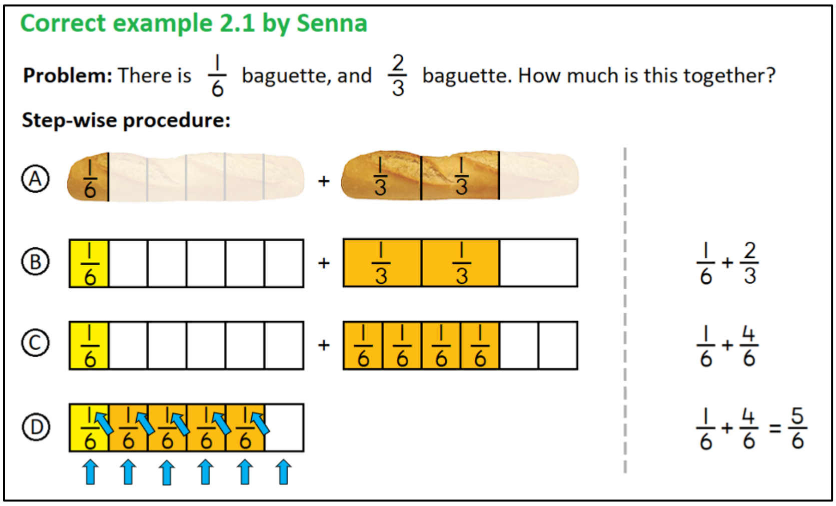


Figure A4. Screenshot of all steps of video 2.1

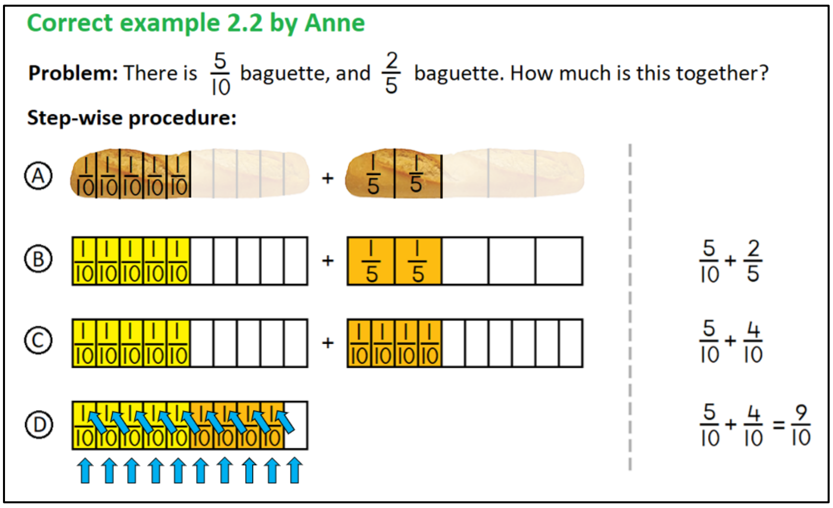


Figure A5. Screenshot of all steps of video 2.2 correct

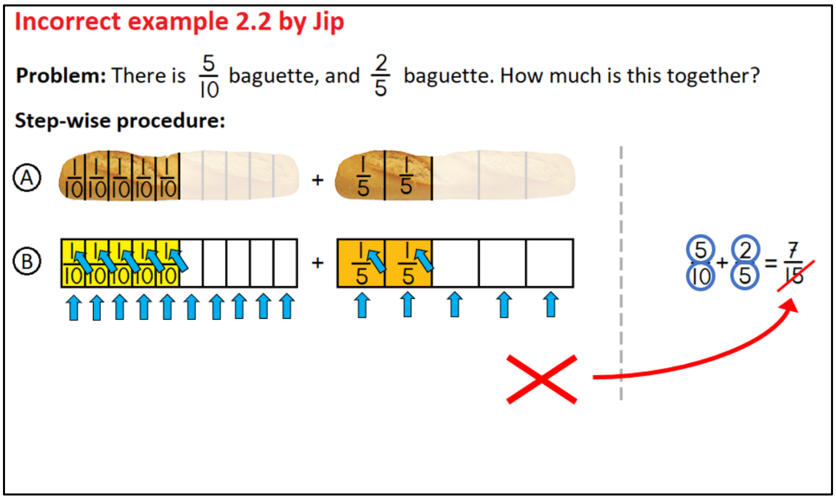


Figure A6. Screenshot of all steps of video 2.2 incorrect

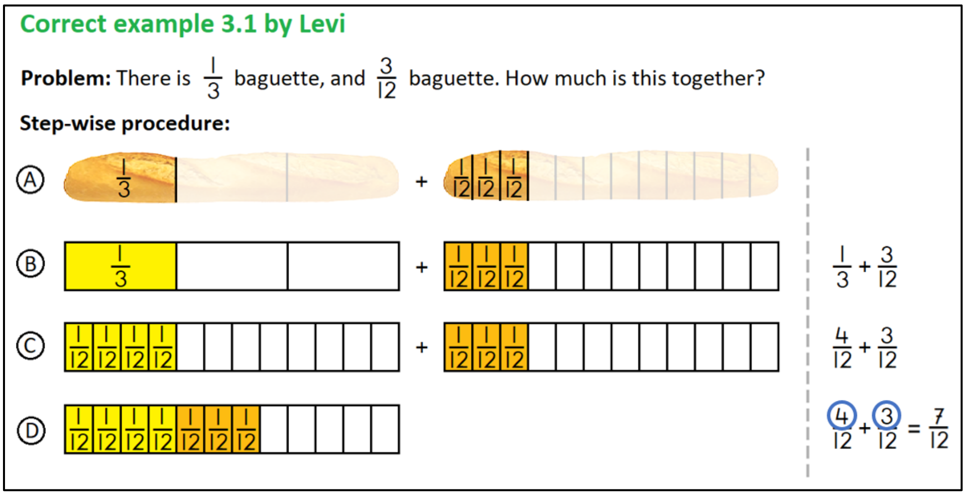


Figure A7. Screenshot of all steps of video 3.1

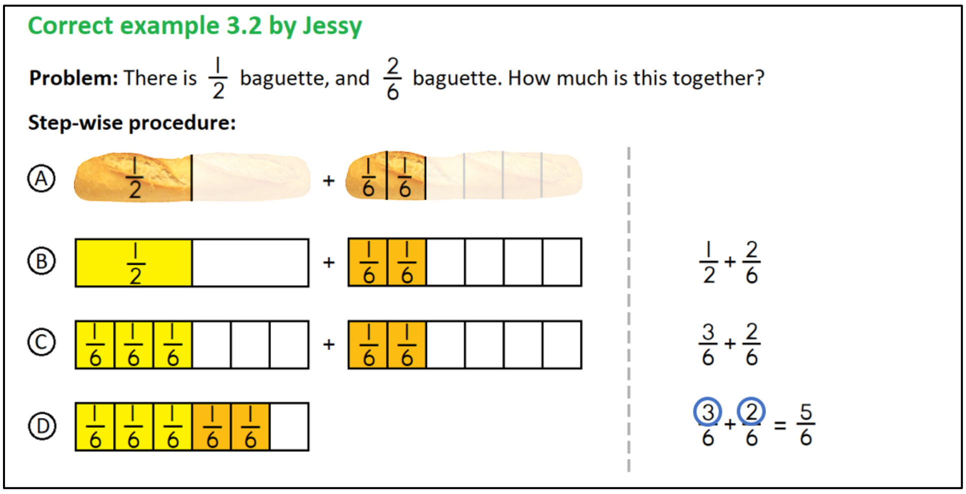


Figure A8. Screenshot of all steps of video 3.2 correct

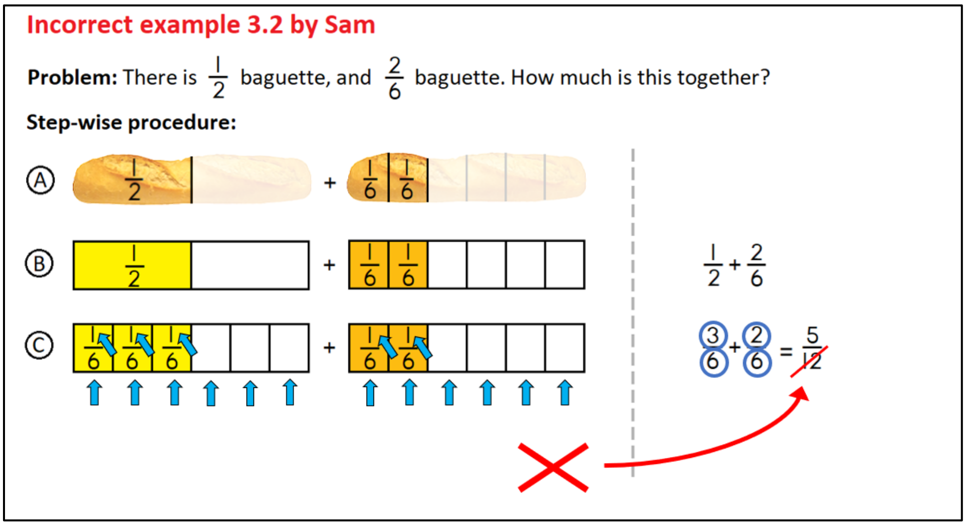


Figure A9. Screenshot of all steps of video 3.2 incorrect

Appendix C

Absolute Video Lengths in Seconds

Video Type	Video 1.1	Video 1.2	Video 2.1	Video 2.2	Video 3.1	Video 3.2	Total
Correct	223	218	215	209	169	149	1183
Incorrect		138		146		180	1071

