Journal of Experimental Child Psychology 217 (2022) 105353

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

No transfer effect of a fraction number line game on fraction understanding or fraction arithmetic: A randomized controlled trial

Sevim Nuraydin *, Johannes Stricker, Michael Schneider

Department of Psychology, University of Trier, 54286 Trier, Germany

ARTICLE INFO

Article history:
Received 12 December 2020
Revised 12 December 2021
Accepted 13 December 2021

Keywords:
Fraction
Number line
Arithmetic
Intervention
Numerical cognition
game

ABSTRACT

Fractions are hard to understand for students. According to Siegler's integrated theory, magnitude understanding is central to numerical development. Whole-number magnitude understanding can be improved by serious games that include practicing mapping numbers onto number lines or similar one-dimensional visuospatial representations. These games were effective even when they did not include direct instruction on whole numbers. Previous studies have also evaluated whether similar number line estimation (NLE) interventions improve fraction learning. We contribute to this literature by evaluating how the catch-the-monster game with fractions affects NLE, magnitude understanding, and arithmetic with fractions in a randomized controlled pretest–intervention–posttest design. The game included NLE with feedback but no direct instruction. A sample of 188 fifth- to eighth-graders participated in a fraction number line game condition, a scaffolded fraction number line games condition, or an active control condition. In fifth- and sixth-graders, participation in the intervention improved the trained measure of fraction understanding (NLE on 0–1 line) but left unchanged two untrained measures of fraction understanding (NLE on 0–5 lines and fraction comparison) and a measure of fraction arithmetic. Seventh- and eighth-graders showed no intervention effects. The lack of transfer in fifth- and sixth-graders indicates that gamified NLE interventions can complement but not replace more direct instruction on fraction concepts and fraction arithmetic.

* Corresponding author.
E-mail address: nuraydin@uni-trier.de (S. Nuraydin).

https://doi.org/10.1016/j.jecp.2021.105353
0022-0965/

Introduction

Competence in using numerical fractions is crucial for the development of children's broader mathematical competence. Fraction competence strongly correlates with mathematical achievement. This has consistently been found in countries with different educational systems and on several continents (Torbeyns, Schneider, Xin, & Siegler, 2015). A study with a cross-lagged panel model showed that these correlations are due to a unidirectional influence of fraction competence on achievement (Bailey, Hoard, Nugent, & Geary, 2012). From Grade 6 to Grade 7, fraction competence predicted improvements in mathematical achievement, whereas mathematical achievement did not predict improvements in fraction competence. In two nationally representative datasets from the United States and the United Kingdom, fraction competence measured in elementary school predicted students' algebra competence and mathematical achievement in high school, that is, 5 to 6 years later (Siegler et al., 2012).

Given this importance of fraction competence, it is unfortunate that many children struggle with learning fractions (Lortie-Forgues, Tian, & Siegler, 2015; Vamvakoussi & Vosniadou, 2010). For example, only 50% of a nationally representative sample of U.S. eighth-graders correctly ordered 2/7, 5/9, and 1/12 from smallest to largest (Martin, Strutchens, & Elliott, 2007). Eighth-graders who were asked to choose the closest whole number of the sum 12/13 + 7/8 showed a similarly poor performance (i.e., only 27% provided the correct answer; Lortie-Forgues et al., 2015). Even adult community college students displayed an error rate of 30% for magnitude comparison problems, including fractions with unequal numerators and denominators (Schneider & Siegler, 2010).

In the current study, we investigated to what extent the *catch-the-monster game* (originally introduced by Rittle-Johnson, Siegler, & Alibali, 2001) improves fraction understanding and fraction arithmetic skills. The catch-the-monster game is an example of a serious game, that is, a game with the intention that the players reach certain learning goals by playing. The catch-the-monster game aims at improving mental numerical magnitude representations by letting learners practice to map numbers onto a visuospatial dimension and by giving learners feedback on the accuracy of the estimated positions. In each trial of the game, learners are presented with a number and a number line. Only the start and end of the number line are marked and labeled with a number. Participants are told that a monster is hiding at the position on the line indicated by the number and that they can catch the monster by moving a monster basket to this position. Thus, participants need to estimate the position of the number on the line. After entering their answer, the monster appears at the correct position and participants receive feedback on whether the monster was caught or escaped. The position of the monster shows participants the correct position of the number on the line. So far, the game has been used as a tool to study the development of conceptual and procedural knowledge about decimal fractions (Rittle-Johnson et al., 2001; Schneider & Stern, 2010) and as part of more complex fraction interventions that also included direct instruction on fraction concepts and procedures (Braithwaite & Siegler, 2021; Fazio, Kennedy, & Siegler, 2016). In our study, we investigated whether the game alone, without any explicit instruction, improves fraction magnitude understanding and fraction arithmetic.

The accuracy of students' estimates on the number line can then be computed as percent absolute error (PAE) with PAE = 100% $\times$ |correct position–estimated position| / range of line. Number line estimation with fractions requires participants to combine information from the numerator and denominator (Rinne, Ye, & Jordan, 2017; Schneider & Siegler, 2010). Students use various strategies for solving the task, for example, rounding or simplifying fractions to easier numbers (e.g., 7/70 to 1/10 to 0.1) or dividing the number line, for example, in whole-number units or denominator units (e.g., in five parts of equal length when estimating the position of 2/5) (Peeters, Degrande, Ebersbach, Verschaffel, & Luwel, 2016; Siegler, Thompson, & Schneider, 2011).

The game is based on theoretical ideas summarized in Siegler's *integrated theory of numerical development* (Siegler & Lortie-Forgues, 2014; Siegler et al., 2011), which poses that mental magnitude representations, the so-called mental number line, help learners to better understand numbers and to better understand how number symbols represent positions on a continuous magnitude dimension and how the magnitudes of numbers relate to each other. This requires learners to understand how the numerator and denominator together determine the numerical magnitude of a fraction. This understanding of numbers and their interrelations then aids further mathematical learning, for example, arithmetic learning and mathematical problem solving (Siegler, 2016).

In addition to Siegler's theory, there are at least four reasons to expect that the game can improve fraction understanding and arithmetic even when it is not combined with direct instruction. First, serious games in general have been shown to be effective in meta-analyses. Second, NLE interventions similar to the catch-the-monster game improved whole-number competence in previous studies. Third, NLE proficiency with fractions predicts fraction magnitude understanding, fraction arithmetic, and broader mathematical competence. Finally, interventions including fraction NLE or even the catch-the-monster game in combination with other visualizations, practice tasks, or direct instruction partly had positive causal effects on fraction understanding or arithmetic. We explain each of these arguments in turn.

Reasons to expect that fraction NLE games can improve fraction competence

The general effectivity of serious games

Meta-analyses found that, on average, serious games tend to be more effective than conventional instruction. A meta-analysis of 77 effect sizes found a positive effect on learning (d = 0.29) and retention (d = 0.36) but not on motivation (Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013). This finding is robust, as shown by a more recent meta-analysis of 57 effect sizes, which used different meta-analytic techniques and obtained a similar main effect of serious games on learning (Clark, Tanner-Smith, & Killingsworth, 2016). The effects were still positive when only studies in the domains of mathematics were analyzed (Wouters et al., 2013). The effects were stronger for schematic games, such as the catch-the-monster game, than for cartoon-like or photorealistic games. The effects were also stronger for games without a narrative, such as the catch-the-monster game, than for games with an evolving story (Clark et al., 2016; Wouters et al., 2013). Descriptively, the effects were lowest for games providing feedback only about the correctness of the students' answers, higher for feedback also presenting the correct answer (as done in the catch-the-monster game), and even higher for games combining this feedback with additional scaffolds such as cues about how to construct the correct problem solution in the game (Clark et al., 2016).

