Pizzas or no pizzas: an advantage of word problems in fraction arithmetic?

Mostert, T.M.M.; Hickendorff, M.

Learning and Instruction 86 (2023) 101775

Contents lists available at ScienceDirect

Learning and Instruction

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

Pizzas or no pizzas: An advantage of word problems in fraction arithmetic?

T.M.M. Mostert, M. Hickendorff *

Institute of Education and Child Studies, Leiden University, PO Box 9555, 2300 RB, Leiden, the Netherlands

ARTICLE INFO

Keywords:
Word problems
Arithmetic
Mathematics education
Fractions
Primary school
Secondary school

ABSTRACT

Fractions are an important but notoriously difficult domain in mathematics education. Situating fraction arithmetic problems in a realistic setting might help students overcome their difficulties by making fraction arithmetic less abstract. The current study therefore investigated to what extent students (106 sixth graders, 187 seventh graders, and 192 eighth graders) perform better on fraction arithmetic problems presented as word problems compared to these problems presented symbolically. Results showed that in multiplication of a fraction with a whole number and in all types of fraction division, word problems were easier than their symbolic counterparts. However, in addition, subtraction, and multiplication of two fractions, symbolic problems were easier. There were no performance differences by students' grade, but higher conceptual fraction knowledge was associated with higher fraction arithmetic performance. Taken together this study showed that situating fraction arithmetic in a realistic setting may support or hinder performance, dependent on the problem demands.

1. Introduction

Fractions play an important role in mathematics education because they require a deeper understanding of numbers and they play an important role in advanced mathematics and in the use of mathematics in other fields such as economics, physics, biology, and chemistry (Bruin-Muurling, 2010; Siegler et al., 2013). However, students experience profound difficulties in acquiring proficiency in this domain (Hiebert, 1985; Lortie-Forgues et al., 2015; Siegler et al., 2013; Siegler & Lortie-Forgues, 2017). One of the reasons for these difficulties is that fractions are abstract and formal, and as such difficult to understand (Lortie-Forgues et al., 2015). A possible way to overcome these difficulties is to make the fractions more concrete and easier to understand by situating fraction arithmetic in a realistic setting, as is done in word problems (Hiebert, 1985). The current study aimed to provide empirical evidence for this mechanism, by investigating the extent to which there is a performance advantage of word problems (e.g., "Thomas has 3 liters lemonade. How many bottles of $\frac{3}{5}$ liters can he fill?") compared to their symbolic counterparts ($3 : \frac{3}{5}$) in fraction arithmetic.

1.1. Fractions

It has long been known that fractions are difficult for students in primary and secondary school (Hiebert, 1985; Lortie-Forgues et al., 2015; Siegler et al., 2013; Siegler & Lortie-Forgues, 2017). Because fraction understanding in primary school is considered an important predictor of later success in high school mathematics (Siegler et al., 2012) and is related to general mathematical abilities in different countries (Torbeyns et al., 2015), it is no surprise that a large body of literature is focused on the understanding of fractions. However, fewer studies focus on performing arithmetic operations (addition, subtraction, multiplication, and division) on fractions (Lortie-Forgues et al., 2015) whereas that is crucial in everyday life, for instance when following a recipe, as well as for several occupations, such as dosage calculation in medical occupations and for statistical computations and probability calculations.

The difference described above between fraction understanding and performing arithmetic operations reflects the distinction between conceptual fraction knowledge and procedural fraction knowledge. Procedural fraction knowledge can be interpreted as the ability to solve fraction arithmetic problems and to know how fraction arithmetic procedures work, whereas conceptual fraction knowledge includes the implicit or explicit understanding of the fraction domain (Lenz et al., 2019; Rittle-Johnson et al., 2001). Relations have been found between procedural and conceptual fraction knowledge. However, there has been no consensus on whether procedural knowledge is needed to acquire conceptual fraction knowledge or the other way around (Hallett et al., 2010; Rittle-Johnson et al., 2001). Although conceptual and procedural fraction knowledge are strongly correlated, research showed that these two knowledge types are empirically separable (Lenz et al., 2019).

* Corresponding author. Institute of Education and Child Studies, Leiden University, PO Box 9555, 2300 RB, Leiden, the Netherlands.
E-mail addresses: t.m.m.mostert@fsw.leidenuniv.nl (T.M.M. Mostert), hickendorff@fsw.leidenuniv.nl (M. Hickendorff).

https://doi.org/10.1016/j.learninstruc.2023.101775
Received 5 April 2022; Received in revised form 13 January 2023; Accepted 26 March 2023
Available online 6 April 2023

Therefore, it is important to study these knowledge types as separate constructs. While ample studies focus on conceptual fraction knowledge, this study will focus on procedural fraction knowledge.

1.2. Difficulties in fraction arithmetic

Students are faced with several difficulties when doing fraction arithmetic. According to Hiebert (1985), the complexity of fraction arithmetic is caused by difficulties in connecting fraction *form*, the numbers, symbols and operations, and fraction *understanding*, the ideas and intuitions about the reality of mathematics. One of those difficulties is that the three-parted fraction structure $\frac{a}{b}$, including the numerator a, denominator b and the fraction bar separating those numbers, can be quite confusing leading students to misread fractions (Lortie-Forgues et al., 2015). For example, research showed that many Dutch ninth graders did not see fractions as a division (Bruin-Muurling, 2010, Chapter 4).

A common difficulty is that students interpret the numerator and denominator as two separate numbers instead of the ratio of these numbers (Gabriel et al., 2013; Lortie-Forgues et al., 2015), the so-called whole-number bias (Ni & Zhou, 2005; Siegler et al., 2011). Consequently, children have difficulties understanding that many properties of whole-number arithmetic are not valid for fraction arithmetic. This leads to overgeneralization of procedures for whole numbers to fractions, such as when procedures for whole numbers are applied to numerators and denominators separately, for example when $\frac{3}{4}+\frac{2}{5}=$ is incorrectly solved as $\frac{3+2}{4+5}=\frac{5}{9}$ (e.g., Braithwaite et al., 2017). Furthermore, students are biased by the direction of effects in whole-number arithmetic, and this bias seems rather persistent and independent of fraction understanding (Gabriel et al., 2013; Siegler & Lortie-Forgues, 2015). For example, US middle-school students and Canadian preservice teachers who understood fraction magnitudes and were able to correctly execute fraction arithmetic procedures in addition and subtraction problems, erroneously expected the whole-number principles – that the outcome of a multiplication is larger than each of the operands while the division outcome is smaller than each of the operands – to hold for multiplication and division problems with fractions smaller than 1 (Siegler & Lortie-Forgues, 2015).

Another inherent difficulty is that fraction arithmetic requires different procedures depending on the operations, the equality of the denominators, and the type of fraction (Braithwaite et al., 2017; Lortie-Forgues et al., 2015; Siegler & Pyke, 2013). In adding or subtracting two fractions the denominators must be equal, whereas this is not necessary in multiplying fractions. The procedure of dividing two fractions might be even more difficult and abstract because the divisor needs to be inverted and then multiplied with the dividend (Braithwaite et al., 2017; Siegler & Pyke, 2013). Moreover, the underlying mechanisms for these fraction arithmetic procedures are complex to understand, particularly for fraction division, and many students do not have the necessary knowledge prior to learning fraction arithmetic skills (Lortie-Forgues et al., 2015). As a result, students seem to know how to apply different fraction arithmetic procedures, but they do not understand why these procedures work (Brown & Quinn, 2006). Hence, fraction arithmetic procedures remain complex and abstract for many students.

In short, students have difficulties understanding fractions due to the abstract structure and the lack of meaning, since students do not see them as divisions. Furthermore, the whole-number bias, the direction-of-effects bias, and the diversity of arithmetic procedures for different operations make solving fraction arithmetic problems more difficult and abstract for students. Since competence with fractions is important for future mathematical competence and everyday life (Siegler et al., 2012; Torbeyns et al., 2015), it is important to investigate ways to overcome these difficulties.