Appendix D

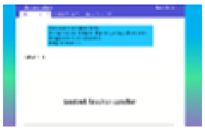
Introduction Page of Booklet 2

Instructions

In a minute, you will watch videos and you will practice in this booklet.
During practice, you are NOT allowed to watch the videos again.

In this booklet, you will encounter the following icons:

This icon means you are going to watch videos.



This icon means you are going to make practice tasks in the booklet.

WAIT UNTIL MISS MARION TELLS YOU TO CONTINUE

1

Appendix E

Pre-Training: Dividing Bars Together with the Instructor

WITH DIVIDING BARS

WAIT to start with this practice page until miss Marion tells you to.

2

4

8

3

6

9

4

12

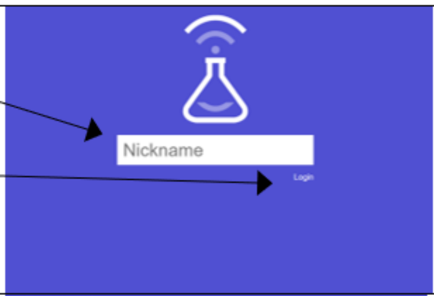
5

10

Appendix F

Instructions on Entering the Online Environment to view Videos 1.1 and 1.2

You will log on:

1. Type at Nickname: **875** Click on login. *A new window will appear.*	
2. Read the light blue text and do what it says.	

Appendix G

Self-Efficacy Questionnaire

How good are you at:

1.	mathematics?	Very good 1	2	3	4	5	6	Very bad 7
2.	adding fractions?	Very good 1	2	3	4	5	6	Very bad 7
3.	subtracting fractions?	Very good 1	2	3	4	5	6	Very bad 7
4.	calculating $\frac{3}{4} - \frac{1}{4}$?	Very good 1	2	3	4	5	6	Very bad 7
5.	adding $1\frac{1}{4}$ and $2\frac{1}{8}$?	Very good 1	2	3	4	5	6	Very bad 7
6.	calculating $\frac{2}{3} - \frac{1}{4}$?	Very good 1	2	3	4	5	6	Very bad 7
7.	adding $\frac{1}{4}$ and $\frac{2}{4}$?	Very good 1	2	3	4	5	6	Very bad 7
8.	adding $\frac{1}{8}$ and $\frac{1}{3}$?	Very good 1	2	3	4	5	6	Very bad 7
9.	calculating $2\frac{3}{5} - 1\frac{1}{5}$?	Very good 1	2	3	4	5	6	Very bad 7

Appendix H

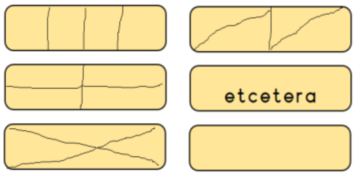
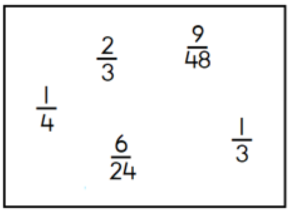
Self-Regulation Questionnaire

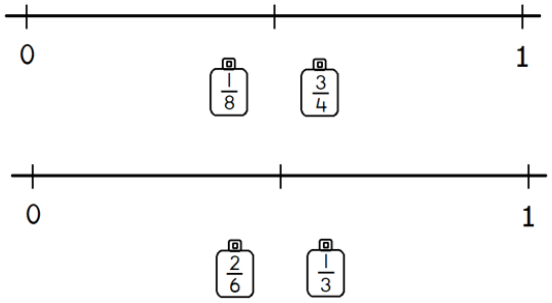
How good are you at:

1.	planning your work?	Very good 1	2	3	4	5	6	Very bad 7
2.	explaining how you solved something?	Very good 1	2	3	4	5	6	Very bad 7
3.	remembering instructions?	Very good 1	2	3	4	5	6	Very bad 7
4.	concentrating during class?	Very good 1	2	3	4	5	6	Very bad 7
5.	solving problems?	Very good 1	2	3	4	5	6	Very bad 7
6.	check your work?	Very good 1	2	3	4	5	6	Very bad 7
7.	noticing if something goes right or wrong?	Very good 1	2	3	4	5	6	Very bad 7

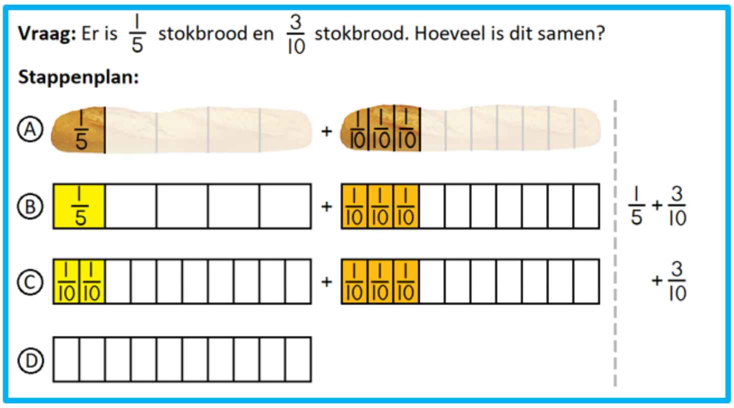
Appendix I

Codebook for all Tests (in Dutch)

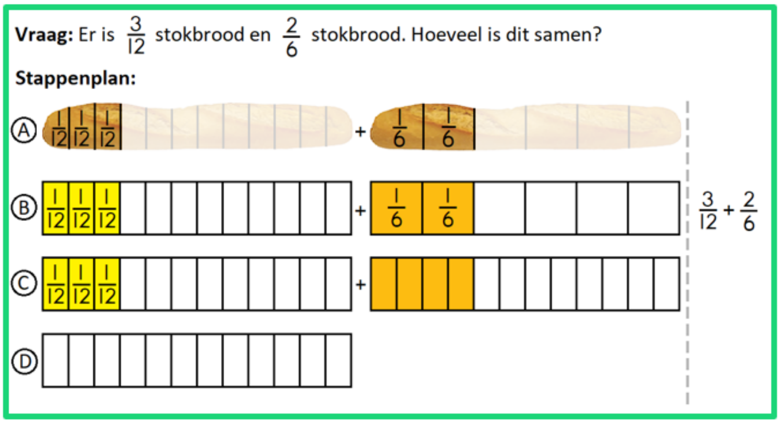
Booklet 1 Pre-test	**antwoord**	**score**
1. Verdeel de cake in kwarten. 	bb1=4 gelijke delen Antwoord =  	1*
2. Zet rondjes om wat evenveel is als $\frac{3}{12}$ *Er zijn er meer goed.*  	bb21=1/4 Antwoord =wel omcirkelen bb22=2/3 Antwoord =niet omcirkelen bb23=9/48 Antwoord =niet omcirkelen bb24=6/24 Antwoord =wel omcirkelen bb25=1/3 Antwoord =niet omcirkelen	1 punt indien 5 antwoorden goed. Anders 0 punten.
3. Schrijf naast het plaatje de breuk op van het deel dat gekleurd is. 	bb3=blauw gedeelte van het geheel Antwoord =5/8	1*
4. Zet een rondje om wat meer is. $\frac{2}{3}$ of $\frac{3}{6}$	bb4=grootste breuk Antwoord =alleen 2/3 omcirkeld	1*
5. Er is $\frac{1}{3}$ cake en $\frac{1}{6}$ cake. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan deze som.*	nb5=uitkomst 1/3 + 1/6 Antwoord =3/6 of 1/2	1*
6. Er is $\frac{1}{3}$ cake en $\frac{5}{12}$ cake. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan.*	nb6=uitkomst 1/3 + 5/12 Antwoord =9/12 of 3/4	1*