Still, serious games are no panacea, and likely many learning goals cannot be reached by playing serious games. Therefore, the meta-analyses provide only indirect evidence showing that the catch-the-monster game might be effective for improving fraction knowledge, but this hypothesis still needs to be tested directly. The meta-analyses also leave open whether the combination of games with additional non-game instruction is beneficial or not. The meta-analysis by Clark et al. (2016) found that games without non-game instruction are about equally effective (g = 0.32) as games with non-game instruction (g = 0.36). In the meta-analysis by Wouters et al. (2013), learners benefited from games with or without non-game instruction, but the effect size was about twice as large for games with additional instruction (d = 0.41) than for games without it (d = 0.20).

Whole-number NLE games improve whole-number understanding

A second reason to expect that the catch-the-monster game improves fraction understanding and arithmetic is that similar games improve whole-number understanding. In a randomized controlled experiment, playing a board game that required young children to map whole numbers onto a number-line-like structure on the board improved their magnitude understanding of these numbers (Siegler & Ramani, 2008). This effect was stronger for boards with a linear layout than for boards with a circular layout, indicating the importance of mapping numerical magnitudes onto a single spatial dimension (Siegler & Ramani, 2009). Effects of the board game transferred to whole-number magnitude comparison, counting, and numeral identification (Ramani & Siegler, 2008) as well as to whole-number arithmetic (Siegler & Ramani, 2009). Just one hour of gameplay led to improvements in magnitude understanding that could still be found at a follow-up test after nine weeks (Ramani

& Siegler, 2008). The games were effective not only when used by experimenters in a lab setting but also when used by practitioners as a small-group learning activity in classrooms (Ramani et al., 2012).

Fraction NLE correlates with broader mathematical competence measures

A third reason to expect that a fraction number line game can improve fraction competence is that fraction NLE is closely related to broader measures of fraction competence and general mathematical competence (Schneider, Thompson, & Rittle-Johnson, 2018b). A meta-analysis of 263 effect sizes obtained with more than 10,000 participants found that NLE proficiency and broader mathematical competence correlated with $r^+ = .443$, 95% confidence interval (CI) [.406, .480] (Schneider et al., 2018a). The correlation was even higher for NLE with fractions ($r^+ = .523$) than for NLE with whole numbers ($r^+ = .409$). The correlation was high when NLE and math competence were measured simultaneously ($r^+ = .427$) but also when NLE proficiency was used as a predictor of mathematical competence over time ($r^+ = .496$). The findings were robust in that NLE proficiency predicted counting, mental arithmetic, written arithmetic, school grades in mathematics, and scores on standardized mathematical achievement tests.

Fraction number line interventions

The fourth reason to expect that a fraction number line game would be effective is the effectiveness of previous fraction number line interventions. Three studies with fourth- and fifth-graders showed that several week-long interventions for students at risk for mathematical difficulties are more effective when they place greater emphasis on number lines and the measurement interpretation of fractions than regular school instruction does (Fuchs et al., 2013, 2014; Schumacher et al., 2018). Three other experiments with students from Grade 2 to Grade 5 found that explanations of fraction concepts combined with fraction visualization training were more effective when number lines were used than when area models (i.e., pie charts) were used in the interventions (Gunderson, Hamdan, Hildebrand, & Bartek, 2019; Hamdan & Gunderson, 2017; Tian, Bartek, Rahman, & Gunderson, 2021). These learning gains partly also transferred to fraction magnitude comparison. By contrast, experiments with 9- to 12-year-old students found that a very brief fraction NLE intervention did not improve fraction magnitude comparison compared with a pie chart condition (Hurst, Massaro, & Cordes, 2020).

Two studies by Siegler and colleagues (Braithwaite & Siegler, 2021; Fazio, DeWolf, & Siegler, 2016) used the catch-the-monster game to evaluate a new conceptual framework for teaching fractions (Braithwaite & Siegler, 2021), The framework, termed *putting fractions together,* "emphasizes that both individual fractions and sums of fractions are composed of unit fractions and can be represented by concatenating them (putting them together)" (Braithwaite & Siegler, 2021, p. 556). Fazio et al. (2016) investigated this idea in two empirical studies with 51 fifth-graders who played the catch-the-monster game with fractions. Before the game, students received conceptual instructions about unit fractions, about ways to segment the number line into the number of segments indicated by the denominator, and about the relation of denominator magnitude and position on the number line (e.g., unit fractions with larger denominators are closer to 0). During the game, students received detailed conceptual feedback providing additional opportunities to improve fraction understanding. For example, multiples of unit fractions were shown together with the monster on the line. For instance, when the monster's position was 6/10, the 0/10, 1/10, . . ., 10/10 all were shown on the line. The fraction number line training group showed stronger increases in fraction NLE proficiency and fraction magnitude comparison proficiency than the control group.

Braithwaite and Siegler (2021) extended this approach to sums of fractions. Participants saw strips with the length of all unit fractions between 0 and 1 (1/2, 1/3, . . ., 1/9) on the computer screen and were taught how to arrange sequences of these strips on the number line to find the position of a fraction (e.g., 3/7) or the position of a sum of fractions (e.g., 3/7 + 2/5). Subsequently, participants played the catch-the-monster game. In each trial, they concatenated several fraction unit strips to find the position of the monster on the line. Braithwaite and Siegler compared different versions of the intervention in two studies with 167 fourth-, fifth-, and sixth-graders. The interventions improved children's accuracy on a fraction NLE task and a fraction magnitude comparison task with individual fractions and sums of fractions.

Overall, previous studies on fraction number line interventions had encouraging results but also raised new questions. All previous studies except the very brief intervention by Hurst et al. (2020) used the fraction NLE task combined with explicit instruction on fraction concepts and procedures. This raises the question of whether the observed learning gains were due to the direct instruction, to the game itself, or to the combination of both.

The two previous studies evaluating the catch-the-monster game with fractions (Braithwaite & Siegler, 2021; Fazio et al., 2016) used explicit instruction and visualizations emphasizing the decomposition of fractions into series of unit fractions. This is a diversion from previous NLE studies, which emphasized the importance of integrated magnitude representations that combine information from the numerator and denominator into a single piece of information (i.e., a single position on the otherwise empty number line). This raises the question of whether the catch-the-monster game is effective only in the context of instruction emphasizing that fractions and sums of fractions are composed of unit fractions. Previous studies also did not investigate whether the beneficial effect of the catch-the-monster game with fractions transfers to symbolic fraction arithmetic problems (e.g., 2/7 + 3/5 = ?). We investigated these questions in the current study.