1.3. Word problems

One of the possible ways to overcome students' difficulties in fraction arithmetic could be to situate the fraction arithmetic in a realistic setting or story, since that could enhance understanding (Hiebert, 1985). Word problems, in which a story sketching a realistic situation that requires mathematical modeling is presented, are an important part of mathematics education throughout the world and play a central role in contemporary mathematics education in the Netherlands (Hickendorff, 2021; van den Heuvel-Panhuizen & Drijvers, 2014). Word problems are thought to have several potential benefits for students: the realistic situations can be motivational, they can be seen as a starting point in instruction as a tool for mathematical modeling, i.e., the process of going from concrete real-life situations to abstract symbolic mathematical problems, and they create opportunities to practice the application of mathematical knowledge and skills in real-life situations (Verschaffel et al., 2020). Furthermore, the situations described in word problems might also elicit more informal, and possibly more effective, solution strategies, particularly when the mathematics is abstract for the students (Koedinger & Nathan, 2004).

As such, using word problems could contribute to overcoming the inherent difficulties students experience in fraction arithmetic. The abstract three-parted fraction structure and the concept of fractions as divisions are probably less abstract and more transparent when the fraction is referring to a meaningful situation. For example, the fraction $\frac{3}{25}$ is more meaningful and less abstract when it represents three students from a class of 25 students. Additionally, other difficulties, such as the whole-number bias or the direction-of-effects bias, might be partly overcome when a fraction arithmetic task is presented in a realistic situation. For example, it is more obvious that the outcome of 5 times $\frac{1}{4}$ is smaller than 5 when the fraction represents a quarter of a cake, compared to the symbolic task $5 \times \frac{1}{4}$ that is not linked to concrete objects. Furthermore, the realistic situations might promote the use of intuitive fraction arithmetic strategies, because of the "*naturalness of corresponding solution methods*" (Bruin-Muurling, 2010, p. 94). For instance, when the task is to calculate the total content of 20 milk cartons holding $\frac{3}{4}$ liters each, a natural solution strategy fitting the situation is repeated addition of $\frac{3}{4}$. By contrast, when the task is to calculate $\frac{3}{4}$ of 20 kg, the natural solution strategy would be to divide 20 kg into four parts of 5 kg and taking three of those parts. (Bruin-Muurling, 2010, p. 94). As mentioned earlier, fraction division has a particularly abstract formal solution procedure. Solving fraction division problems such as $\frac{1}{4}:\frac{1}{12}$ might be less abstract and therefore easier when the fractions refer to pizzas or pies, because this might elicit the mental model: how many pieces of $\frac{1}{12}$ fit into $\frac{1}{4}$ of a pizza (Fig. 1).

Previous studies in Dutch third to sixth graders showed no performance difference between word problems and their symbolic counterparts in whole-number arithmetic (Hickendorff, 2013, 2021). However, studies in the domain of algebra showed that US students were more successful in solving word problems than in solving mathematically equivalent symbolically presented problems because they more often used intuitive, informal strategies to solve the word problems (Koedinger et al., 2008; Koedinger & Nathan, 2004). Koedinger and Nathan (2004) suggested that the situation presented in word problems might help to overcome the abstractness of symbolically presented problems. Since fraction arithmetic in symbolic format is also abstract to students, word problems may help overcome this abstractness in a similar way. Furthermore, they also found that students were better able to detect and avoid errors when solving word problems. For example, students achieved better on decimal word problems than on their symbolic counterparts because it prevented them from incorrectly adding or subtracting dollars to cents. Misconceptions in fraction arithmetic, such as arithmetic errors because of the whole-number bias, might also be less frequent when presenting a fraction arithmetic problem in a realistic situation.

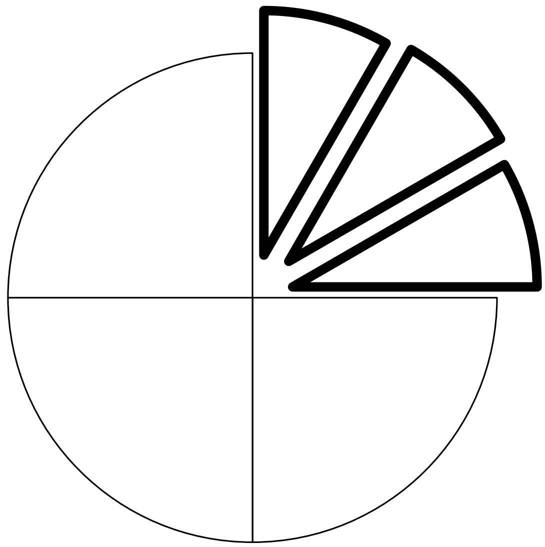

Fig. 1. Pie chart as mental model in solving fraction division problems such as $\frac{1}{4} : \frac{1}{12}$.

In conclusion, there are several possible ways in which situating fraction arithmetic in realistic settings might aid students' problem solving. The main hypothesis of the current study was therefore that in fraction arithmetic, students perform better on word problems than on their symbolic counterparts, whereas we did not expect to find this word-problem advantage for whole-number arithmetic, since that is not abstract for students at the end of primary school anymore. In what follows, we discuss two factors that could play a role in this hypothesized mechanism: conceptual fraction knowledge and students' instructional experiences with fractions.

1.4. Conceptual fraction knowledge

Better understanding of the concept of fractions and of fraction magnitude may help to perform better on fraction arithmetic (Hallett et al., 2010; Lortie-Forgues et al., 2015). Rittle-Johnson et al. (2001) found conceptual understanding, procedural knowledge, and problem representation of decimal fraction problems to be interrelated in US fifth and sixth graders. High conceptual knowledge may help students to construct a better mental representation of a problem and thereby increasing the chance of selecting an effective procedure. Furthermore, fraction magnitude knowledge allows students to reject procedures that generate implausible answers (Siegler & Pyke, 2013). One of our proposed mechanisms for a word-problem advantage in fraction arithmetic was that the situation of a fraction word problem might elicit mental representations that may aid in selecting and conducting effective procedures. Students with high conceptual knowledge might not need this situation to construct this mental model, but use their conceptual knowledge instead. Therefore, we hypothesize that the expected performance advantage of word problems is largest for students with low conceptual knowledge.

1.5. Instructional experiences

In the Dutch mathematics curriculum, fraction arithmetic is taught in sixth, seventh and eighth grade. During these years, students make a transition from primary school (sixth grade) to secondary school (seventh grade and further) (Meelissen et al., 2020). In Dutch secondary education, students are placed in differentiated tracks: prevocational education ("VMBO", further differentiated in three ordered tracks "VMBO-basis", "VMBO-kader", and "VMBO-theoretisch"), senior general secondary education ("HAVO") and pre-university education ("VWO").

Bruin-Muurling (2010) showed that there is a discontinuity in the fractions learning trajectory in the Dutch mathematics curriculum: in primary education, the focus is on informal, situational, solution strategies, which shifts into a focus on formal mathematical reasoning with numbers as mathematical objects in secondary education. Consequently, in sixth grade students solve fraction arithmetic problems primarily as word problems, whereas in seventh and eighth grade students primarily solve symbolic fraction arithmetic problems. The strong connection between situations and procedures in primary school is thus abolished in secondary education and possibly complicate the generalization to more formal, symbolic, fraction procedures. We expect to see these differences in instructional experiences reflected in the word problem advantage. That is, we hypothesize that the advantage of word problems is higher in sixth graders, for whom situated fraction arithmetic is common but symbolic fraction arithmetic is new and abstract, than in seventh and eighth graders, who have been taught procedures to solve symbolic fraction arithmetic.