7. Trek lijnen van de kaartjes naar de goede plaats op de getallenlijnen. 	bb71=kaartje 1/8 Antwoord =op lijn tussen 1/16 en 3/16 bb72=kaartje 3/4 Antwoord =op lijn tussen 11/16 en 13/16 bb73=kaartje 2/6 Antwoord =op lijn tussen 5/18 en 7/18 bb74=kaartje 1/3 Antwoord =op lijn tussen 5/18 en 7/18	Elk onderdeel 0,25 punt. Alle onderdelen = samen max 1 punt.
8. Schrijf de breuk op. 5 stukjes van $\frac{1}{6}$ is …… 3 stukjes van $\frac{1}{9}$ is ……	bb81=5 keer 1/6 Antwoord =5/6 bb82=3 keer 1/9 Antwoord =3/9 of 1/3	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.
9. Zet in de twee breuken hieronder rondjes om de tellers en pijlen naar de noemers. $\frac{1}{3}$ $\frac{4}{9}$	bb91=teller in 1/3 Antwoord =1 omcirkelen bb92=teller in 4/9 Antwoord =4 omcrikelen bb93=noemer in 1/3 Antwoord =pijl naar 3 bb94=noemer in 4/9 Antwoord =pijl naar 9	Elk onderdeel 0,25 punt. Alle onderdelen = samen max 1 punt.
10. Reken de sommen uit. $\frac{2}{4} + \frac{1}{4} =$ $\frac{4}{6} + \frac{1}{6} =$	bb101=bovenste som Antwoord =3/4 bb102=onderste som Antwoord =5/6	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.

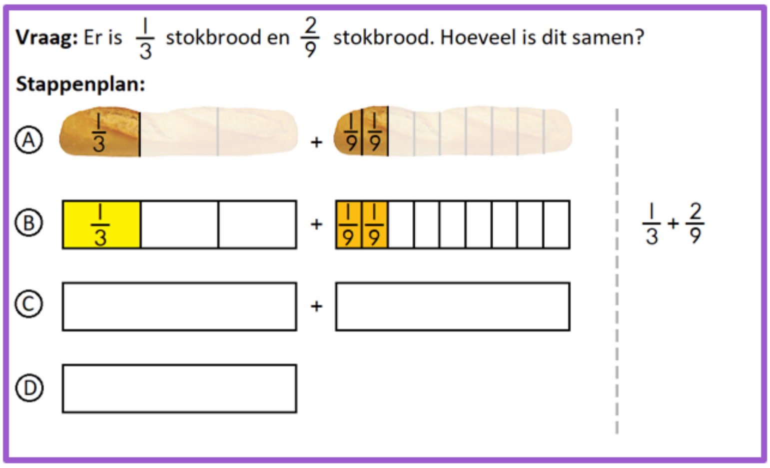
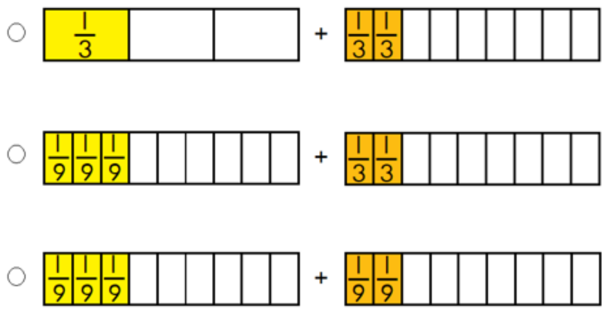
Booklet 2 Practice	**antwoord**	**score**
1. Zet in de breuk hieronder: *een rondje om de teller* *en een pijl naar de noemer.* $\frac{1}{5}$	np111=teller Antwoord =1 np112=noemer Antwoord =5	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.
2. Hoe kun je breuken handig optellen? *Kruis 1 antwoord aan.* ○ Je kunt meteen de cijfers boven de streep en onder de streep optellen. ○ Je kunt meteen tellen hoeveel stukjes je hebt. ○ Je moet eerst zorgen dat ze in evenveel stukjes verdeeld zijn.	np12=MPC Antwoord =3^{e}	1*



3. In welke breuk wordt $\frac{1}{5}$ veranderd? *ANTWOORD*: ...	np13=stap C 1e strook Antwoord =2/10	1*
4. Bij stap C is de nieuwe som naast de stippellijn nog niet helemaal opgeschreven. Wat moet er voor de + staan? *Schrijf het daar op.*	np14 = breuk in blauw vak vóór 3/10 Antwoord = 2/10	1*
5.1 In hoeveel stukken is de strook van stap D verdeeld? *ANTWOORD: ...*	np151 = tellen van stukken in strook D Antwoord = 10	1*
5.2 *Kleur het gele stokbrood en het oranje stokbrood samen in die strook.*	np1521 = kleuren gele stokbrood in strook D Antwoord = 2 stukjes gekleurd np1522 = kleuren oranje stokbrood in strook D Antwoord = 3 stukjes gekleurd	1 punt indien beide antwoorden goed
5.3 *Schrijf de breuken in de gekleurde stukken.*	np1531 = opschrijven breuken in strook D Antwoord = 1/10 staat in de twee stukjes OF 2/10 staat over de stukjes heen geschreven (samengenomen) np1532 = opschrijven breuken in strook D Antwoord = 1/10 staat in de drie stukjes staat 1/10 OF 3/10 staat over de stukjes heen geschreven (samengenomen)	1 punt indien beide antwoorden goed
6. Hoeveel stukjes heb je in totaal gekleurd in de strook bij D? *ANTWOORD*: ...	np16 = aantal stukjes tellen dat gekleurd had moeten worden Antwoord = 5 (indien anders = fout)	1*
7. Welke breuk hoort bij die stukken van stap D? *ANTWOORD*: ...	np17 = totaalbreuk bij stukjes die in D gekleurd/geschreven hadden moeten worden Antwoord = 5/10 OF 1/2 (indien anders = fout)	1*



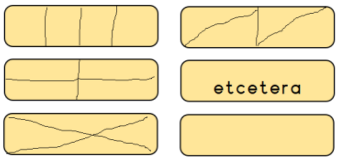
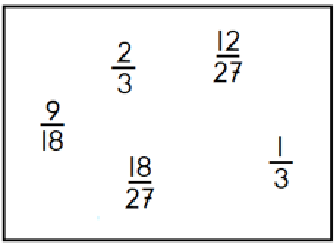
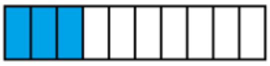
1. Kijk naar stap C. Welke breuk moet er in de oranje stukjes staan? *Schrijf het in de oranje stukjes.*	np21 = breuk bij opgesplitste oranje stukjes in stap C Antwoord = 1/12 in elk stukje OF 4/12 over alle stukjes heengeschreven (samengenomen)	1*
2. *Schrijf de nieuwe som bij stap C naast de stippellijn.*	np221 = breukenstukjes eerste strook Antwoord = 3/12. Mag ook achteraan in som. np222 = breukenstukjes tweede strook Antwoord = 4/12. Mag ook vooraan in som.	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.
3.1 In hoeveel stukken is de strook van stap D verdeeld? *ANTWOORD: …*	np231 = tellen stukjes strook D Antwoord =12	1*
3.2 *Kleur het gele stokbrood en het oranje stokbrood samen in die strook.*	np2321 = kleuren gele stokbrood in strook D Antwoord = 3 stukjes gekleurd np2322 = kleuren oranje stokbrood in strook D Antwoord = 4 stukjes gekleurd	1 punt indien beide antwoorden goed
3.3 *Schrijf de breuken erin.*	np2331 = opschrijven breuken eerste stokbrood in strook D Antwoord = 1/12 in de drie stukjes. OF 3/12 staat over de stukjes heen (samengenomen) np2332 = opschrijven breuken tweede stokbrood in strook D Antwoord = 1/12 in de vier stukjes OF 4/12 staat over de stukjes heen (samengenomen)	1 punt indien beide antwoorden goed
4. Welke breuk hoort bij de gekleurde stukjes van stap D? *ANTWOORD: …*	np24 = totaalbreuk bij stukjes die in D Antwoord = 7/12 (indien anders = fout)	1*
5. Wanneer kun je breuken handig optellen? *Kruis 1 antwoord aan.* o Als de tellers hetzelfde zijn o Als de noemers hetzelfde zijn o Als de teller hetzelfde is als de noemer	np25 = MPC Antwoord = 2e	1*