Scaffolding as a means to reduce cognitive load in game-based learning

Developers of serious games sometimes use scaffolding to improve learning (Chen & Law, 2016). These scaffolds can, for example, be small cues or prompts or a presentation order from easier problems to harder problems in the game. The need for such scaffolds arises from the fact that learning from a serious game without any direct non-game instruction can be considered a form of discovery learning. Nonguided discovery learning is relatively ineffective because of the high cognitive load associated with it (Kirschner, Sweller, & Clark, 2006). In line with this, meta-analytic findings show that serious games with additional scaffolds are more effective ($g = 0.48$) than games without scaffolds ($g = 0.26$) (Clark et al., 2016). Previous NLE interventions displayed hatch marks as scaffolding cues (e.g., Rittle-Johnson et al., 2001) or let students estimate sums of fractions on the number line to foster transfer to arithmetic (Braithwaite & Siegler, 2021). In the current study, we investigated to what extent these and further scaffolds improve the effectiveness of the catch-the-monster game.

The current study

In sum, fractions are an important but difficult to learn content of school mathematics. Previous intervention studies have shown that number lines combined with direct instruction on fraction concepts and procedures can improve fraction magnitude understanding. These studies left several questions open, and these are the research questions of the current study.

First, does the catch-the-monster game without any direct instruction on fraction concepts or strategies improve fraction magnitude understanding? We hypothesized that this was the case given that serious games tend to be effective because fraction NLE predicts fraction understanding and arithmetic and because previous fraction interventions with number lines found mostly positive effects. If the game really improves understanding, improvements should show in fraction NLE tasks and in other tasks assessing fraction magnitude understanding that were not trained in the intervention. We expected to find evidence for such near transfer because previous studies found near transfer from fraction NLE to alternative measures of fraction magnitude understanding (Fazio et al., 2016; Hamdan & Gunderson, 2017).

Second, do in-game scaffolds, such as cues and problem sequencing, increase the effectiveness of the game? We hypothesized that this was the case because such scaffolds turn free discovery learning into guided discovery learning, thereby reducing learners' cognitive load and leaving more working memory resources free for learning.

Third, do any beneficial effects of the catch-the-monster game transfer from NLE proficiency to fraction arithmetic? We hypothesized to find transfer from NLE to symbolic fraction arithmetic because NLE proficiency is an excellent predictor of fraction arithmetic competence (Schneider et al., 2018a) and because Braithwaite and Siegler (2021) found transfer from their version of the catch-the-monster game with direct instruction and fraction strips to a fraction sum comparison task.

Finally, is the effect of the catch-the-monster game moderated by grade level? Previous fraction NLE intervention studies mainly investigated students from Grades 4, 5, and 6. However, learning from serious games without additional non-game instruction is demanding in terms of working memory resources and learning strategies. Therefore, we hypothesized that students in Grades 7 and 8 would profit more from the game than students in Grades 5 and 6.

In the current study, we tested these hypotheses by letting fifth- to eighth-graders play the catch-the-monster game. We randomly assigned students to one of three experimental groups. A *fraction estimation group* used the standard version of the catch-the-monster game, in which they estimated the positions of single-digit random fractions on a number line. A *scaffolded estimation group* also played the game but received more scaffolding (i.e., tick marks on the line and trials ordered by difficulty). The third experimental group was an *active control group,* which played the catch-the-monster game with whole numbers. During this task, participants had no opportunity to learn something about fractions.

Method

Sample

We recruited students by contacting secondary schools in geographic proximity to our university, requesting them to promote the study among their fifth- to eighth-graders. A total of 188 German secondary school students volunteered to participate in this study in exchange for a small monetary compensation. In Germany, fraction concepts and procedures are mainly introduced in Grades 5 and 6. The students learn that fractions can be represented in various ways, including number lines, but instruction on number lines is usually brief (Bicker et al., 2003).

A sensitivity power analysis with G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for α = .05 and $1 - \beta$ = .80 showed that our sample size of N = 188 allowed us to detect effects with a strength of η^2 = .050 or higher for measures assessed only at posttest and effects with a strength of η^2 = .013 or higher for measures assessed at pretest and posttest. Thus, our design allowed us to detect even small effect sizes (cf. Cohen, 1992).

Of the 188 participants, 86 reported to be boys and 102 reported to be girls. The sample included 37 fifth-graders, 40 sixth-graders, 70 seventh-graders, and 41 eighth-graders. The mean age was 12.1 years (SD = 1.37). Table 1 provides an overview of the demographic characteristics of the students in the three experimental groups. There were no statistically significant differences among the three experimental groups regarding gender, $\chi^2(2, N = 188) = 3.82$, p = .148, grade, $\chi^2(6, N = 188) = 4.16$, p = .655, and age, $F(2, 185) = 2.00$, p = .138, η^2_G = .021. The study was approved by a local ethical committee.

Procedure

We used an experimental pretest–training–posttest design with three groups to investigate the effects of fraction number line training on fraction understanding and fraction arithmetic proficiency.

Table 1
Demographic characteristics for the three experimental groups.

Sample characteristic	Group		
	Fraction estimation	Scaffolded estimation	Active control
n	69	57	62
% Girls	60.9	43.9	56.5
% Boys	39.1	56.1	43.5
Mean age (*SD*) (years)	12.0 (1.3)	12.4 (1.3)	12.0 (1.5)
% in Fifth grade	17.4	17.5	24.2
% in Sixth grade	27.5	15.8	19.4
% in Seventh grade	33.3	42.1	37.1
% in Eighth grade	21.7	24.6	19.4

The computer program with the pretest, game, and posttest, the data, and the analysis syntax are available online (https://osf.io/cb8he/?view_only=15a5d5d3d97f4b3583f0813f51308ca3).

The study took place in a group setting at the last author's university's computer lab. On average, 17 students (*SD* = 7.28, range = 2–26) participated together. Study participation was voluntary. Parents provided informed consent for the study participation of their children. On arrival, students were randomly assigned to one of the three experimental groups (between-participant design). Students completed the pretest, participated in the intervention, and completed the posttest. The average training duration was 8.2 min (*SD* = 3.3).

Interventions in the experimental groups

Participants in the three experimental groups played different versions of a computerized number line game (*Catch the Monster With Fractions;* similar to Fazio et al., 2016; Rittle-Johnson et al., 2001; Schneider & Stern, 2010). Most aspects of the game were the same in all three conditions. In all three conditions, the number line game comprised three levels with 20 trials each (i.e., 60 trials in total per condition). In each trial, children were presented with a number line and a number. They were told that a monster was hiding at the position on the line indicated by the number. They could catch the monster by moving a monster basket to this position. After a student entered the answer, a monster appeared at the correct position of the number on the number line. In trials where a student indicated the position correctly, the monster basket turned green and the monster was "caught." Otherwise, the monster basket turned red and the monster escaped. The numbers of caught and escaped monsters across all trials were counted and displayed on the computer screen. The start point and end point of the number line were marked and labeled with numbers. The number line was otherwise empty except that unlabeled tick marks were shown in some trials, as explained below. At the end of each level of the game, a message window popped up and informed participants that they had completed the level and that a new part of the program was about to begin. The width of the monster basket decreased from one level to the next so that participants needed to make progressively more precise estimates over time to catch the monsters.