1.6. The current study

Fractions are a critical concept in mathematics education, but they are also notoriously difficult for students. Situating fractions problems in a realistic story might help students overcoming these difficulties, because it is expected to make fractions easier to understand and fraction arithmetic easier to perform. The aim of the current study was therefore to provide empirical evidence for this suggested mechanism, by investigating the effect of problem representation (symbolic vs. word problems) in fraction arithmetic performance. A sample of sixth, seventh, and eighth graders solved addition, subtraction, multiplication, and division fraction arithmetic problems in word-problem format and in symbolic format. To ensure that the expected word-problem advantage is specific for fraction arithmetic, we also included whole-number arithmetic problems in symbolic and word-problem format as a control baseline, since previous studies showed no performance difference between word problems and their symbolic counterparts in whole-number arithmetic (Hickendorff, 2013, 2021). Furthermore, to address the potential moderating effect of conceptual knowledge of fractions, students also solved conceptual fraction knowledge problems. Finally, to address the potential impact of the discontinuity in the Dutch fractions learning trajectory from primary to secondary education (Bruin-Muurling, 2010), we included students from the final year of primary school (sixth graders) and the first two years of secondary school (seventh and eighth graders). We had the following expectations.

Hypothesis 1. We expected a performance advantage of word problems compared to their symbolic counterparts in fraction arithmetic, which we did not expect in whole number arithmetic.

Hypothesis 2. We expected the performance advantage of word problems in fraction arithmetic to be affected by students' conceptual fraction knowledge: the lower students' conceptual fraction knowledge, the larger the performance advantage of word problems.

Hypothesis 3. We expected the performance advantage of word problems in fraction arithmetic to decrease when students move from primary to secondary school, due to the discontinuous learning trajectory. Therefore, we expected the word-problem advantage to be highest in sixth graders and lower in seventh and eighth graders.

2. Materials and methods

2.1. Participants

In total, 498 students from nine primary schools and nine secondary from different regions in the Netherlands participated. Thirteen students (one sixth grader, six seventh graders and six eighth graders were not able to reach the end of the task (a string of the last eight problems or more were left unanswered) and were excluded from further analysis. The effective sample thus consisted of 485 students: 106 sixth graders, 187 seventh graders, and 192 eighth graders. Table 1 presents descriptive statistics of the students' gender and age. There were seven to 68 students participating per school (average 26.9 students per school). The research protocol was approved by the Institute's IRB [number ECPW-2016/141] and only children with written parental consent and individual consent participated.

For 77 sixth graders whose parents gave consent to register background information we collected their most recent score on the standardized mathematics subtest of CITO's student monitoring system (Hop et al., 2017) divided in five population-referenced quantiles I–V. There were 28 students (36%) in the highest quantile I, 14 students (18%) in quantile II, 17 students (22%) in quantile III, 10 students (13%) in quantile IV, and 8 students (10%) in the lowest quantile V.

From secondary education, there were 74 students from the highest track of the prevocational education program ("VMBO-theoretisch"), 26 students from the combined track of prevocational education theoretical program and senior general secondary education ("VMBO-theoretisch/HAVO"), 89 from senior general secondary education ("HAVO"), 58 from the combined track of senior general secondary education and pre-university education ("HAVO/VWO"), and 132 from pre-university education ("VWO"). Note that there were no secondary school students from the lowest prevocational education programs.

2.2. Materials and design

A 40-item task was developed, consisting of 24 fraction arithmetic problems, eight whole-number arithmetic problems, and eight conceptual fraction knowledge problems. There were eight different task booklets, constructed by crossing three factors: (a) two different random orders of the arithmetic problems (whole-number and fraction arithmetic problems were mixed, (b) two different orders of which problems were presented first: arithmetic problems or conceptual fractions knowledge problems, and (c) two different versions of which arithmetic problem version was presented in symbolic format and which in word-problem format (see below). In the task a maximum of three problems were printed on a page (A4 size).

2.2.1. Fraction arithmetic problems

Table 2 shows the 24 fraction arithmetic problems, covering addition, subtraction, multiplication, and division. There were 12 problem types and for each problem type there were two parallel versions *a* and *b* with numbers and solution steps as similar as possible (see Table 2). Either problem versions *a* were presented in word-problem format and versions *b* in symbolic format or vice versa. The word problems were similar to those in Dutch educational textbooks and assessments, which are mostly written in a triparted structure, for instance: "Jesse still has $\frac{5}{6}$ of a loaf or bread. He eats $\frac{2}{3}$ of a loaf of bread. How much loaf of bread does he has left?"

To ensure measuring fraction *arithmetic* instead of other fraction competences (e.g., regrouping of improper fractions or simplifying) a few restrictions were made on the fraction arithmetic problems. The problems were constructed so that simplification of the operands and outcome was not possible or necessary. Mixed fractions (e.g. $1\frac{1}{2}$, $3\frac{3}{4}$, etc.) or improper fractions (e.g, $\frac{8}{3}$) were not included as operands or outcomes. The fraction denominator had a maximum value of 12, to ensure the face validity of the fractions regarding real-life situations and to keep the complexity of the fraction problems manageable.

Table 1
Descriptive statistics of students' gender and grade.

	N	boys	girls	age *M* (SD)
grade 6	106	51	56	11.66 (0.450)
grade 7	187	101	72	12.53 (0.663)
grade 8	192	99	86	13.56 (0.448)

Note: Missing information on students' gender 1–7%; missing information on students' age 2–7%.

Table 2
Fraction arithmetic and whole-number arithmetic problems.

Fraction arithmetic problem		Version *a*	Version *b*
A1	Fraction + fraction (*equal* denominators)	$\frac{2}{9}+\frac{5}{9}$	$\frac{2}{7}+\frac{3}{7}$
A2	Fraction + fraction (*unequal* denominators)	$\frac{1}{4}+\frac{3}{8}$	$\frac{1}{5}+\frac{7}{10}$
S1	Fraction – fraction (*equal* denominators)	$\frac{5}{7}-\frac{2}{7}$	$\frac{3}{5}-\frac{1}{5}$
S2	Fraction – fraction (*unequal* denominators)	$\frac{8}{9}-\frac{1}{3}$	$\frac{5}{6}-\frac{2}{3}$
M1	Fraction x whole number	$9\times\frac{1}{3}$	$8\times\frac{1}{4}$
M2	Whole number x fraction	$\frac{4}{7}$ x 210	$\frac{5}{12}$ x 240
M3	Fraction x fraction (*equal* denominators)	$\frac{7}{10}$ x $\frac{3}{10}$	$\frac{1}{5}$ x $\frac{3}{5}$
M4	Fraction x fraction (*unequal* denominators)	$\frac{1}{2}$ x $\frac{5}{9}$	$\frac{3}{8}$ x $\frac{9}{10}$
D1	Fraction: whole number	$6:\frac{3}{4}$	$3:\frac{3}{5}$
D2	Whole number: fraction	$\frac{6}{7}:3$	$\frac{8}{11}:4$
D3	Fraction: fraction (*equal* denominators)	$\frac{4}{5}:\frac{2}{5}$	$\frac{6}{7}:\frac{2}{7}$
D4	Fraction: fraction (*unequal* denominators)	$\frac{3}{4}:\frac{1}{8}$	$\frac{2}{3}:\frac{1}{6}$
Whole-number arithmetic problem			
A	Addition	283 + 368	386 + 238
S	Subtraction	432 – 185	423 – 158
M	Multiplication	24 × 36	23 × 34
D	Division	238 : 14	216 : 12

The problems were constructed to systematically cover different complexity factors in fraction arithmetic (Bruin-Muurling, 2010). For each of the four operations, problems with fractions with equal denominators and fractions with unequal denominators were included. Moreover, in multiplication and division problems there were two more problems with a whole number and a fraction as operands, in two different orders. Those mixed-operand problems were not included in addition and subtraction problems, because that would lead to mixed and improper fractions.

The inter-item reliability of these 24 fraction arithmetic problems was high, λ-*2* = 0.842.

2.2.2. Whole-number arithmetic problems

The eight whole-number arithmetic problems (Table 2) were selected from the study by Hickendorff (2021) and consisted of four problems with two parallel versions *a* and *b*, with numbers and solution steps as similar as possible. In each test booklet, half of the problems were presented in word-problem format and half in symbolic format. Like in fraction arithmetic, each task version contained either problem versions *a* in word-problem format and versions *b* in symbolic format, or vice versa. Again, the word problems were similar to those in Dutch mathematics textbooks and tests, for instance "Elisa has been on a trip. The hotel costed 283 euros, and she also spent 368 euros. How much did the trip cost in total?" The inter-item reliability of these eight problems was rather low, λ-*2* = 0.593, possibly due to the small number of problems. However, it is important to note that the statistical analyses did not use the scale scores but rather the scores on individual problems.