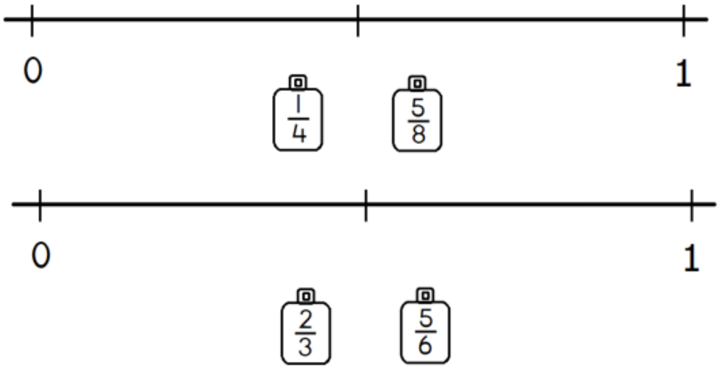
 1. Verander je de breuk $\frac{1}{3}$ of $\frac{2}{9}$? *ANTWOORD*	np31 = de breuk die veranderd kan worden zodat beide noemers gelijk zijn Antwoord = 1/3	1*
2. Waarom verander je die? *Kruis 1 antwoord aan.* ○ omdat $\frac{1}{3}$ maar één stuk is en $\frac{2}{9}$ twee stukken ○ omdat je van $\frac{1}{3}$ wel negenden kunt maken en van $\frac{2}{9}$ geen derden ○ omdat je van $\frac{2}{9}$ wel derden kunt maken en van $\frac{1}{3}$ geen negenden	np32 = MPC Antwoord = 2e	1*
3. Op welke noemer kom je uit? *ANTWOORD*: ...	np33 = de noemer in breuk die je uiteindelijk krijgt, oftewel: het aantal stukjes waarin je de strook verdeelt Antwoord = 9	1*
4. Welk antwoord hoort bij stap C? *Kruis 1 rondje aan.* 	np34 = stroken die beide in evenveel stukjes zijn verdeeld en waarin in elk gekleurd stukje 1/9 staat Antwoord = 3e	1*
5. *Schrijf de nieuwe som bij stap C naast de stippellijn.*	np351 = naast stap C staat de breuk van de eerste strook Antwoord = 3/9. Mag ook achteraan in som. np352 = naast stap C staat de breuk van de tweede strook Antwoord = 2/9. Mag ook vooraan in som.	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.
6.1 *Verdeel de strook bij stap D in het goede aantal stukjes.*	np361 = strook D in 9 stukjes verdelen, net als bij strook C. Antwoord = 9. Goed indien het 9 stukjes zijn, ook al zijn ze niet heel gelijk verdeeld.	1*

6.2 *Kleur de stokbroden erin.*	np3621 = kleuren gele stokbrood in strook D Antwoord = 3 stukjes gekleurd. Alleen goed als np361 goed is. np3622 = kleuren oranje stokbrood in strook D Antwoord = 2 stukjes gekleurd. Alleen goed als np361 goed is.	1 punt indien beide antwoorden goed zijn én np361 goed is.
6.3 *Schrijf de breuken erin.*	np3631 = opschrijven breuken eerste stokbrood in strook D Antwoord = 1/9 staat in elk gekleurd stukje of indien niet gekleurd: in drie stukjes staat 1/9 OF 3/9 staat over de stukjes heen geschreven (samengenomen) np3632 = opschrijven breuken tweede stokbrood in strook D Antwoord = 1/9 staat in elk gekleurd stukje of indien niet gekleurd: in twee stukjes staat 1/9 OF 2/9 staat over de stukjes heen geschreven (samengenomen)	1 punt indien beide antwoorden goed
7. Wat is het antwoord op de som? *ANTWOORD: …*	np37 = totaalbreuk bij stukjes die in D gekleurd/geschreven hadden moeten worden Antwoord = 5/9 (indien anders = fout)	1*
8. Wat is waar als je breuken wilt optellen? *Kruis 1 antwoord aan.* ○ Als de noemers hetzelfde zijn kun je de noemers optellen. ○ Als de noemers hetzelfde zijn kun je de tellers optellen. ○ Als de tellers hetzelfde zijn kun je de noemers optellen.	np38 = MPC Antwoord = 2^{e}	1*

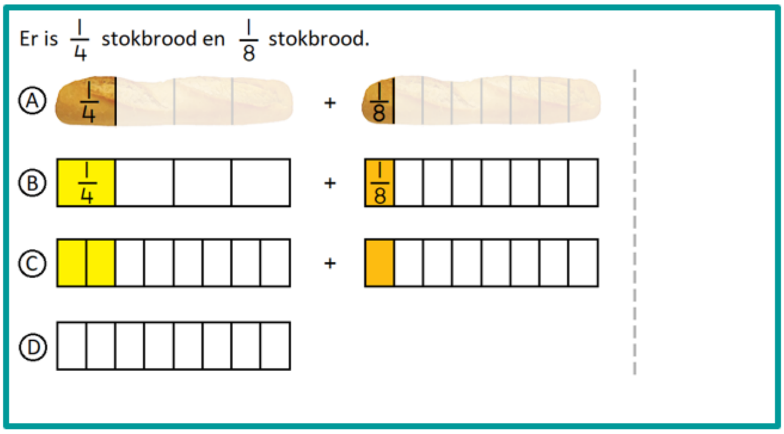
Booklet 3

Pre-test repetition

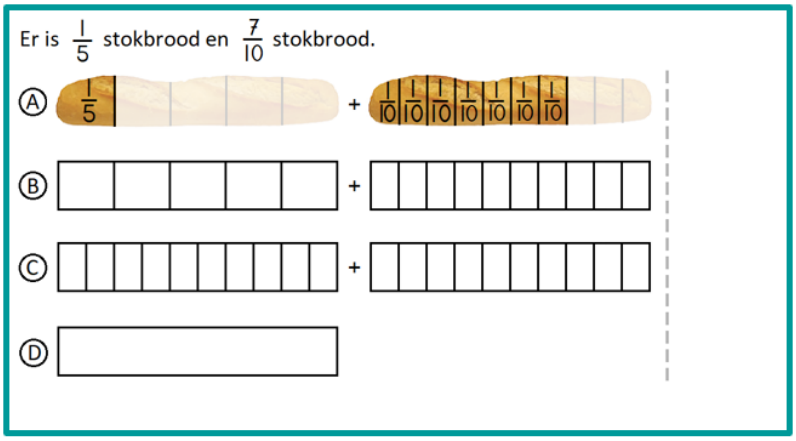
1. Verdeel de ontbijtkoek in kwarten. 	ba1 = 4 gelijke delen Antwoord = 	1*
2. Zet rondjes om wat evenveel is als $\frac{6}{9}$ *Er zijn er meer goed.* 	ba21 = 9/18 Antwoord = niet omcirkelen ba22 = 2/3 Antwoord = wel omcirkelen ba23 = 12/27 Antwoord = niet omcirkelen ba24 = 18/27 Antwoord = wel omcirkelen ba25 = 1/3 Antwoord = niet omcirkelen	1 punt indien 5 antwoorden goed. Anders 0 punten.
3. Schrijf naast het plaatje de breuk op van het deel dat gekleurd is. 	ba3 = blauw gedeelte van het geheel Antwoord = 3/10	1*
4. Zet een rondje om wat meer is. $\frac{3}{5}$ of $\frac{5}{10}$	ba4 = grootste breuk Antwoord = alleen 3/5 omcirkeld	1*

5. Trek lijnen van de kaartjes naar de goede plaats op de getallenlijnen. 	ba51 = kaartje 1/4 Antwoord = op lijn tussen 3/16 en 5/16 ba52 = kaartje 5/8 Antwoord = op lijn tussen 9/16 en 11/16 ba53 = kaartje 2/3 Antwoord = op lijn tussen 11/18 en 12/18 ba54 = kaartje 5/6 Antwoord = op lijn tussen 14/18 en 16/18	Elk onderdeel 0,25 punt. Alle onderdelen = samen max 1 punt.
6. Schrijf de breuk op. 5 stukjes van $\frac{1}{8}$ is …… 2 stukjes van $\frac{1}{3}$ is ……	ba61 = 5x 1/8 Antwoord =5/8 ba62 = 2 x 1/3 Antwoord = 2/3	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.
7. Zet in de twee breuken hieronder rondjes om de tellers en pijlen naar de noemers. $\frac{1}{4}$ $\frac{4}{8}$	ba71 = teller in 1/4 Antwoord = 1 omcirkelen ba72 = teller in 4/8 Antwoord = 4 rechts omcirkelen ba73 = noemer in 1/4 Antwoord = 4 links omcirkelen ba74 = noemer in 4/8 Antwoord = 8 omcirkelen	Elk onderdeel 0,25 punt. Alle onderdelen = samen max 1 punt.
8. Reken de sommen uit. $\frac{2}{5} + \frac{1}{5} =$ $\frac{6}{8} + \frac{1}{8} =$	ba81 = bovenste som Antwoord = 3/5 ba82 = onderste som Antwoord = 7/8	Elk onderdeel 0,5 punt. Alle onderdelen = samen max 1 punt.