Table 2 shows the differences between the versions of the game played in the three experimental conditions. Participants in the fraction estimation group estimated the positions of single-digit fractions on a number line from 0 to 1. The fractions were presented in random order, but the same fractions were presented in each level of the game to aid students' hypothesis testing and memorization. The fractions were chosen to span the whole range of the number line. In half the trials, unlabeled tick marks were shown to make the relation between the denominator and the position on the line more salient. For example, when the presented fraction was 4/7, the tick marks were shown at the position of 1/7 and its multiples (2/7, . . ., 6/7).

Participants in the scaffolded estimation group also engaged in NLE but received scaffolds designed to make the relation among whole-number division, fraction magnitude estimation, and fraction arithmetic more salient. In the first level of the game, students first estimated the results of

Table 2
Differences between the interventions for the three experimental groups.

Game level	Fraction estimation group	Scaffolded fraction estimation group	Active control group
1	Single-digit fractions (e.g., 2/7); 0–1 range; random order; tick marks shown in half the trials	Whole-number divisions (e.g., 6/2) and single-digit fractions (e.g., 2/7); 0–3 range; trials ordered from easy to hard; tick marks shown in every trial	Whole numbers (e.g., 475); 0–1000 range; random order; tick marks shown in half the trials
2	Single-digit fractions; 0–1 range; random order; tick marks shown in half the trials	Single-digit fractions; 0–1 range; random order; tick marks shown in half the trials	Whole numbers; 0–1000 range; random order; tick marks shown in half the trials
3	Single-digit fractions; 0–1 range; random order; tick marks shown in half the trials	Sums of single-digit fractions (e.g., 1/4 + 2/7); 0–1 range; random order; tick marks shown in half the trials	Whole numbers; 0–1000 range; random order; tick marks shown in half the trials

whole-number divisions (e.g., 6 divided by 2) on number lines from 0 to 3 in 5 trials. Then, students estimated 1/2 and its multiples for 5 trials (i.e., 1/2, 2/2, 3/2, 4/2, 5/2). Next, students estimated the positions of multiples of 1/3 (i.e., 0/3, 1/3, 3/3, 5/3, 6/3). Finally, students indicated the positions of five fractions with numerator 3 and increasing denominators (i.e., 3/1, 3/2, 3/3, 3/6, 3/9). We expected that varying the numerator only (in Trials 6–15) or varying the denominator only (in Trials 16–20) while holding the other constant would help students to better understand how they relate to the overall fraction magnitude. In Level 1, unlabeled tick marks were shown in each trial. Level 2 for the scaffolded fraction estimation group was the same as Level 2 for the estimation-only group. In Level 3, the scaffolded fraction estimation group indicated the position of a sum of two fractions on the number line (e.g., "What is the position of 3/5 + 2/5 on the line?") in each trial. We included fraction pairs with equal denominators (5 trials) and fraction pairs with unequal denominators (15 trials). Fraction sums were included to foster the integration of knowledge about fraction magnitudes and about arithmetic procedures. We expected this to facilitate transfer from the NLE training to fraction arithmetic measured at posttest.

In the active control group, students indicated the location of whole numbers on a number line ranging from 0 to 1000 in all three levels of the game. We determined these whole numbers by multiplying the fractions used in the fraction estimation group by 1000 and rounding them to whole numbers (e.g., $2/7 \times 1000 \approx 286$).

Measures

Students completed a fraction NLE task (ranging from 0 to 1) and a fraction addition task immediately before the intervention (pretest). After the intervention, students completed the same tasks as in the pretest and also completed an NLE task (ranging from 0 to 5), a fraction subtraction task, and a fraction comparison task. We presented all measures in a randomized order at pretest. At posttest, we presented the tasks ordered by how similar they were to the intervention, that is, in the following order: NLE with fractions (0–1), NLE with fractions (0–5), fraction comparison task, fraction addition task, and fraction subtraction task. We used this order to avoid contaminating test effects from the arithmetic test tasks to the NLE tasks.

The NLE with fractions (0–1), which was presented at pretest and posttest, contained the same fractions that the fraction estimation group trained in three levels and that the scaffolded fraction estimation group trained in Level 2. Thus, this task served as an implementation check. The NLE tasks with fractions (0–5) and the fraction comparison task served as indicators of fraction understanding. The fraction addition task and the fraction subtraction task served as indicators of transfer to fraction arithmetic proficiency.

For NLE tasks, we quantified task performance as the percent of absolute error (PAE = $100 \times$ |correct number–estimated number| / numerical range of the number line). For all other tasks, we quantified task performance as the percentage of correctly solved trials.

NLE with fractions (0–1)

Students indicated the position of 10 single-digit fractions on a number line ranging from 0 to 1 (i.e., 0/8, 1/8, 2/7, 1/3, 3/7, 2/4, 3/5, 7/9, 4/5, 2/2). Only the start point and end point of the lines were marked and labeled. All the presented single-digit fractions were also presented in the single-digit fractions training blocks.

NLE with fractions (0–5)

Students indicated the position of 10 fractions on an empty number line ranging from 0 to 5 (i.e., 3/9, 12/15, 5/3, 18/10, 9/4, 8/3, 17/5, 4/1, 9/2, 14/3). This task was a transfer task in that none of these fractions was presented in the number line training and none of the number lines in the training had a range from 0 to 5. Thus, students could not solve this task by memorizing the positions of the fractions in the number line trainings but instead needed to learn the underlying principle (i.e., how the numerator and denominator together determine the numerical magnitude of a fraction) and transfer it to new problems.

Fraction comparison task

Students were presented with eight single-digit fractions (i.e., 2/9, 3/8, 4/9, 4/7, 5/8, 3/4, 5/6, 8/9). For each fraction, participants decided whether it was larger or smaller than 3/5. Half the presented fractions were smaller and half were larger than 3/5. All fractions were presented four times. The fraction comparison task is an indicator of fraction understanding that is related to fraction NLE (e.g., Fazio et al. 2016) but is conceptually distinct from fraction NLE (Schneider, Thompson, & Rittle-Johnson, 2018b).

Fraction addition task

Students solved six addition problems with single-digit fractions (i.e., 4/8 + 1/8, 2/5 + 2/5, 1/2 + 2/5, 1/4 + 1/3, 2/7 + 2/3, 2/3 + 1/6). In two of these six addition problems, the denominators of the addends were equal.

Fraction subtraction task

Analogous to the fraction addition task, students solved six subtraction tasks with single-digit fractions (i.e., 3/5 – 2/5, 5/7 – 2/7, 1/2 – 2/7, 2/3 –1/4, 1/7 – 1/8, 3/4 – 1/8). Again, in two of the six subtraction problems, the denominators of the two fractions were equal.

Statistical analyses

For each task completed at pretest and posttest, we conducted a mixed analysis of variance (ANOVA) with experimental group (fraction estimation, scaffolded fraction estimation, or active control) and grade (fifth, sixth, seventh, or eighth grade) as between factors, time (pretest or posttest) as a within factor, and the performance in the respective task as the dependent variable. A significant interaction effect of experimental group and time indicates that pretest–posttest changes in fraction competence differed between the experimental groups. In addition, we computed Bayes factors indicating the probability of our data under the alternative hypothesis (H1; i.e., there is an interaction effect of group and time) or the null hypothesis (H0; i.e., there is no interaction effect of group and time) (Dienes, 2014).