2.2.3. Conceptual fraction knowledge problems

There were eight conceptual fraction knowledge problems, where, in contrast to the fraction arithmetic problems, no computations were needed. Four problems addressed fraction magnitude knowledge in part-whole representations. In two of these problems students had to name the fraction that corresponded to the shaded part of a circle: two out of seven equally sized parts (correct answer $\frac{2}{7}$; and one out of five equally sized parts (correct answer $\frac{1}{5}$). In the other two problems students had to shade the part of a circle corresponding to a given fraction themselves: $\frac{7}{10}$ and $\frac{3}{8}$.

There were two problems addressing fraction magnitude knowledge in number-line representation. In these multiple-choice items, students had to circle the correct arrow pointing to a specific fraction ($\frac{3}{5}$ and $\frac{1}{4}$) on the 0–1 number line, segmented in ten equally sized segments. Finally, there were two fraction comparison problems, in which students had to compare two fractions, $\frac{3}{4}$ or $\frac{4}{9}$ and $\frac{2}{5}$ or $\frac{2}{7}$, and circle the largest fraction.

The inter-item reliability of the eight-item scale was rather low (λ-2 = 0.602) and the distribution very skewed to the left. Therefore, in the analyses we used a categorical variable: CFK level I (0–6 problems correct; $n = 105$), level II (7 problems correct, $n = 131$) and level III (8 problems correct, $n = 249$).

2.3. Procedure

The test was administered as a paper-and-pencil test in a classroom setting by one of eleven trained research assistants. The eight different task booklets were randomly distributed within the classrooms. Students were instructed that they could work through the test booklet at their own speed, within a 50-minute time slot. Furthermore, students were instructed that they could use the space next to each problem to write down their solution steps if they wanted to, and that they had to write their final answer to each problem on the designated answer line. Furthermore, they were instructed to give their answer either as fraction or whole number, but not as decimal number. Students' answers to each of the 40 problems were coded as either correct (1) or incorrect (0). Problems with no answer were scored as incorrect, assuming that students skipped the problem because they did not know how to solve it, as is common in untimed educational assessments such as TIMSS (Foy et al., 2019) and PISA (OECD, 2020). Improper fractions (e.g. $\frac{6}{2}$) and non-simplified fractions (e.g., $\frac{6}{8}$) that were equivalent to the correct answer were scored as correct.

2.4. Statistical analyses

The dependent variable in each of the research questions was arithmetic performance. In the current design, items are fully crossed with students, and students are nested in schools. To fully use all the information in the data, we did not aggregate the responses across items into scale scores, but instead used the responses of individual students to individual items as the unit of analysis. That is, we estimated a Rasch model, which is an item response theory (IRT) model in which differences between items are captured with fixed item effects (item difficulty/easiness parameters) and differences between individuals with a random student intercept (also called ability). To account for the fact that students were nested in schools, and in secondary schools also within tracks, the random student intercept was split into two components: one random intercept accounting for variation between students within tracks-within-schools, and one accounting for variation between tracks-within-school (37 different categories). Thus, we estimated the multilevel extension of the Rasch model (Doran et al., 2007).

To explain variance on the level of persons and/or items, predictor variables can be added to the regression, which has been called explanatory IRT modeling (De Boeck & Wilson, 2004). Such an approach has been used before in other studies addressing the effects of task/-problem features in the domain of mathematics (Fagginger Auer et al., 2016, 2018; Hickendorff, 2020), reading (Pavias et al., 2016), and analogical reasoning (Stevenson et al., 2013).

In all analyses in the current study, we estimated multilevel Rasch models on the correct/incorrect scores on the 24 fraction arithmetic problems (hypotheses 1–3) or the eight whole-number arithmetic problems (hypothesis 1). For hypotheses 2 and 3, students' conceptual fraction knowledge (CFK) level (three categories) and grade (three categories) were added as predictors to this model, respectively.

Such multilevel Rasch models can be conceptualized as a random effect multivariate logistic regression models, and can therefore be estimated with software for generalized linear mixed models (Rijmen et al., 2003). We used the glmer-function in the lme4-package to estimate the multilevel Rasch models (Bates et al., 2015; Doran et al., 2007) in R (R Core Team, 2021). Predictor effects with one degree of freedom were tested for significance with Wald tests and predictor effects with more than one degree of freedom (e.g., the main effect of categorical predictors such as grade) were tested with likelihood ratio (LR) tests. LR-tests statistically evaluate the improvement in model fit (log-likelihood) of the more complex model containing a specific predictor compared with the simpler model without that predictor, with a chi-square statistic. All tests were based on a Type I error probability of .05.

To control whether the eight different task booklets had the same difficulty level, we first analyzed the effect of Task Booklet (eight categories), which was not significant for fraction arithmetic (LR χ^2 (df = 7) = 5.60, p = .587) nor for whole-number arithmetic (LR χ^2 (df = 7) = 8.60, p = .283).

3. Results

Table 3 presents descriptive statistics of performance on all arithmetic problems. For whole-number arithmetic, there were only minimal differences between symbolic and word problems – although the difference of .08 for subtraction problems is an exception. For fraction arithmetic the differences between symbolic and word problems are somewhat larger, but more importantly, they are in both directions. Particularly the division problems stand out with higher performance on word problems than on their symbol counterparts.

3.1. Hypothesis 1: Word-problem advantage in fraction arithmetic performance

First, we analyzed the fraction arithmetic scores with a multilevel Rasch model with random intercepts for students and (tracks within) schools and fixed item effects β_i – which represent item *easiness* parameters in the glmer-parametrization. The variance of the random

Table 3
Descriptive statistics of performance (proportion correct) for fraction and whole-number arithmetic.

Fraction artithmetic problem		Symbolic	Word problem
A1	Fraction + fraction (*equal* denominators)	.91	.81
A2	Fraction + fraction (*unequal* denominators)	.81	.75
S1	Fraction – fraction (*equal* denominators)	.91	.82
S2	Fraction – fraction (*unequal* denominators)	.78	.76
M1	Fraction x whole number	.36	.71
M2	Whole number x fraction	.70	.79
M3	Fraction x fraction (*equal* denominators)	.48	.11
M4	Fraction x fraction (*unequal* denominators)	.56	.13
D1	Fraction: whole number	.50	.68
D2	Whole number: fraction	.27	.54
D3	Fraction: fraction (*equal* denominators)	.30	.83
D4	Fraction: fraction (*unequal* denominators)	.31	.70
Whole-number arithmetic problem			
A	Addition	.88	.87
S	Subtraction	.79	.71
M	Multiplication	.66	.69
D	Division	.66	.67

student-within-schools intercept was 0.784 and of the random school intercept 0.624. Table 4 presents the item easiness parameters β_i, which ranged between −2.77 and 2.67. Furthermore, it presents the difference in easiness parameters between the two problem formats of the same problem type $\beta_{diff} = \beta_{WP} - \beta_{sym}$, with its associated standard error and Wald test. Positive values of β_{diff} indicate that the word-problem format was easier than the symbolic counterpart of the same problem type, whereas negative values of β_{diff} indicate the opposite pattern. These differences were statistically significant for all but one (S2) of the twelve problem types. Notably, the direction of the difference was not consistent. For the two addition problems, one of the two subtraction problems and two of the four multiplication problems, word problems were significantly more difficult than their counterparts presented as symbolic problems. For the other two multiplication problems and the four division problems, word problems were significantly easier than their counterparts presented as symbolic problems. Fig. 2 shows a graphical representation of these differences (where item easiness parameters are transformed into the estimated probability correct per item, for easier interpretation).