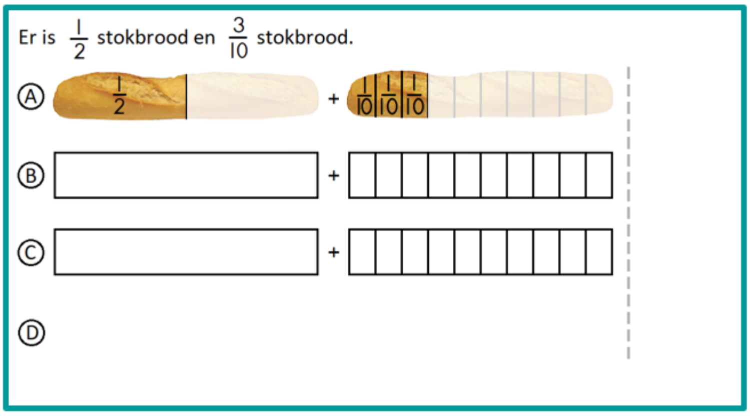
Immediate post-test

 1.1 *Schrijf bij stap B de som naast de stippellijn.*	ni111 = breukenstukje eerste strook bij B Antwoord = 1/4. Mag ook achteraan in som. ni112 = breukenstukje tweede strook bij B Antwoord = 1/8	1 punt indien beide antwoorden goed
1.2 *Schrijf bij stap C de breuken in de stukjes.*	ni121 = opschrijven breuken eerste stokbrood in strook C Antwoord = 1/8 staat in beide stukjes of 2/8 staat over de stukjes samengenomen ni122 = opschrijven breuken tweede stokbrood in strook C Antwoord = 1/8 staat in het oranje stukje geschreven	1 punt indien beide antwoorden goed
1.3 *Schrijf de nieuwe som naast de stippellijn.*	ni131 = breukenstukjes eerste strook bij C Antwoord = 2/8. Mag ook achteraan in som. ni132 = breukenstukje tweede strook bij C Antwoord = 1/8. Mag ook vooraan in som.	1 punt indien beide antwoorden goed
1.4 *Kleur bij stap D de stukjes stokbrood samen in de strook.*	ni141 = kleuren gele stokbrood in D Antwoord = 2 stukjes gekleurd, kleur maakt niet uit. ni142 = kleuren oranjes stokbrood in D Antwoord = 1 stukje gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed

1.5 *Schrijf de breuken in de stukjes.*	ni151 = opschrijven breuken eerste stokbrood in strook D Antwoord = 1/8 staat in beide stukjes, OF 2/8 over beide stukjes samengenomen, OF 3/8 over stukjes van ni151 en ni152 samengenomen ni152 = opschrijven breuken tweede stokbrood in D Antwoord = 1/8 staat in laatste stukje, OF 3/8 over stukjes van ni151 en ni152 samengenomen	1 punt indien beide antwoorden goed
1.6 *Schrijf de som* ***en het antwoord*** *naast de stippellijn.*	ni161 = breukenstukjes eerste stuk stokbrood Antwoord = 2/8. Mag ook achteraan in som. ni162 = breukenstukjes tweede stuk stokbrood Antwoord = 1/8. Mag ook vooraan in som. ni163 = antwoord op som 2/8+1/8 Antwoord = 3/8 (indien anders = fout) Mag ook achter de som bij ni131&ni132 staan.	1 punt indien ni163 goed

 2.1 *Teken bij stap B zelf de stokbroden na met de breuken erin.*	ni211 = kleuren linkerstokbrood in strook B Antwoord = 1 stukje kleuren ni212 = kleuren rechterstokbrood in strook B Antwoord = 7 stukjes kleuren ni213 = opschrijven breuk in linkerstokbrood in linkerstrook B Antwoord = 1/5 ni214 = opschrijven breuken in rechterstokbrood in rechterstrook B Antwoord = 7 keer 1/10 OF 7/10 over alle stukjes samengenomen	1 punt indien beide antwoorden goed 1 punt indien beide antwoorden goed
2.2 *Schrijf de som naast de stippellijn.*	ni221 = breukenstukje linkerstrook bij B Antwoord = 1/5. Mag ook achteraan. ni222 = breukenstukjes rechterstrook bij B Antwoord = 7/10. Mag ook vooraan	1 punt indien beide antwoorden goed
2.3 *Bedenk zelf wat je in de stroken van stap C moet doen en doe dat.*	ni231 = Kleuren linkerstukjes van A in strook C Antwoord = 2 gele stukjes, mag ook andere kleur ni232 = Kleuren rechterstukjes van A in strook Antwoord = 7 oranje stukjes, mag ook andere kleur ni233 = opschrijven breuken linkerstokbrood van strook C Antwoord = 1/10 staat in beide stukjes, OF 2/10 over beide stukjes samengenomen ni234 = opschrijven breuken rechterstokbrood van strook C Antwoord = 1/10 staat 7 stukjes, OF 7/10 over beide stukjes samengenomen	1 punt indien beide antwoorden goed 1 punt indien beide antwoorden goed
2.4 *Schrijf de nieuwe som naast de stippellijn.*	ni241 = breukenstukjes linkerstokbrood van C Antwoord =2/10. Mag ook achteraan in som. ni242 = breukenstukjes rechterstokbrood van C Antwoord =7/10. Mag ook vooraan in som.	1 punt indien beide antwoorden goed

2.6 *Kleur de stukjes stokbrood er samen in.*	ni261 = Kleuren linkerstokbrood in D Antwoord = 2 stukjes gekleurd, kleur maakt niet uit ni262 = kleuren rechterstokbrood in D Antwoord = 7 stukjes gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed
2.7 *Schrijf de breuken in de stukjes.*	ni271 = opschrijven breuken linkerstokbrood in strook D Antwoord = 1/10 staat in beide stukjes, OF 2/10 over beide stukjes samengenomen, OF 9/10 over stukjes van ni271 en ni272 ni272 = opschrijven breuken rechterstokbrood in strook D Antwoord = 1/10 staat in 7 stukjes, OF 7/10 over beide stukjes samengenomen, OF 9/10 over stukjes van ni271 en ni272	1 punt indien beide antwoorden goed
2.8 *Schrijf de som **en het antwoord** naast de stippellijn.*	ni281 = breukenstukjes linkerstokbrood Antwoord = 2/10. Mag ook achteraan in de som ni282 = breukenstukjes rechterstokbrood Antwoord = 7/10. Mag ook vooraan in de som ni283 = Antwoord op som 2/10+7/10 Antwoord = 9/10 (indien anders = fout). Mag ook achter de som bij ni241&ni242 staan.	1 punt indien ni283 goed