For each task completed only at posttest, we conducted an ANOVA with experimental group (fraction estimation, scaffolded fraction estimation, or active control) and grade (fifth, sixth, seventh, or eighth grade) as between factors and performance in the respective task as the dependent variable. In addition, we computed Bayes factors indicating the probability of our data under the H1 (i.e., the groups differ at posttest) or the H0 (i.e., the groups do not differ at posttest).

In case of significant main effects, we conducted Bonferroni-corrected pairwise post hoc tests to determine which groups differed in their means. As standardized effect sizes, we report the generalized eta-squared index η^2_G (Bakeman, 2005). We conducted the analyses in R (R Core Team, 2020). We conducted the ANOVAs with the package *ez* (Lawrence, 2016), the post hoc tests with *emmeans* (Lenth et al., 2021), and the Bayesian analyses with *BayesFactor* (Morey & Rouder, 2018).

Results

Preliminary analyses

Table 3 displays the means and standard deviations for the two measures that had been assessed at pretest as well as at posttest: NLE on the 0–1 line and fraction addition. The three experimental groups did not differ in their performance at pretest, neither in NLE, $F(2, 185) = 0.08$, $p = .923$, $\eta^2_G = .001$, nor in fraction addition, $F(2, 185) = 0.26$, $p = .774$, $\eta^2_G = .003$. The means and standard deviations of the dependent variables measured only at posttest can be found in the online supplementary material (Table SM1).

The bivariate correlations and Cronbach's α for all tasks are listed in the supplemental material (Tables SM2–SM5). The intercorrelations of all tasks were strong (for all *r*s: $.45 < |r| < .92$). For estimation and addition, pretest performance positively correlated with posttest performance (all *r*s $\geq .80$).

Table 3
Numbers of participants, means, and standard deviations for pretest and posttest by measure, experimental group, and grade.

	Total sample			Grade 5			Grade 6			Grade 7			Grade 8		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Number line estimation task (0–1)															
Total sample															
Pretest	188	12.66	11.70	37	26.16	11.14	40	12.12	11.54	70	9.06	9.06	41	7.15	5.82
Posttest	188	9.77	10.13	37	21.11	10.34	40	7.52	8.05	70	7.15	9.26	41	6.21	4.89
Fraction estimation group															
Pretest	69	12.38	11.27	12	29.19	9.94	19	11.91	9.96	23	7.89	6.99	15	6.42	4.62
Posttest	69	7.25	8.14	12	17.40	8.71	19	5.37	5.13	23	5.50	8.58	15	4.17	2.05
Scaffolded fraction estimation group															
Pretest	57	12.47	12.15	10	24.24	11.63	9	15.84	17.96	24	9.50	9.54	14	6.99	4.49
Posttest	57	9.72	10.62	10	23.86	11.19	9	8.80	10.44	24	6.77	8.46	14	5.26	3.70
Active control group															
Pretest	62	13.15	11.91	15	25.01	11.95	12	9.67	7.57	23	9.78	10.54	12	8.26	8.36
Posttest	62	12.63	11.04	15	22.23	10.78	12	9.98	9.55	23	9.18	10.63	12	9.87	6.65
Addition task															
Total sample															
Pretest	188	59.57	42.36	37	0.00	0.00	40	71.25	34.59	70	72.14	36.20	41	80.49	28.60
Posttest	188	59.75	41.68	37	2.25	8.01	40	67.92	32.33	70	73.10	36.37	41	80.89	29.48
Fraction estimation group															
Pretest	69	59.90	41.35	12	0.00	0.00	19	68.42	33.75	23	73.91	35.11	15	75.56	33.85
Posttest	69	61.11	40.59	12	2.78	9.62	19	65.79	31.17	23	77.54	34.31	15	76.67	34.39
Scaffolded fraction estimation group															
Pretest	57	62.28	40.40	10	0.00	0.00	9	53.70	41.48	24	77.78	25.38	14	85.71	27.62
Posttest	57	63.74	40.60	10	1.67	5.27	9	53.70	37.99	24	79.86	27.79	14	86.90	27.09
Active control group															
Pretest	62	56.72	45.61	15	0.00	0.00	12	88.89	22.84	23	64.49	45.87	12	80.56	23.39
Posttest	62	54.57	43.96	15	2.22	8.61	12	81.94	26.07	23	61.59	44.22	12	79.17	26.71

Note. We quantified performance in the number line estimation tasks as percent absolute error (PAE), with lower scores indicating better performance, and as percent of correctly solved trials, with higher scores indicating better performance in all other tasks.

Tables SM3, SM4, and SM5 display the bivariate correlations and Cronbach's α for all tasks separately for the three experimental groups. Cronbach's α was acceptable given the heterogeneity of provided problems in each task (median α = .82, range = .62–.94).

Main analyses

Implementation check

Table 4 displays the results of all ANOVAs. As an implementation check, we tested whether the training groups differed in their changes in NLE proficiency from pretest to posttest. We computed an ANOVA with experimental group and grade as between factors, time point as a within factor, and performance in the fraction NLE task (0–1) as the dependent variable. The results revealed a main effect of grade, a main effect of time, an interaction effect of group and time, an interaction effect of grade and time, and a three-way interaction of group, grade, and time. The three-way interaction is visualized by box plots in Fig. 1. The *y* axis represents improvements in PAE, computed as $PAE_{pretest} - PAE_{posttest}$. Positive values on the *y* axis indicate learning gains because they imply that the error was smaller at posttest than it was at pretest. As shown in Fig. 1, the average learning gains were much larger for fifth-graders in the estimation-only group than for all other grades and training groups, whereas seventh- and eighth-graders showed virtually no improvement in NLE proficiency. ANOVAs separately for each grade showed a significant interaction of group and time for Grade 5 (p = .009, η^2_G = .051) and Grade 6 (p = .047, η^2_G = .027), but not for Grade 7 (p = .293, η^2_G = .003) or Grade 8 (p = .114, η^2_G = .027). As shown in Fig. 1, the interaction effect in Grade 5 mainly stems from the large learning gains of the fraction estimation group compared with the other two groups. In Grade 6, both estimation groups had larger learning gains than the control group. Overall, these findings indicate that the intervention was only partly successful. The aspect of fraction understanding that is assessed by NLE on 0–1 lines improved selectively only in some treatment groups and grade levels.

As indicated by the main effect of grade, and as shown by the mean values in Table 4, fraction NLE (0–1) PAE at pretest decreased plausibly with increasing age. Bonferroni-corrected pairwise post hoc tests indicated differences between Grade 5 and Grade 6, $t(176) = 6.058$, $p < .001$, η^2_G = .282, between Grade 5 and Grade 7, $t(176) = 8.675$, $p < .001$, η^2_G = .411, and between Grade 5 and Grade 8, $t(176) = 8.608$, $p < .001$, η^2_G = .547.

NLE task (0–5)

An ANOVA with experimental group and grade as between factors and performance on the NLE task (0–5) at posttest as the dependent variable revealed no significant main effect of the experimental group (see Table 4). The main effect of grade level was statistically significant. Bonferroni-corrected pairwise post hoc tests revealed that fifth-graders had a higher PAE than sixth-graders, $t(176) = 5.306$, $p < .001$, η^2_G = .348, seventh-graders, $t(176) = 8.610$, $p < .001$, η^2_G = .457, and eighth-graders, $t(176) = 7.262$, $p < .001$, η^2_G = .495. All other comparisons between grades did not reach statistical significance. The interaction effect between grade and experimental group was not statistically significant. The Bayes factor for group differences at posttest was BF_{01} = 14.0, indicating that our data were 14.0 times more likely under the H0 than under the H1. According to the classification by Jarosz and Wiley (2014), this is strong evidence for the H0 that there was no transfer from the NLE training to fraction NLE on the 0–5 line.