To control whether these representation effects are unique for fraction arithmetic, we conducted a similar analysis on the whole-number arithmetic problems. Again we first estimated a multilevel Rasch model, this time with the correct/incorrect scores on the eight whole-number problems as dependent variables. The variance of the random student intercept was 0.795 and of the random school intercept 0.130. Item easiness parameters ranged between 0.75 and 2.31 (Table 4). The differences between symbolic and word problems were statistically significant for the subtraction problem only (word problem more difficult than symbolic problem); for the other three problem types differences were not significant. Fig. 3 shows the estimated probability correct on the whole-number arithmetic problems.

In all, findings for hypothesis 1 were mixed. We expected a word-problem advantage in fraction arithmetic which we indeed found in half of the fraction multiplication problems and in all fraction division problems. However, in fraction addition, one of the fraction subtraction problems, and half of the fraction multiplication problems the opposite pattern was found, with word problems being more difficult than symbolic problems. In whole-number arithmetic there was no word-problem advantage in addition, multiplication, or division (as expected); in subtraction the word problem was significantly more difficult than the symbolic problem.

Table 4
Item easiness parameters β_i and differences in item easiness between word problems and symbolic problems $\beta_{diff} = \beta_{WP} - \beta_{sym}$ from multilevel Rasch models for fraction arithmetic and whole-number arithmetic.

Fraction arithmetic					
Problem	Symbolic	Word problem	Difference (word problem – symbolic)		
	β_i (SE)	β_i (SE)	β_{diff} (SE)	z	p
A1	2.64 (0.218)	1.70 (0.190)	−0.94 (0.210)	−4.49	<.001
A2	1.70 (0.190)	1.26 (0.183)	−0.44 (0.174)	−2.52	.012
S1	2.67 (0.219)	1.75 (0.191)	−0.92 (0.212)	−4.34	<.001
S2	1.48 (0.186)	1.35 (0.184)	−0.13 (0.171)	−0.78	.438
M1	−0.87 (0.176)	0.98 (0.179)	1.85 (0.155)	11.92	<.001
M2	0.93 (0.179)	1.56 (0.187)	0.62 (0.167)	3.73	<.001
M3	−0.26 (0.174)	−2.77 (0.213)	−2.50 (0.191)	−13.14	<.001
M4	0.17 (0.174)	−2.49 (0.204)	−2.66 (0.181)	−14.70	<.001
D1	−0.18 (0.173)	0.79 (0.177)	0.96 (0.150)	6.41	<.001
D2	−1.45 (0.181)	0.04 (0.174)	1.49 (0.155)	9.59	<.001
D3	−1.22 (0.179)	1.82 (0.193)	3.04 (0.175)	17.41	<.001
D4	−1.20 (0.178)	0.93 (0.179)	2.13 (0.158)	13.48	<.001

Whole-number arithmetic					
Problem	Symbolic	Word problem	Difference (word problem – symbolic)		
	β_i (SE)	β_i (SE)	β_{diff} (SE)	z	p
A	2.31 (0.167)	2.14 (0.160)	−0.17 (0.200)	−0.83	.405
S	1.50 (0.141)	1.05 (0.13)	−0.45 (0.159)	−2.84	.005
M	0.75 (0.129)	0.93 (0.131)	0.18 (0.148)	1.21	.225
D	0.75 (0.129)	0.82 (0.130)	0.07 (0.147)	0.45	.652

3.2. *Hypothesis 2: Students' conceptual fraction knowledge*

We added the effect of Conceptual Fraction Knowledge (CFK in three categories: level I, II or III) to the fraction arithmetic model multilevel logistic regression model on the fraction arithmetic problems. The main effect of CFK was significant (LR χ^2 ($df = 2$) = 36.57; $p < .001$). Students with CFK at level I performed significantly lower than students at level II ($\beta_{\text{I-II}} = -0.56$ (SE = 0.138), $z = -4.08$, $p < .001$) and level III ($\beta_{\text{I-III}} = -0.78$ (SE = 0.128), $z = -6.11$, $p < .001$); the difference between the latter two levels was not significant ($\beta_{\text{II-III}} = -0.22$ (SE = 0.115), $z = -1.89$, $p = .059$).

Next, Table 5 shows the results of the statistical tests of the interaction effects between students' CFK and Problem Representation, per problem type. On four of the twelve problems, this interaction effect was significant — A1, M2, M3, and D4 — meaning that the problem representation effect was moderated by students' level of conceptual fraction knowledge for these problems.

Two of these problems showed an advantage of symbolic problems (A1 and M3) in the previous analyses. Post-hoc analyses for problem A1 (adding two fractions with equivalent denominators) showed that only for students with CFK level III had a significant advantage of symbolic problems ($\beta_{diff} = -1.76$) whereas problem representation differences were not significant or students with CKF level I or level II. Post-hoc analyses for problem M3 (multiplying two fractions with equivalent denominators) showed that for students with CFK level I ($\beta_{diff} = -3.61$), level II ($\beta_{diff} = -2.13$), and level III ($\beta_{diff} = -2.40$) there was a significant advantage of symbolic problems; this advantage was largest for students at CFK level I.

The other two problems showed an advantage of word problems (M2 and D4) in the previous analyses. Post-hoc analyses for problem M2 (multiplying a fraction with a whole number) showed that only for students with CFK level II ($\beta_{diff} = 1.36$) and level III ($\beta_{diff} = -0.75$) had a significant advantage of word problems, whereas problem representation differences were not significant for students with CKF level I. Post-hoc analyses for problem D4 (dividing two fractions with non-equivalent denominators) showed that for students with CFK level I ($\beta_{diff} = 1.37$), level II ($\beta_{diff} = 2.19$), and level III ($\beta_{diff} = 2.41$) there was an advantage of word problems; this advantage was larger for students with higher levels of CFK.

In all, these results are largely in contrast with hypothesis 2, since we expected a decrease in the word-problem advantage for higher levels of conceptual fraction knowledge. For three of the four items (M2, M3 and D4) the opposite pattern was found, with an increased word-problem advantage or decrease in symbolic advantage for higher levels of conceptual fraction knowledge. Only one problem (A1) showed the expected pattern, with an increase in symbolic advantage for students with higher levels of conceptual fraction knowledge.

3.3. *Hypothesis 3: Students' grade*

Finally, we added the effect of students' Grade (three categories: sixth, seventh, or eighth grade) to the fraction arithmetic model. The main effect of Grade was not significant (LR χ^2 ($df = 2$) = 5.54; $p = .063$). Table 6 shows the results of the statistical tests of the interaction effects between Problem Representation and students' Grade. This interaction was significant for two of the twelve problem types: for A1 and M2. In problem A1 (adding two fractions with equal denominators), for which an advantage of symbolic problems was found in the overall analyses, post-hoc analyses showed that symbolically presented problems were significantly easier than word problems only in seventh ($\beta_{diff} = -1.47$) and eighth grade ($\beta_{diff} = -1.20$), whereas there was no performance

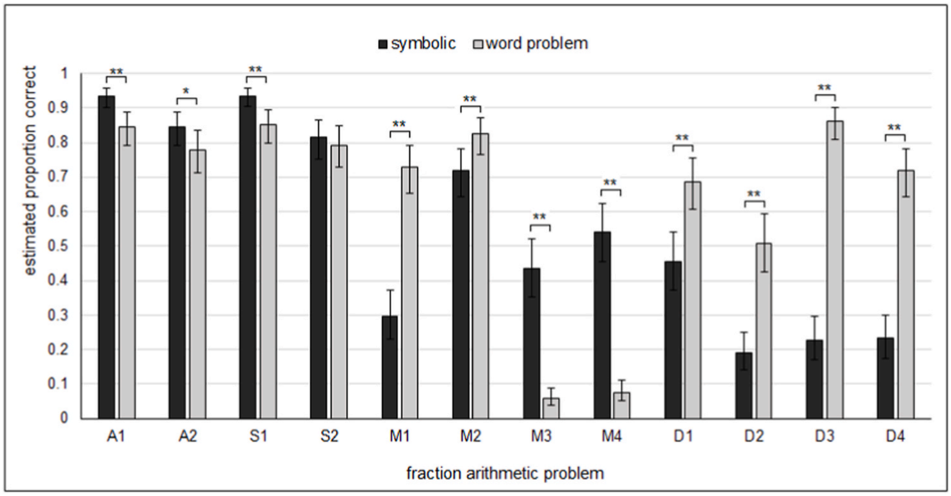

Fig. 2. Estimated proportion correct on fraction arithmetic problems, with 95% confidence intervals. Problem labels from Table 2. *p < .05; **p < .01.