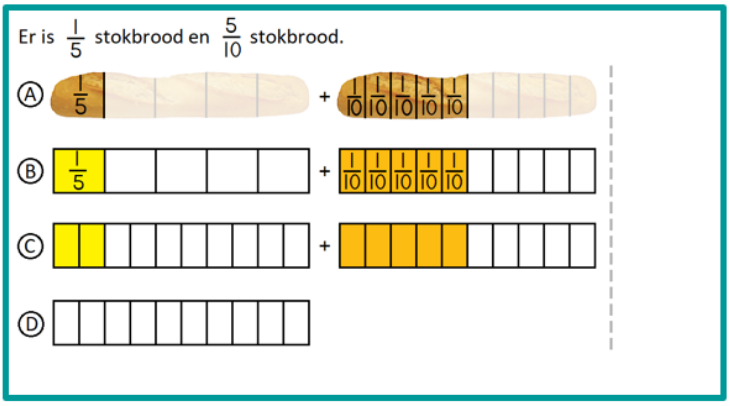
 	ni311 = verdelen linkerstokbrood bij B Antwoord = in twee ongeveer gelijke stukken verdeeld ni312 = kleuren linkerstokbrood bij B Antwoord = 1 stuk kleuren ni313 = kleuren rechterstokbrood bij B Antwoord = 3 stukjes kleuren	1 punt indien 3 antwoorden goed
3.1 *Teken bij stap B zelf de stokbroden na met de breuken erin.*	ni314 = opschrijven breuk in linkerstokbrood in linkerstrook B Antwoord = 1/2 ni315 = opschrijven breuken in rechterstokbrood in rechterstrook B Antwoord = 3/10	1 punt indien beide antwoorden goed
3.2 *Schrijf de som naast de stippellijn.*	ni321 = breukenstukje linkerstrook bij B Antwoord =1/2 ni322 = breukenstukjes rechterstrook bij B Antwoord = 3/10	1 punt indien beide antwoorden goed
3.3 *Verdeel bij stap C zelf de stokbroden in evenveel stukjes.*	ni33 = linkerstrook bij C in stukken verdelen Antwoord = 10 stukken	1*
3.4 *Maak de stroken af.*	ni341 = kleuren linkerstokbrood van A in strook C Antwoord = 5 stukjes gekleurd, kleur maakt niet uit ni342 = kleuren rechterstokbrood van A in strook C Antwoord = 3 stukjes gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed
	ni343 = opschrijven breuken linkerstokbrood van strook C Antwoord = 1/10 staat in 5 stukjes, OF 5/10 over 5 stukjes samengenomen ni344 = opschrijven breuken rechterstokbrood van strook C Antwoord = 1/10 staat in 3 stukjes, OF 3/10 over 3 stukjes samengenomen	1 punt indien beide antwoorden goed
3.5 *Schrijf de nieuwe som naast de stippellijn*	ni351 = breukenstukjes linkerstokbrood van C Antwoord = 5/10. Mag ook achteraan in som. ni352 = breukenstukjes rechterstokbrood van C Antwoord = 3/10. Mag ook vooraan in som	1 punt indien beide antwoorden goed

3.6 *Teken bij stap D zelf een strook in het goede aantal stukken.*	ni36 = Bij D zelf strook tekenen en in stukken verdelen Antwoord = getekende strook is in 10 stukken verdeeld	1*
3.7 *Maak de strook af.*	ni371 = Kleuren linkerstokbrood in D Antwoord = 5 stukjes gekleurd, kleur maakt niet uit ni372 = Kleuren rechterstokbrood in D Antwoord = 3 stukjes gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed
	ni373 = opschrijven breuken linkerstokbrood in D Antwoord = 1/10 staat in 5 stukjes, OF 5/10 over 5 stukjes samengenomen, OF 8/10 over stukjes van ni373 en ni374 ni374 = opschrijven breuken rechterstokbrood in D Antwoord = 1/10 staat in 3 stukjes, OF 3/10 over 3 stukjes samengenomen, OF 8/10 over stukjes van ni373 en ni374	1 punt indien beide antwoorden goed
3.8 *Schrijf de som* ***en het antwoord*** *naast de stippellijn.*	ni381 = breukenstukjes linkerstokbrood Antwoord = 5/10 ni382 = breukenstukjes rechterstokbrood Antwoord = 3/10 ni383 = antwoord op som 5/10+3/10 Antwoord = 8/10 OF 4/5 (indien anders = fout). Mag ook achter de som bij ni3511&ni352 staan.	1 punt indien ni383 goed

4. Er is $\frac{2}{6}$ stokbrood en $\frac{5}{12}$ stokbrood. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan.*	ni4 = antwoord op som 2/6 + 5/12 Antwoord = 9/12 OF 3/4 (indien anders = fout)	1*
5. Reken uit. $\frac{2}{4} + \frac{1}{8} =$ Je mag het tekenen. *Schijf het antwoord onderaan.*	ni5 = antwoord op som 2/4 + 1/8 Antwoord = 5/8 (indien anders = fout)	1*
6. $\frac{1}{4} + \frac{1}{12} =$ Je mag het tekenen. *Schijf het antwoord onderaan.*	ni6 = antwoord op som 1/4 + 1/12 Antwoord = 4/12 OF 1/3 (indien anders = fout)	1*

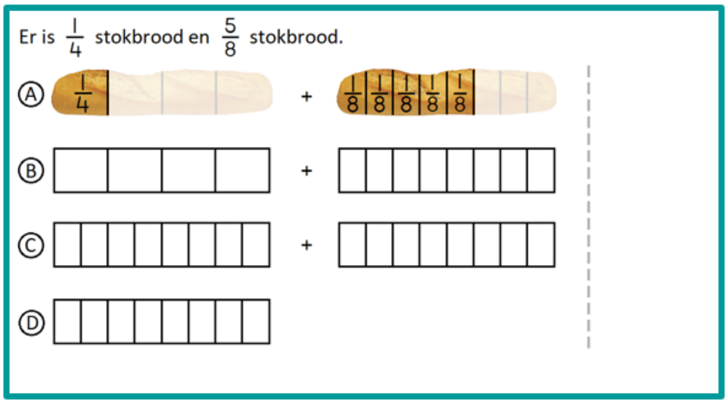
Booklet 4

Delayed post-test

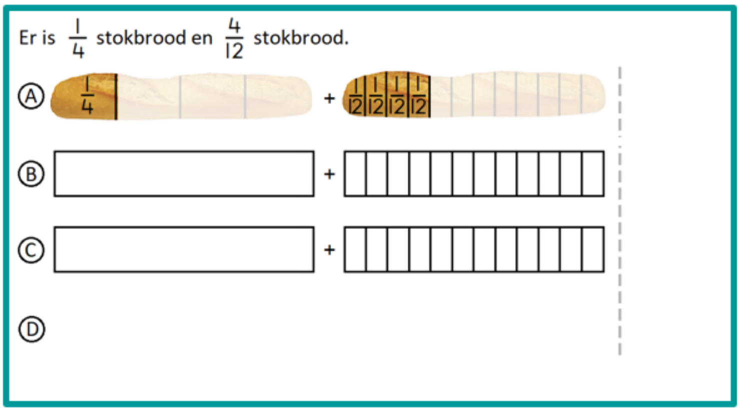


1.1 *Schrijf bij stap B de som naast de stippellijn.*	nd111 = breukenstukje eerste strook bij B Antwoord = 1/5. Mag ook achteraan in som. nd112 = breukenstukjes tweede strook bij B Antwoord = 5/10	1 punt indien beide antwoorden goed
1.2 *Schrijf bij stap C de breuken in de stukjes.*	nd121 = opschrijven breuken eerste stokbrood in strook C Antwoord = 1/10 staat in beide stukjes of 2/10 staat over de stukjes samengenomen nd122 = opschrijven breuken tweede stokbrood in strook C Antwoord = 1/10 staat in de oranje stukjes of 5/10 over de stukjes samengenomen	1 punt indien beide antwoorden goed
1.3 *Schrijf de nieuwe som naast de stippellijn.*	nd131 = breukenstukjes eerste strook bij C Antwoord = 2/10. Mag ook achteraan in som. nd132 = breukenstukje tweede strook bij C Antwoord = 5/10. Mag ook vooraan in som.	1 punt indien beide antwoorden goed
1.4 *Kleur bij stap D de stukjes stokbrood samen in de strook.*	nd141 = kleuren gele stokbrood in D Antwoord = 2 stukjes gekleurd, kleur maakt niet uit. nd142 = kleuren oranjes stokbrood in D Antwoord = 5 stukjes gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed

1.5 *Schrijf de breuken in de stukjes.*	nd151 = opschrijven breuken eerste stokbrood in strook D Antwoord = 1/10 staat in beide stukjes, OF 2/10 over beide stukjes samengenomen, OF 7/10 over stukjes van nd151 en nd152 samengenomen nd152 = opschrijven breuken tweede stokbrood in D Antwoord = 1/10 staat in 5 stukjes, OF 5/10 over 5 stukjes, OF 7/10 over stukjes van nd151 en nd152 samengenomen	1 punt indien beide antwoorden goed
1.6 *Schrijf de som* ***en het antwoord*** *naast de stippellijn.*	nd161 = breukenstukjes eerste stuk stokbrood Antwoord = 2/10. Mag ook achteraan in som. nd162 = breukenstukjes tweede stuk stokbrood Antwoord = 7/10. Mag ook vooraan in som. nd163 = antwoord op som 2/10+5/10 Antwoord = 7/10 (indien anders = fout) Mag ook achter de som bij nd131&nd132 staan.	1 punt indien nd163 goed

 2.1 *Teken bij stap B de stokbroden na als stroken.*	nd211 = kleuren linkerstokbrood in strook B Antwoord = 1 stukje kleuren nd212 = kleuren rechterstokbrood in strook B Antwoord = 5 stukjes kleuren nd213 = opschrijven breuk in linkerstokbrood in linkerstrook B Antwoord = 1/4 nd214 = opschrijven breuken in rechterstokbrood in rechterstrook B Antwoord = 5 keer 1/8 OF 5/8 over alle stukjes samengenomen	1 punt indien beide antwoorden goed 1 punt indien beide antwoorden goed
2.2 *Schrijf de som naast de stippellijn.*	nd221 = breukenstukje linkerstrook bij B Antwoord = 1/4. Mag ook achteraan. nd222 = breukenstukjes rechterstrook bij B Antwoord = 5/8. Mag ook vooraan	1 punt indien beide antwoorden goed
2.3 *Maak bij stap C de stroken af.*	nd231 = Kleuren linkerstukjes van A in strook C Antwoord = 2 gele stukjes, mag ook andere kleur nd232 = Kleuren rechterstukjes van A in strook Antwoord = 5 oranje stukjes, mag ook andere kleur nd233 = opschrijven breuken linkerstokbrood van strook C Antwoord = 1/8 staat in beide stukjes, OF 2/8 over beide stukjes samengenomen nd234 = opschrijven breuken rechterstokbrood van strook C Antwoord = 1/8 staat 5 stukjes, OF 5/8 over beide stukjes samengenomen	1 punt indien beide antwoorden goed 1 punt indien beide antwoorden goed
2.4 *Schrijf de nieuwe som naast de stippellijn.*	nd241 = breukenstukjes linkerstokbrood van C Antwoord = 2/8. Mag ook achteraan in som. nd242 = breukenstukjes rechterstokbrood van C Antwoord =5/8. Mag ook vooraan in som.	1 punt indien beide antwoorden goed

2.6 *Kleur de stukjes stokbrood er samen in.*	nd261 = Kleuren linkerstokbrood in D Antwoord = 2 stukjes gekleurd, kleur maakt niet uit nd262 = kleuren rechterstokbrood in D Antwoord = 5 stukjes gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed
2.7 *Schrijf de breuken in de stukjes.*	nd271 = opschrijven breuken linkerstokbrood in strook D Antwoord = 1/8 staat in beide stukjes, OF 2/8 over beide stukjes samengenomen, OF 7/8 over stukjes van nd271 en nd272 nd272 = opschrijven breuken rechterstokbrood in strook D Antwoord = 1/8 staat in 5 stukjes, OF 5/8 over beide stukjes samengenomen, OF 7/8 over stukjes van nd271 en nd272	1 punt indien beide antwoorden goed
2.8 *Schrijf de som* ***en het antwoord*** *naast de stippellijn.*	nd281 = breukenstukjes linkerstokbrood Antwoord = 2/8. Mag ook achteraan in de som nd282 = breukenstukjes rechterstokbrood Antwoord = 5/8. Mag ook vooraan in de som nd283 = Antwoord op som 2/8+5/8 Antwoord = 7/8 (indien anders = fout). Mag ook achter de som bij nd241&nd242 staan.	1 punt indien nd283 goed

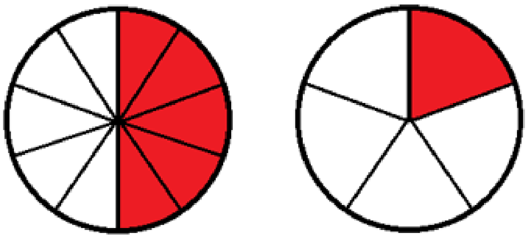


3.1 *Teken bij stap B zelf de stokbroden na met de breuken erin.*	nd311 = verdelen linkerstokbrood bij B Antwoord = in vier ongeveer gelijke stukken verdeeld nd312 = kleuren linkerstokbrood bij B Antwoord = 1 stuk kleuren nd313 = kleuren rechterstokbrood bij B Antwoord = 4 stukjes kleuren nd314 = opschrijven breuk in linkerstokbrood in linkerstrook B Antwoord = 1/4 nd315 = opschrijven breuken in rechterstokbrood in rechterstrook B Antwoord = 4/12	1 punt indien 3 antwoorden goed 1 punt indien beide antwoorden goed
3.2 *Schrijf de som naast de stippellijn.*	nd321 = breukenstukje linkerstrook bij B Antwoord =1/4 nd322 = breukenstukjes rechterstrook bij B Antwoord = 4/12	1 punt indien beide antwoorden goed
3.3 *Verdeel bij stap C zelf de stokbroden in evenveel stukjes.*	nd33 = linkerstrook bij C in stukken verdelen Antwoord = 12 stukken	1*
3.4 *Maak de stroken af.*	nd341 = kleuren linkerstokbrood van A in strook C Antwoord = 2 stukjes gekleurd, kleur maakt niet uit nd342 = kleuren rechterstokbrood van A in strook C Antwoord = 4 stukjes gekleurd, kleur maakt niet uit nd343 = opschrijven breuken linkerstokbrood van strook C Antwoord = 1/12 staat in 3 stukjes, OF 3/12 over 3 stukjes samengenomen nd344 = opschrijven breuken rechterstokbrood van strook C Antwoord = 1/12 staat in 4 stukjes, OF 4/12 over 4 stukjes samengenomen	1 punt indien beide antwoorden goed 1 punt indien beide antwoorden goed
3.5 *Schrijf de nieuwe som naast de stippellijn*	nd351 = breukenstukjes linkerstokbrood van C Antwoord = 3/12. Mag ook achteraan in som. nd352 = breukenstukjes rechterstokbrood van C Antwoord = 4/12. Mag ook vooraan in som	1 punt indien beide antwoorden goed

3.6 *Teken bij stap D zelf een strook in het goede aantal stukken.*	nd36 = Bij D zelf strook tekenen en in stukken verdelen Antwoord = getekende strook is in 12 stukken verdeeld	1*
3.7 *Maak de strook af.*	nd371 = Kleuren linkerstokbrood in D Antwoord = 3 stukjes gekleurd, kleur maakt niet uit nd372 = Kleuren rechterstokbrood in D Antwoord = 4 stukjes gekleurd, kleur maakt niet uit	1 punt indien beide antwoorden goed
	nd373 = opschrijven breuken linkerstokbrood in D Antwoord = 1/12 staat in 3 stukjes, OF 3/12 over 3 stukjes samengenomen, OF 7/12 over stukjes van nd373 en nd374 nd374 = opschrijven breuken rechterstokbrood in D Antwoord = 1/12 staat in 4 stukjes, OF 4/12 over 4 stukjes samengenomen, OF 7/12 over stukjes van nd373 en nd374	1 punt indien beide antwoorden goed
3.8 *Schrijf de som* ***en het antwoord*** *naast de stippellijn.*	nd381 = breukenstukjes linkerstokbrood Antwoord = 3/12 nd382 = breukenstukjes rechterstokbrood Antwoord = 4/12 nd383 = antwoord op som 3/12+4/12 Antwoord = 7/12 (indien anders = fout). Mag ook achter de som bij nd351&nd352 staan.	1 punt indien nd383 goed

4. Er is $\frac{4}{6}$ stokbrood en $\frac{3}{12}$ stokbrood. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan.*	nd4 = antwoord op som 4/6 + 3/12 Antwoord = 11/12 (indien anders = fout)	1*
5. Reken uit. $\frac{2}{5} + \frac{1}{10} =$ Je mag het tekenen. *Schijf het antwoord onderaan.*	nd5 = antwoord op som 2/5 + 1/10 Antwoord = 5/10 OF 1/2 (indien anders = fout)	1*
6. $\frac{1}{3} + \frac{1}{9} =$ Je mag het tekenen. *Schijf het antwoord onderaan.*	nd6 = antwoord op som 1/3 + 1/9 Antwoord = 4/9 (indien anders = fout)	1*