Fraction comparison task

An ANOVA with experimental group and grade as between factors and performance on the fraction comparison task as the dependent variable revealed no significant main effect of the experimental group (see Table 4). The main effect of grade level reached statistical significance. The interaction did not reach statistical significance. For grade level, Bonferroni-corrected pairwise post hoc tests revealed that fifth-graders performed significantly worse than seventh-graders, $t(176) = -4.010$, $p = .001$, η^2_G = .184, and eighth-graders, $t(176) = -4.500$, $p < .001$, η^2_G = .256. The Bayes factor BF_{01} for group differences at posttest was 10.4, which provided strong support in favor of the H0 that there was no transfer from the NLE training to fraction comparison.

Table 4
Results from the analyses of variance with experimental group and grade as between factors, time point as a within factor, and performance in all measures completed as the dependent variables.

		Number line estimation (0–1)			Number line estimation (0–5)			Fraction comparison			Fraction addition			Fraction subtraction		
Effect	$df_{between}$	*F*	*p*	η^2_G	*F*	*p*	η^2_G	*F*	*p*	η^2_G	*F*	*p*	η^2_G	*F*	*p*	η^2_G
Group	2	1.01	.367	.010	0.08	.924	.000	0.44	.642	.005	0.13	.880	.001	0.19	.827	.002
Grade	3	33.28	<.001	.330	27.15	<.001	.316	7.80	<.001	.117	62.53	<.001	.494	48.52	<.001	.453
Time	1	36.47	<.001	.026	–	–	–	–	–	–	0.00	.960	.000	–	–	–
Group × Time	2	10.41	<.001	.015	–	–	–	–	–	–	0.77	.467	.001	–	–	–
Group × Grade	6	0.28	.948	.008	0.78	.584	.026	0.09	.997	.003	1.85	.092	.055	0.65	.688	.022
Grade × Time	3	3.81	.011	.008	–	–	–	–	–	–	0.66	.575	.001	–	–	–
Group × Grade × Time	6	2.81	.012	.012	–	–	–	–	–	–	0.16	.987	.001	–	–	–

Note. df_{within} = 176 for all analyses.

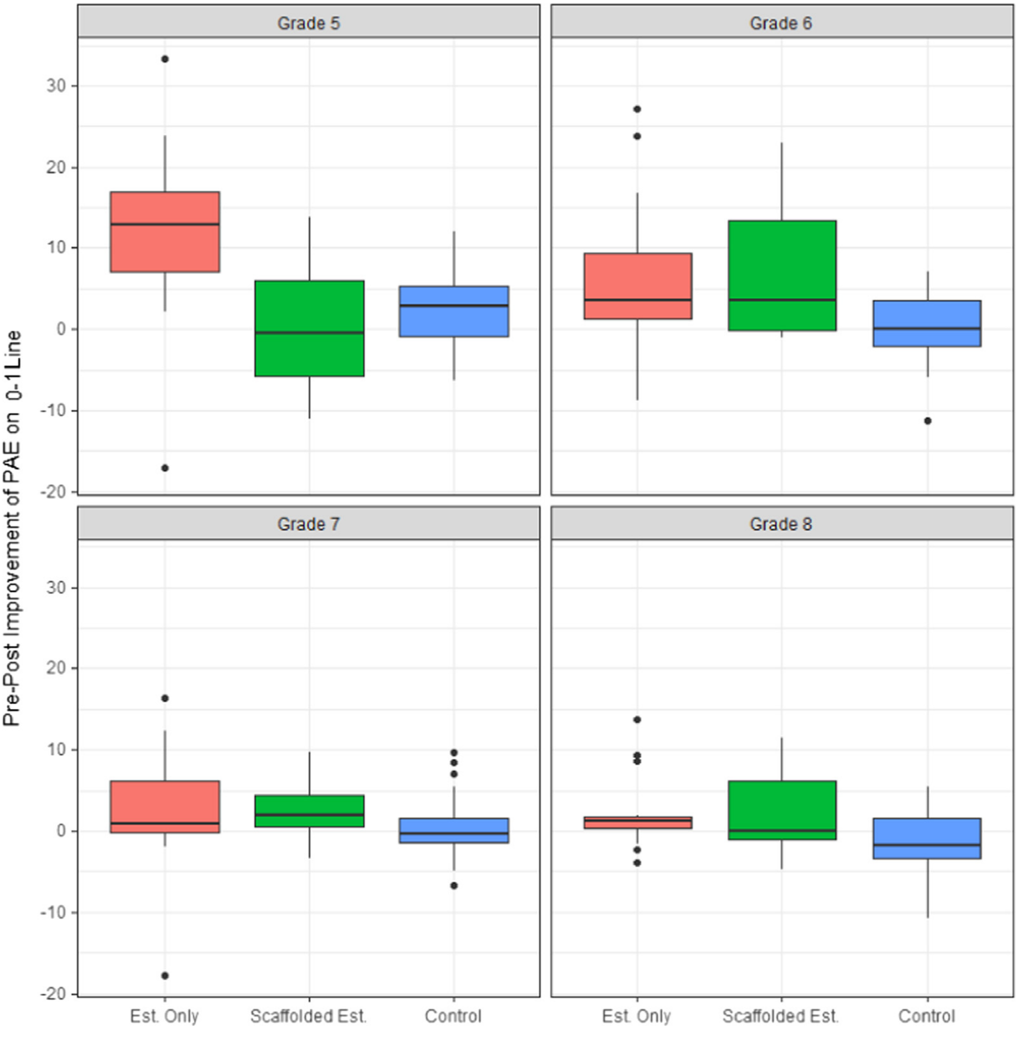

Fig. 1. Pretest–posttest improvements in number line estimation on the 0–1 line (computed as $PAE_{pretest} - PAE_{posttest}$, where PAE is percent absolute error, so that positive differences indicate decreasing error). Est. = estimation.

Fraction addition task

As shown in Table 4, an ANOVA with experimental group and grade as between factors, time point as a within factor, and performance in the fraction addition task as the dependent variable revealed neither a significant main effect of time point nor a significant interaction effect of experimental group and time point. Thus, the performance of students in all three experimental groups did not increase from pretest to posttest. The main effect of grade reached statistical significance. At pretest, fifth-graders showed overall a floor effect (M = 0.00, SD = 0.00). Thus, follow-up t tests for the pretest revealed that fifth-graders had a lower performance than sixth-graders, $t(176) = -9.920$, $p < .001$, $\eta_G^2 = .677$, seventh-graders, $t(176) = -11.645$, $p < .001$, $\eta_G^2 = .582$, and eighth-graders, $t(176) = -11.674$, $p < .001$, $\eta_G^2 = .794$. All other comparisons between grades did not reach statistical significance. The Bayes factor for group differences at posttest was $BF_{01} = 9.4$, which provides substantial support (cf. Jarosz & Wiley, 2014) for the H0 that there was no transfer to fraction addition.