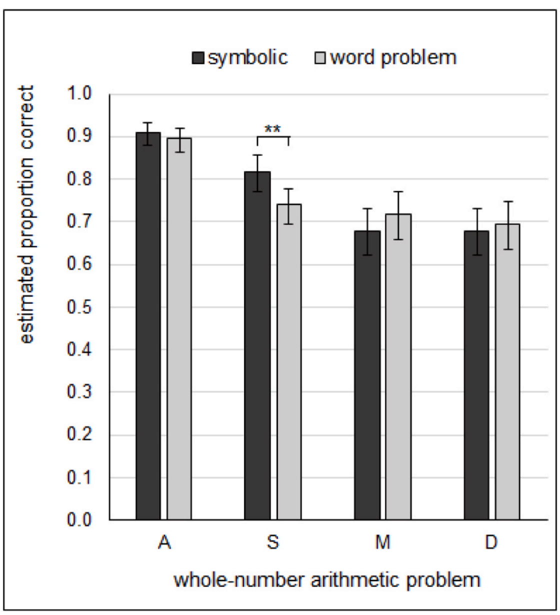

Fig. 3. Estimated proportion correct on whole-number arithmetic problems, with 95% confidence intervals. *p < .05; **p < .01.

difference in sixth grade. For problem M2 (multiplying a whole number with a fraction), for which an advantage of word problems was found in the overall analyses, post-hoc analyses showed that word problems were significantly easier than their symbolic counterparts only in seventh ($\beta_{diff} = 0.63$) eighth grade ($\beta_{diff} = 1.12$), whereas there was no performance difference in sixth grade.

In all, the performance patterns displayed in Fig. 2 hold for students in sixth, seventh, and eighth grade on most of the problems, which is contrary to hypothesis 3. On one of the two problems with differential effects the pattern aligned with our hypothesis (increase in the advantage of symbolic problems for secondary students) but on the other problem the pattern was opposite to what we expected (increase in the advantage of word problems for secondary students).

4. Discussion

Fractions are a critical concept in mathematics education. They play a central role in the transition from early to more advanced mathematics and in the use of mathematics in other fields and in everyday life. However, fraction problems are also notoriously difficult for students, likely due to their abstractness which hampers understanding (Brown & Quinn, 2006; Hiebert, 1985). The main hypothesis of the current study was that situating fractions arithmetic in a realistic story might help students overcoming these difficulties. This makes fractions and fraction arithmetic more concrete, and as such can help students to use mental models that fit the problem situation, as was shown in algebra

Table 5
Likelihood Ratio (LR) tests for the interaction effects between Problem Representation and students' Conceptual Fraction Knowledge (CFK) in fraction arithmetic, with post-hoc tests of the difference in item easiness $\beta_{diff} = \beta_{WP} - \beta_{sym}$ per level of CFK.

Problem	LR chi^2 ($df = 2$)	p	CFK level I			CFK level II			CFK level III		
			β_{diff} (SE)	z_{diff}	p	β_{diff} (SE)	z_{diff}	p	β_{diff} (SE)	z_{diff}	p
A1	10.75	.004	−0.64 (0.359)	−1.79	.074	−0.16 (0.400)	−0.40	.686	−1.76 (0.337)	−5.22	<.001
A2	3.55	.169									
S1	0.68	.711									
S2	5.85	.053									
M1	5.45	.065									
M2	13.95	<.001	−0.31 (0.318)	−0.96	.337	1.36 (0.323)	4.21	<.001	0.75 (0.251)	2.99	.003
M3	8.01	.018	−3.61 (0.454)	−7.96	<.001	−2.13 (0.337)	−6.34	<.001	−2.40 (0.240)	−9.98	<.001
M4	2.97	.227									
D1	0.27	.872									
D2	1.23	.541									
D3	1.26	.533									
D4	6.48	.039	1.37 (0.337)	4.06	<.001	2.19 (0.300)	7.30	<.001	2.41 (0.225)	10.71	<.001

Table 6
Likelihood Ratio (LR) tests for the interaction effects between problem Representation and students' Grade in fraction arithmetic, with post-hoc tests of the difference in item easiness $\beta_{diff} = \beta_{WP} - \beta_{sym}$ per Grade.

Problem	LR chi^2 ($df = 2$)	p	Grade 6			Grade 7			Grade 8		
			β_{diff} (SE)	z_{diff}	p	β_{diff} (SE)	z_{diff}	p	β_{diff} (SE)	z_{diff}	p
A1	7.63	.022	−0.13 (0.366)	−0.37	.714	−1.47 (0.356)	−4.13	<.001	−1.20 (0.359)	−3.35	<.001
A2	2.31	.315									
S1	3.10	.212									
S2	2.26	.323									
M1	1.19	.550									
M2	6.92	.031	0.00 (0.315)	0.00	1.00	0.63 (0.272)	2.315	.021	1.12 (0.285)	3.930	<.001
M3	1.82	.402									
M4	5.56	.062									
D1	4.45	.108									
D2	0.47	.791									
D3	0.81	.667									
D4	4.46	.107									

(Koedinger et al., 2008; Koedinger & Nathan, 2004). In the current study we therefore investigated the effect of problem representation (symbolic vs. word problem) on fraction arithmetic performance of students at the end of primary school and the start of secondary school. Furthermore, we contrasted the problem representation effect for fraction arithmetic with that for whole-number arithmetic, and investigated the potential moderating role of students' conceptual fraction knowledge and grade (sixth, seventh, eighth).

Our first hypothesis, that there would be a word-problem advantage in fraction arithmetic but not in whole-number arithmetic, was partly confirmed but also partly rejected. We found that in fraction arithmetic, problem representation effects were found in both directions. In adding, subtracting, and multiplying two fractions, word problems were more difficult than symbolic problems, contrary to our hypothesis. In multiplication of a fraction with a whole number and in all types of fraction division, word problems were easier than their symbolic counterparts. In contrast to these results for fraction arithmetic, in whole-number arithmetic word problems were just as difficult (addition, multiplication, and division) or more difficult (subtraction) compared to their symbolic counterparts, but never easier, replicating previous studies' results (Hickendorff, 2013, 2021). In all, it thus seems that there may indeed be a word-problem advantage in fraction arithmetic, but only in particular problems and particularly in fraction division, where the differences were large, whereas in addition and subtraction we found a rather small advantage of symbolic problems. This pattern suggests that the more abstract the fraction arithmetic procedure, the larger the likelihood that there is a word-problem advantage (with multiplication as an exception, which will be discussed later).

Our second hypothesis, that the word-problem advantage in fraction arithmetic would decrease as students have higher conceptual fraction knowledge, was not supported. Although students with higher levels of conceptual fraction knowledge had higher performance in fraction arithmetic overall, conceptual fraction knowledge did not moderate the problem representation effect in eight of the twelve problems. In the four problems in which we did find an interaction between students' conceptual fraction knowledge and problem representation, we found a pattern opposite to our expectations, with an increased word-problem advantage or decrease in symbolic advantage for higher levels of conceptual fraction knowledge, for three problems. Only one problem showed the expected pattern.

Our third hypothesis, that the word-problem advantage in fraction arithmetic would decrease across grades six to eight as students move to secondary education, was also not supported. There were no significant differences in overall fraction arithmetic performance between sixth, seventh, and eighth graders. Grade did not moderate the problem representation effect in ten of the twelve problems. In the two problems for which we did find an interaction effect between students' grade and problem representation, the pattern was in the expected direction for one problem and in the opposite direction for the other problem.