Transfer test

1. Hoeveel is dit samen? Je mag het tekenen. *Schrijf je antwoord onderaan deze som.* 	t1 = rode stukken samengenomen Antwoord = 7/10 OF in 1 cirkel getekend die in 10 stukken is verdeeld waarvan er 7 zijn gekleurd. (indien anders = fout)	1*
2. Reken uit. Je mag het tekenen. *Schrijf je antwoord onderaan.* $\frac{11}{12} - \frac{2}{3} =$	t2 = antwoord op som 11/12 – 2/3 Antwoord = 3/12 OF 1/4 (indien anders = fout)	1*
3. Er is $\frac{1}{2}$ ontbijtkoek en $\frac{1}{3}$ ontbijtkoek. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan deze som.*	t3 = antwoord op som 1/2 + 1/3 Antwoord = 5/6 OF 10/12 (indien anders = fout)	1*
4. Er is $\frac{1}{4}$ ontbijtkoek en $\frac{3}{5}$ ontbijtkoek. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan.*	t4 = antwoord op som 1/4 + 3/5 Antwoord = 17/20 (indien anders = fout)	1*
5. Er is $\frac{2}{3}$ ontbijtkoek en $\frac{7}{9}$ ontbijtkoek. Hoeveel is dit samen? Je mag het tekenen. *Schijf het antwoord onderaan.*	t5 = antwoord op som 2/3 + 7/9 Antwoord = 1 4/9 OF 13/9 (indien anders = fout)	1*
6. Job heeft $\frac{3}{4}$ van zijn ontbijtkoek opgegeten en geeft de rest aan Lisa. Mila heeft $\frac{7}{12}$ van haar ontbijtkoek opgegeten en geeft de rest ook aan Lisa. Hoeveel ontbijtkoek heeft Lisa nu? Je mag het tekenen. *Schijf het antwoord onderaan.*	t6 = antwoord op som 1/4 + 5/12 Antwoord = 8/12 OF 4/6 OF 2/3 OF 16/24 (indien anders = fout)	1*
7. In de fles cola zit $\frac{11}{15}$ liter, Sep drinkt $\frac{3}{5}$ liter op. Hoeveel is er nog over? Je mag het tekenen. *Schijf het antwoord onderaan.* 	t7 = antwoord op som 11/15 – 3/5 Antwoord = 2/15 (indien anders = fout)	1*

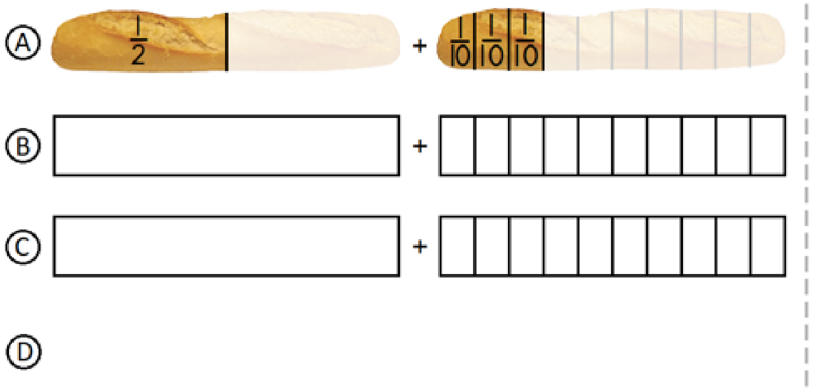
Note. 1* Betekent 1 punt indien antwoord goed.

Appendix J

Immediate Post-test Problem 3

3. The box below is incomplete. You are going to complete it.
   Now start with task 3.1 beneath the box.

There is $\frac{1}{2}$ baguette, and $\frac{3}{10}$ baguette. How much is this together?



3.1 *At step B, you copy the baguettes and fill them with fractions.*
3.2 *Write down the operation next to the dotted line.*

3.3 *At step C, you divide the baguettes in an equal number of pieces.*
3.4 *Complete the bars.*
3.5 *Write down the new operation next to the dotted line.*

3.6 *At step D, you draw a bar in the right number of pieces.*
3.7 *Complete the bar.*
3.8 *Write down the operation* ***and the solution*** *next to the dotted line.*

GO ON TO THE NEXT PAGE OF THIS BOOKLET

Appendix K

Mean Percentages of Correct Scores and Standard Deviations on Play

	Video 1.1		Video 1.2		Video 2.1		Video 2.2		Video 3.1		Video 3.2		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct-Correct (*n* = 40)	104.4	19.6	87.3	32.0	90.4	33.9	68.3	37.8	73.5	34.8	58.6	39.8	80.4	20.3
Correct-Incorrect (*n* = 41)	111.1	46.6	92.7	31.1	94.3	21.7	89.3	28.2	79.8	34.0	71.8	41.5	89.8	18.7
Total (*N* = 81)	107.8	35.8	90.0	31.5	92.4	28.3	78.9	34.7	76.7	34.3	65.3	40.9	85.2	20.0

Appendix L

Absolute Unique Playtime Scores and Standard Deviations in Seconds

	Video 1.1		Video 1.2[a]		Video 2.1		Video 2.2[a]		Video 3.1		Video 3.2[a]		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct-Correct (*n* = 40)	218	35	184	68	180	58	140	79	121	60	86	60	931	245
Correct-Incorrect (*n* = 41)	209	50	125	63	199	43	130	41	133	56	128	74	923	186
Total (*N* = 81)	213	43	154	63	190	52	135	63	127	58	107	70	927	216

[a]Videos 1.2, 2.2, and 3.2 have different lengths in each condition. For an overview of all video lengths, see Appendix B.

Appendix M

Summary of Spearman Rank Correlations for all Possible Pairs by Condition

Measure	1 SEb	2 SR	3 Pre	4 Play	5 Uplay	6 A. Pl	7 A. Upl	8 Prac	9 SEa	10 Imm	11 Delay	12 Trans
1. Self-efficacy before												
C-C condition	—	.56**	.49**	-.37*	-.48**	-.33*	-.45**	.10	.83**	.22	.37*	.33*
C-I condition	—	.46**	.32*	.09	.15	.09	.14	.20	.84**	.33*	.42**	.51**
2. Self-regulation												
C-C condition		—	.39*	-.25	-.35*	-.21	-.31*	.27	.58**	.44**	.58**	.65**
C-I condition		—	.26	-.29	-.13	-.27	-.13	.20	.58**	.11	.24	.26
3. Pre-test												
C-C condition			—	-.15	-.13	-.16	-.14	.55**	.33*	.45**	.45**	.45**
C-I condition			—	-.04	.09	-.05	.08	.41**	.49**	.49**	.39*	.54**
4. Play												
C-C condition				—	.95**	.98**	.96**	-.02	-.27	-.07	-.18	-.03
C-I condition				—	.85**	1.00**	.85**	.05	.02	.21	.27	-.08
5. Unique play												
C-C condition					—	.91**	.99**	.08	-.36*	-.06	-.19	-.06
C-I condition					—	.83*	1.00**	.21	.13	.31*	.34*	.04
6. Absolute playtime												
C-C condition						—	.94**	-.06	-.24	-.09	-.19	.01
C-I condition						—	.85**	.05	.02	.21	.27	-.08
7. Absolute unique playtime												
C-C condition							—	.07	-.33*	-.05	-.19	-.05
C-I condition							—	.20	.13	.31	.34*	.03
8. Practice												
C-C condition								—	.08	.57**	.61**	.49**
C-I condition								—	.44**	.71**	.74**	.60**
9. Self-efficacy after												
C-C condition									—	.28	.35*	.48**
C-I condition									—	.42**	.58**	.50**
10. Immediate post-test												
C-C condition										—	.74**	.54**
C-I condition										—	.73**	.64**
11 Delayed post-test												
C-C condition											—	.57**
C-I condition											—	.64**
12 Transfer post-test												
C-C condition												—
C-I condition												—

Note. C-C = Correct-Correct; C-I = Correct-Incorrect. In the C-C condition n = 40, in the C-I condition n = 41. In the column heads, abbreviations are used. They correspond to the terms in the leftmost column.
*p < .05. **p < .01.