Fraction subtraction task

An ANOVA with experimental group and grade as between factors and performance on the fraction subtraction task as the dependent variable revealed no significant differences between the three experimental groups and no Group × Grade interaction. The main effect of grade level was statistically significant. Bonferroni-corrected pairwise post hoc tests revealed that the performance of fifth-graders was significantly lower than the performance of sixth-graders, $t(176) = -7.832$, $p < .001$, $\eta^2_G = .620$, seventh-graders, $t(176) = -10.883$, $p < .001$, $\eta^2_G = .552$, and eighth-graders, $t(176) = -10.614$, $p < .001$, $\eta^2_G = .717$. All other comparisons between grades did not reach statistical significance. The Bayes factor BF_{01} for group differences at posttest was 10.5, indicating that our data are 10.5 times more likely under the H0 that there was no transfer to fraction subtraction than under the H1 that there was transfer.

Discussion

Main findings

This study investigated whether the catch-the-monster game with fractions but without direct instruction increases fraction understanding and fraction arithmetic proficiency in a randomized controlled experiment with a large sample of fifth- to eighth-graders. In the game, students practiced estimating the position of fractions on a number line. As expected, the two fraction estimation groups showed larger improvements in fraction NLE proficiency than the control group. Thus, the implementation check was successful. This shows that students were motivated to learn and could learn something from the game in principle. Students on all grade levels showed improvements, but the effects reached statistical significance only for fifth- and sixth-graders because seventh- and eighth-graders already performed well at pretest. This shows that fraction number line games can be more effective for younger children than for older children. The large learning gains of the youngest age group in our study raise the question of whether the game would have been even more effective in younger children, for example, third- and fourth-graders. In line with this, previous studies with second- to fifth-graders had encouraging results (e.g., Fazio et al., 2016; Fuchs et al., 2013; Hamdan & Gunderson, 2017).

Our first research question was whether the game improved not only fraction NLE proficiency but also fraction magnitude understanding in general. If this were the case, students in the fraction estimation groups should outperform students in the control group on tasks that assess magnitude understanding but were not part of the intervention. This near transfer did not occur. The fraction monster game did not improve fraction estimation proficiency on 0–5 lines and did not improve the solution rate on a fraction magnitude comparison task. The fact that the fraction NLE training improved one measure of fraction understanding (i.e., fraction estimation on 0–1 lines), but not two other measures of fraction understanding (i.e., fraction estimation on 0–5 lines and fraction magnitude comparison), shows that fraction magnitude understanding is not a homogeneous construct and cannot be fully equated with NLE proficiency. This finding is in line with studies showing that the NLE task is no pure measure of mental magnitude representations but also reflects estimation strategies, proportional reasoning, and possibly other competencies (cf. Schneider et al., 2018a).

Our second research question was whether in-game scaffolds increase the effectiveness of the game. This was not the case. For Grade 5, students playing the fraction game without scaffolds had much higher learning gains than students playing the fraction game with scaffolds. For Grade 6, students learning gains were equally large in both conditions. This suggests that the scaffolds, which included the estimation of fraction sums on the line, were too difficult for German fifth-graders and did not help older students.

Our third research question was whether there is a transfer from improved fraction magnitude understanding to symbolic fraction arithmetic. Because students' fraction magnitude understanding, as assessed by NLE on 0–5 lines and by magnitude comparison, did not improve, we could not investigate this question. Significance tests showed that the three experimental groups did not differ in their solution rates on fraction addition and fraction subtraction tasks.

Our fourth research question concerned age differences in learning from the game. We found strong and plausible differences between students in different grades. With increasing grade level, students were able to estimate fractions more accurately and to solve more fraction comparison and fraction arithmetic problems. Against our expectations, younger students (i.e., fifth- and sixth-graders) learned more from the game than older students.

Why was there no transfer effect?

The game improved NLE proficiency on 0–1 lines in fifth- and sixth-graders, but these improvements did not transfer to NLE in 0–5 lines, magnitude comparison, or arithmetic with fractions. Several strengths of the study make it unlikely that this lack of transfer is a mere methodological artifact. We used an experimental design with random assignment of students to the experimental groups and an active control group. The sample size was larger than in some previous studies investigating fraction number line training (Fazio et al., 2016; Hamdan & Gunderson, 2017). The statistically significant implementation check and age group differences indicate that the measures had sufficient reliability for effects to reach statistical significance. Accordingly, power sensitivity analyses showed that our sample size was large enough to detect even weak effects. In line with this, the effect sizes for differences between the three experimental conditions were generally so small that these differences were irrelevant from a practical point of view. To assess the intervention outcome, we applied a broad set of well-established measures of fraction understanding and fraction arithmetic. Some of these measures assessed near transfer (e.g., from estimating fractions on 0–1 lines to estimating fractions on 0–5 lines), and some of them assessed transfer over a greater transfer distance (e.g., from estimating fractions on 0–1 lines to fraction subtraction). The generally high and plausible intercorrelations of our measures show that we did not simply measure random noise. They indicate that we captured a high proportion of systematic variance, which demonstrates the reliability of our measures.

As an intervention, we applied two versions of a previously successfully used number line game (Fazio et al., 2016; Rittle-Johnson et al., 2001; Schneider & Stern, 2010). The intervention in the fraction estimation group mirrored prior intervention studies, and the intervention in the scaffolded fraction estimation group gave additional cues and guidance designed to make the task easier, reduce working memory load, and increase transfer. We found significant improvements in fraction NLE proficiency for previously trained single-digit fractions from pretest to posttest in both fraction estimation groups but not in the active control group. This implementation check demonstrates that participating students complied with the instructions and completed the training thoroughly. Neither the training performance nor the test performance indicated any lack of student motivation. The intervention was brief (8 min on average), but with 60 trials the treatment intensity was higher than in previous number line interventions with positive effects. For example, only 14 trials of the catch-the-monster game with decimal fractions led to significant pretest–posttest improvements in the study by Rittle-Johnson et al. (2001). A median number of 30 trials led to significant improvements in the catch-the-monster with fractions study by Fazio et al. (2016). Opfer and Siegler (2007) found that oftentimes a single trial of NLE of feedback can cause stable improvements of second-graders' magnitude understanding of whole numbers. These changes transferred from the trained number to all numbers on the number line. The games played by the fraction estimation group and the enhanced fraction estimation group differed in several characteristics (see Table 2). The fact that our results were the same for both groups demonstrates the robustness of our findings to small choices in designing the game.

The core difference between the fraction intervention in this study and fraction interventions in previous studies is that the number line training was not accompanied by any further direct instruction on fraction concepts or fraction procedures. In our study, fraction NLE was not embedded in an extensive fraction learning intervention over multiple weeks (Fuchs et al., 2013, 2014; Schumacher et al., 2018), and there were no conceptual introductions to fractions or extensive feedback highlighting relations among numerator, denominator, and positions on the number line (Fazio et al., 2016; Hamdan & Gunderson, 2017). Instructions on how to divide the number line into segments (e.g., Braithwaite & Siegler, 2021; Fazio et al., 2016) might, independent of fraction number line training, improve fraction understanding because these instructions integrate common part-of-whole explanations of fractions with the visual representation of a fraction magnitude on a number line.