4.1. Word problems in fraction arithmetic

Our main expectation was that situating fraction arithmetic in a realistic story or setting would make fractions easier to understand and fraction arithmetic easier to perform and as such would help overcome students' difficulties. For fraction division we indeed found a robust word-problem advantage. Since fraction division has the most abstract and complex formal procedure of the four operations (Siegler & Lortie-Forgues, 2015), it is not surprising that this is the operation where we find a word-problem advantage. Situating the fraction division problem in a concrete setting may elicit mental models that fit the problem situation, which can lead students to use informal strategies that they understand rather than the formal procedures they do not understand as well as students checking the plausibility of their answers (Koedinger et al., 2008; Koedinger & Nathan, 2004). Furthermore, the multiplication problems that required multiplying a fraction with a whole number were also significantly easier as word problem than as symbolic problem. Perhaps the realistic situation also elicited other mental models and solution strategies, such as seeing than multiplying with $\frac{1}{3}$ is the same as dividing by 3. Further research should address students' solution strategies and error types, for instance by think-aloud protocols, to put this explanatory mechanism to the test.

However, in adding, subtracting or multiplying two fractions (with equivalent or non-equivalent denominators), we found an advantage of symbolic representation, contrary to our expectations. One possible explanation is that these problems were not as abstract as we expected them to be; in other words, students may have been taught the formal procedures to solve them and thus needed to rely less on their mental models. Another explanation is that perhaps students used more informal strategies on the word problems, and that these were less efficient than the formal procedures. Finally, the fraction-times-fraction problems showed a relatively large advantage of symbolic representation. A possible explanation is that the wording of these word problems was confusing for students. Consider the following word problem "From all children in a class, $\frac{1}{2}$ plays a musical instrument. From all the children playing a musical instrument, $\frac{5}{9}$ th plays the piano. Which part of the children in the class plays the piano?" Inspection of students' answers shows that many students answered $\frac{5}{9}$, suggesting that they did not understand that they had to take part of a part. Perhaps this is an example of a semantically disaligned word problem which have been found to be more difficult to solve than semantically aligned word problem (Martin & Bassok, 2005). It illustrates that situating fraction arithmetic only helps if it elicits a correct mental representation, and otherwise may hinder performance.

4.2. Conceptual fraction knowledge

We expected that students who have difficulties understanding the concept of fractions would be helped by presenting fraction arithmetic problems in a situated way instead of symbolically. However, we did not find this pattern for eight of the problems, and the pattern for three other problems was even in the opposite direction. Possibly, understanding individual fractions' magnitudes is not a strong predictor of understanding how arithmetical operations transform those magnitudes. Further research could focus on students' understanding of fraction arithmetic procedures (Siegler & Lortie-Forgues, 2014). Another possible explanation is that students may need conceptual knowledge to translate the word problems into the appropriate mental representation and arithmetic problem (Rittle-Johnson et al., 2001).

An important question that the current study cannot address is the extent to which situating fraction arithmetic in real-life settings can contribute to *acquiring* fraction knowledge, as is suggested by Realistic Mathematics Education (Van den Heuvel-Panhuizen et al., 2014). Through such better understanding of fractions and informal solution strategies, students are expected to understand the formal (symbolic) fraction arithmetic procedures better. A longitudinal study design in which students having ample experience in solving fraction word problems are followed in their development of conceptual and procedural fraction knowledge would be needed to address that question and arrive at corollary instructional implications.

4.3. Instructional experiences

We expected there would be a difference between students at the end of primary school on the one hand and students at the start of secondary school on the other in the anticipated word-problem advantage. Remarkably, we did not find a main effect of grade, which suggests that the development of students' fraction arithmetic competence stagnates in the end of primary school, as was also reported by Bruin-Muurling (2010, chapter 3 and 4). This is surprising since fraction arithmetic is part of the mathematics curriculum in both the final grades of primary school and the lower grades of secondary school, and it is particularly striking since there were no secondary school students from the lowest educational track in our sample.

For two of the twelve problems, students' grade moderated the effect of problem representation: once in the expected direction and once in the opposite direction. In all, there were thus no robust patterns found. The overall absence of student progress in fraction arithmetic performance may also explain why we did not find robust differential effect of problem representation by grade. If on average students hardly learned something extra in secondary school it is not surprising that secondary school students show the same performance patterns as primary students.

4.4. Limitations

There are several limitations that are important to consider. First, we used parallel problems to investigate differences between symbolic and word problems. Although these problems were carefully matched so that the numbers and solution steps were as similar as possible, there could have been systematic differences between the parallel versions. However, since the sample size is large and the different problem version-problem format combinations were distributed randomly across the students, such potential systematic differences between parallel problems should level out in the analyses at group level, which is supported by the finding that the eight different task booklets did not differ in difficulty level.

Second, we included a specific set of real-life situations, which could have impacted the results. Specifically, the situation that we used for multiplying two fractions seems to be unfortunate, since many students misunderstood it, as mentioned before. If this problem would have been situated in another story in which it is clear to students that they have to take part of a part, such as "How much is $\frac{1}{2}$ of $\frac{5}{9}$ pie?" (Bruin-Muurling, 2010), potentially the word-problem advantage found for division would also extend to multiplication. Furthermore, it illustrates the importance of presenting a situation that, through the mental model, supports using an appropriate solution strategy. Further research could address the match between the situation model and the solution strategies to systematically investigate this suggested mechanisms. Moreover, in future studies more efforts should be undertaken to carefully pilot the problems' numerical characteristics (to accommodate the previous limitation) and situational descriptions.

Third, we did not study all types of fractions, thereby excluding several factors that make fractions abstract and complex. For instance, we excluded improper fractions such as $\frac{3}{2}$ or fractions that can be simplified such as $\frac{6}{8}$, whereas research findings suggest that these kind of fractions may be particularly challenging for students (Bruin-Muurling, 2010). It would be important to extend the current study to fraction arithmetic involving these challenging fractions, such as $\frac{3}{4}+\frac{3}{4}=\frac{6}{4}=\frac{3}{2}=1\frac{1}{2}$. Possibly students are helped if this problem is situated in a concrete context that they can imagine, such as "How much pie do you have if you have three quarters of a pie and another three quarters of a pie?"

Fourth, we used grade level as a proxy for instructional experiences. Although instruction differs between grade levels as a result of general differences in the Dutch mathematics curriculum, this does not mean all students in the same grade have equal instructional experiences. Within grades, relevant differences exist in for example prior knowledge or differentiated instruction. For example, Braithwaite et al. (2017) found individual differences in patterns of strategy use in fraction arithmetic. Siegler and Pyke (2013) also found that differences in fraction arithmetic accuracy between high achieving students and low achieving students was larger in eight grade than in sixth grade. We did not address such within-grade differences in the current study.

Fifth and finally, our analyses are restricted to performance (accuracy) only, which does not give insight into affective components of learning, such as self-efficacy and motivation, or students' *process* of solving the problems. Since the suggested mechanism of the expected word-problem advantage focuses on different solution processes, the fact that we do not have information on the solution strategies or the errors severely limits the interpretation of the results. Earlier research showed there were individual differences in fraction arithmetic strategy use in symbolic fraction arithmetic problems (Braithwaite et al., 2017) and it would be interesting to extend this with word problems in future research. Furthermore, such research could investigate to what extent primary school and secondary school students used the same strategies or whether they used different strategies but with the same overall success ratio. Solution strategy data could also give insights into the mental models that students construct in the word-problem solving process.

4.5. Conclusion

Situating fraction arithmetic in a realistic setting might help students overcome the difficulties they experience in this domain, but only under specific conditions: when the formal procedure is abstract or difficult to understand (as in division) and if the situation presented supports the construction of an appropriate mental model. Importantly, these results challenge common beliefs that word problems are particularly challenging for students because they have to translate the word problem in an arithmetic problem. However, for less abstract or challenging fraction arithmetic procedures, there may be an advantage of symbolic representation. In all, the current study suggests that word problems that are carefully selected can aid fraction arithmetic performance, particularly for complex and abstract arithmetic procedures.

Author contributions

Terry Mostert: Conceptualization; Investigation; Methodology; Project administration; Writing – original draft; Writing – review & editing.

Marian Hickendorff: Conceptualization; Investigation; Formal analysis; Methodology; Writing – original draft; Writing – review & editing.

Declaration of interest statement

The authors report there are no competing interests to declare.

Acknowledgements

We are indebted to all undergraduate students who contributed to data collection.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Package lme4. *Journal of Statistical Software, 67*(1), 1–91. http://lme4.r-forge.r-project.org.

Braithwaite, D. W., Pyke, A. A., & Siegler, R. S. (2017). A computational model of fraction arithmetic. *Psychological Review, 124*(5), 603–625. https://doi.org/10.1037/REV0000072

Brown, G., & Quinn, R. J. (2006). Algebra students' difficulty with fractions. *Australian Mathematics Teacher, 62*(4), 28–40. https://search.informit.org/doi/10.3316/informit.153305808535500.

Bruin-Muurling, G. (2010). *The development of proficiency in the fraction domain (Unpublished doctoral dissertation)*. the Netherlands: Technical University Eindhoven.

De Boeck, P., & Wilson, M. (2004). Explanatory item response models. In *Explanatory item response models*. New York: Springer. https://doi.org/10.1007/978-1-4757-3990-9.

Doran, H., Bates, D., Bliese, P., & Bowling, M. (2007). Estimating the multilevel Rasch model: With the lme4 package. *Journal of Statistical Software, 20*(2), 1–18. https://doi.org/10.18637/JSS.V020.I02

Fagginger Auer, M. F., Hickendorff, M., & van Putten, C. M. (2016). Solution strategies and adaptivity in multidigit division in a choice/no-choice experiment: Student and instructional factors. *Learning and Instruction, 41*, 52–59. https://doi.org/10.1016/j.learninstruc.2015.09.008

Fagginger Auer, M. F., Hickendorff, M., & van Putten, C. M. (2018). Training can increase students' choices for written solution strategies and performance in solving multi-digit division problems. *Frontiers in Psychology, 9*, 1644. https://doi.org/10.3389/fpsyg.2018.01644

Foy, P., Fishbein, B., von Davier, M., & Yin, L. (2019). *Chapter 12: Implementing the TIMSS 2019 scaling methodology. Methods and procedures: TIMSS 2019 technical report*. https://timssandpirls.bc.edu/timss2019/methods/pdf/T19_MP_Ch12-scaling-implementation.pdf.

Gabriel, F., Coché, F., Szucs, D., Carette, V., Rey, B., & Content, A. (2013). A componential view of children's difficulties in learning fractions. *Frontiers in Psychology, 4*(OCT), 1–12. https://doi.org/10.3389/fpsyg.2013.00715

Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology, 102*(2), 395–406. https://doi.org/10.1037/a0017486

van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 521–525). Dordrecht, Heidelberg, New York, London: Springer.

Hickendorff, M. (2013). The effects of presenting multidigit mathematics problems in a realistic context on sixth graders' problem solving. *Cognition and Instruction, 31*(3), 314–344. https://doi.org/10.1080/07370008.2013.799167

Hickendorff, M. (2020). Fourth graders' adaptive strategy use in solving multidigit subtraction problems. *Learning and Instruction, 67*, Article 101311. https://doi.org/10.1016/j.learninstruc.2020.101311

Hickendorff, M. (2021). The demands of simple and complex arithmetic word problems on language and cognitive resources. *Frontiers in Psychology, 12*, 4494. https://doi.org/10.3389/FPSYG.2021.727761/BIBTEX

Hiebert, J. (1985). Children's knowledge of common and decimal fractions. *Education and Urban Society, 17*(4), 427–437.

Hop, M., Janssen, J., & Engelen, R. (2017). *Wetenschappelijke verantwoording Rekenen-Wiskunde 3.0 voor groep 7 [Scientific report student monitoring system 3.0 mathematics subtest for grade 5]*. CITO, National Institute for Educational Measurement.

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32*(2), 366–397. https://doi.org/10.1080/03640210701863933

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences, 13*(2), 129–164. https://doi.org/10.1207/s15327809jls1302_1

Lenz, K., Dreher, A., Holzäpfel, L., & Wittmann, G. (2019). Are conceptual knowledge and procedural knowledge empirically separable? The case of fractions. *British Journal of Educational Psychology*. https://doi.org/10.1111/bjep.12333

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

Martin, S. A., & Bassok, M. (2005). Effects of semantic cues on mathematical modeling: Evidence from word-problem solving and equation construction tasks. *Memory & Cognition 2005, 33*(3), 471–478. https://doi.org/10.3758/BF03193064, 33:3.

Meelissen, M., Hamhuis, E., & Weijn, L. (2020). Netherlands. In D. L. Kelly, V. A. S. Centurino, M. O. Martin, & I. V. S. Mullis (Eds.), *TIMSS 2019 encyclopedia: Education policy and curriculum in mathematics and science* (pp. 1–9). Boston College: TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2019/encyclopedia/.

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

OECD. (2020). *PISA 2018. Technical report*. OECD Publishing. https://www.oecd.org/pisa/data/pisa2018technicalreport/.

Pavias, M., van den Broek, P., Hickendorff, M., Beker, K., & Van Leijenhorst, L. (2016). Effects of social-cognitive processing demands and structural importance on narrative recall: Differences between children, adolescents, and adults. *Discourse Processes, 53*(5–6). https://doi.org/10.1080/0163853X.2016.1171070

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/.

Rijmen, F., Tuerlinckx, F., De Boeck, P., & Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods*. https://doi.org/10.1037/1082-989X.8.2.185

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346–362. https://doi.org/10.1037/0022-0663.93.2.346

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691–697. https://doi.org/10.1177/0956797612440101

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences, 17*(1), 13–19. https://doi.org/10.1016/j.tics.2012.11.004

Siegler, R. S., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives, 8*(3), 144–150. https://doi.org/10.1111/cdep.12077

Siegler, R. S., & Lortie-Forgues, H. (2015). Conceptual knowledge of fraction arithmetic. *Journal of Educational Psychology, 209*(3), 374–386. https://doi.org/10.1037/edu0000025

Siegler, R. S., & Lortie-Forgues, H. (2017). Hard lessons: Why rational number arithmetic is so difficult for so many people. *Current Directions in Psychological Science, 26*(4), 346–351. https://doi.org/10.1177/0963721417700129

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49*(10), 1994–2004. https://doi.org/10.1037/a0031200

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*(4), 273–296. https://doi.org/10.1016/j.cogpsych.2011.03.001

Stevenson, C. E., Hickendorff, M., Resing, W. C. M., Heiser, W. J., & de Boeck, P. A. L. (2013). Explanatory item response modeling of children's change on a dynamic test of analogical reasoning. *Intelligence, 41*(3), 157–168. https://doi.org/10.1016/j.intell.2013.01.003

Torbeyns, J., Schneider, M., Xin, Z., & Siegler, R. S. (2015). Bridging the gap: Fraction understanding is central to mathematics achievement in students from three different continents. *Learning and Instruction, 37*, 5–13. https://doi.org/10.1016/j.learninstruc.2014.03.002

Van den Heuvel-Panhuizen, M., Drijvers, P., Education, M., Sciences, B., & Goffree, F. (2014). *Encyclopedia of mathematics education*. Encyclopedia of Mathematics Education. https://doi.org/10.1007/978-94-007-4978-8

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM - Mathematics Education, 52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4

Terry Mostert works as researcher at Leiden University and as mathematics teacher at a Dutch secondary school. Her research interests involves the primary and secondary mathematics curriculum. Marian Hickendorff works as an associate professor Educational Sciences at Leiden University. Her research ambition is to give empirical basis to several questions and discussions about primary school mathematics education. Furthermore, she is interested in the application of modern statistical techniques in research into learning and cognition.