The most plausible explanation for the lack of any transfer effects despite the rigorous methods used in this study is that isolated fraction number line training does not improve fraction understanding and fraction arithmetic proficiency. Further instructional elements seem to be a prerequisite for fraction number line training to show transfer effects.

In number line training with whole numbers, additional instructions do not seem to be required to improve numerical knowledge and whole number arithmetic proficiency (e.g., Opfer & Siegler, 2007; Siegler & Ramani, 2009). This discrepancy to effects of fraction number line trainings may be explained by the higher complexity of fractions (particularly fraction arithmetic) in comparison with whole numbers. Whole-number magnitude knowledge and whole-number arithmetic proficiency are closely related (e.g., Booth & Siegler, 2008; Siegler & Ramani, 2009). Comparatively, the relation between fraction arithmetic proficiency and fraction magnitude knowledge is weaker (Braithwaite, Tian, & Siegler, 2018). According to the FARRA (fraction arithmetic reflects rules and associations) model of fraction arithmetic (Braithwaite, Pyke, & Siegler, 2017), the development of fraction arithmetic proficiency is relatively independent of fraction understanding. Thus, even if isolated fraction number line training had positive effects on fraction understanding, without additional instructional elements, transfer effects on fraction arithmetic proficiency do not necessarily occur. To achieve increases in fraction arithmetic proficiency, explicit instructions on the principles of fraction arithmetic seem to be required (Braithwaite et al., 2018).

Limitations and future research directions

Notwithstanding its methodical rigor, this study has several limitations. First, even though our treatment intensity was high compared with previous successful number line interventions, we cannot rule out that an even higher number of training trials would have caused stronger and broader effects in our study.

Second, different tasks in the fraction number line training might have produced stronger effects. Although we combined different fraction number line training variations in the scaffolded fraction estimation group, other approaches may be more promising. Using a high number of easy problems (i.e., with simple fractions such as 1/4) might leave students more cognitive load for inferring general properties of fractions. Perhaps further gamification through linear number board games (cf. Siegler & Ramani, 2009) could further increase students' engagement in the training.

Third, students in this study were 10 to 15 years old. Although no transfer effects occurred in this age group, isolated fraction number line training may be effective in other age groups, for example, third- and fourth-graders. Future research is needed to test which forms of fraction number line training and accompanying instructions are effective in different age groups. Some cell counts of the Condition $\times$ Grade (see Table 3) were quite low. Thus, our exploratory results regarding how grade level moderated the intervention effects need to be interpreted with caution.

Theoretical and practical implications

This study showed that, unlike the fraction number line with additional instructional elements (e.g., Fazio et al., 2016), fraction number line training without additional conceptual instructions does not lead to transfer effects on fraction understanding and fraction arithmetic proficiency. This finding is in line with extensive evidence demonstrating the superiority of guided instruction compared with minimal instruction on various learning outcomes (e.g., Kirschner et al., 2006) and with meta-analytic evidence indicating the effectiveness of direct instruction for learning ($d = 0.59$; Hattie, 2008) that may explain the observed effects in previous studies. In our study, students mapped each fraction onto a point on a continuous number line. This proved to be ineffective. By contrast, Braithwaite and Siegler (2021) used a quite different and much more effective version of the catch-the-monster game. They let students visualize fractions and sums of fractions as a series of unit fractions on the number line. Our findings imply that this putting fractions together framework is a promising approach for the development of effective fraction interventions. As other fraction interventions have shown, number lines are also highly useful tools for visualizing and explaining fractions in direct instruction (Gunderson et al., 2019; Hamdan & Gunderson, 2017; Tian et al., 2021). In particular, number lines are useful for introducing relevant concepts for fraction understanding, such as fraction equivalence (e.g., 3/4 is equal to 6/8 because both fractions share the same location on a number line; Jordan, Rodrigues, Hansen, & Resnick, 2017). However, fraction number line trainings are not a panacea. For improving fraction understanding and fraction arithmetic proficiency, fraction number line training is effective only when it is accompanied by additional instructions or when it is embedded in an elaborated intervention program. Thus, fraction number line trainings may enrich but not replace other instructional practices.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2021.105353.

References

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113*, 447–455.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*, 379–384.

Bicker, U., Ding, K., Euteneuer, A., Hein, F., Jacoby, J., Martin, K., . . . Vicary, R., (2003). *Rahmenlehrplan Mathematik (Klassenstufen 5–9/10).* Mainz, Germany: Ministerium für Bildung, Wissenschaft, Jugend und Kultur Rheinland Pfalz. Retrieved from https://www.edoweb-rlp.de/resource/edoweb%3A2686122.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*, 1016–1031.

Braithwaite, D. W., Pyke, A. A., & Siegler, R. S. (2017). A computational model of fraction arithmetic. *Psychological Review, 124*, 603–625.

Braithwaite, D. W., & Siegler, R. S. (2021). Putting fractions together. *Journal of Educational Psychology, 113*, 556–571.

Braithwaite, D. W., Tian, J., & Siegler, R. S. (2018). Do children understand fraction addition? *Developmental Science, 21*, e12601.

Chen, C. H., & Law, V. (2016). Scaffolding individual and collaborative game-based learning in learning performance and intrinsic motivation. *Computers in Human Behavior, 55*(Pt B), 1201–1212.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*, 79–122.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00781.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1–16.

Fazio, L. K., Kennedy, C. A., & Siegler, R. S. (2016). Improving children's knowledge of fraction magnitudes. *PLoS One, 11*(10), e165243.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., ... Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*, 683–700.

Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., ... Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude–treatment interaction. *Journal of Educational Psychology, 106*, 499–514.

Gunderson, E. A., Hamdan, N., Hildebrand, L., & Bartek, V. (2019). Number line unidimensionality is a critical feature for promoting fraction magnitude concepts. *Journal of Experimental Child Psychology, 187*, 104657.

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial–numerical representation: Evidence from a fraction intervention. *Developmental Psychology, 53*, 587–596.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to*. Achievement, London: Routledge.

Hurst, M. A., Massaro, M., & Cordes, S. (2020). Fraction magnitude: Mapping between symbolic and spatial representations of proportion. *Journal of Numerical Cognition, 6*, 204–230.

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *Journal of Problem Solving, 7*(1), 2. https://doi.org/10.7771/1932-6246.1167.

Jordan, N. C., Rodrigues, J., Hansen, N., & Resnick, I. (2017). Fraction development in children: Importance of building numerical magnitude understanding. In D. C. Geary, D. B. Berch, R. J. Ochsendorf, & K. M. Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 125–140). Cambridge, MA: Academic Press.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Lawrence, M. A. (2016). *ez: Easy analysis and visualization of factorial experiments.* R package Version 4.4-0. URL: https://CRAN.R-project.org/package=ez.

Lenth, R. V., Buerkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2021). *emmeans: Estimated marginal means, aka least-squares means.* R package Version 1.6.3. URL: https://CRAN.R-project.org/package=emmeans.

Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review, 38*, 201–221.

Martin, W. G., Strutchens, M. E., & Elliott, P. C. (2007). *The learning of mathematics* (Vol. 69) Reston, VA: National Council of Teachers of Mathematics.

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes factors for common designs.* R package Version 0.9.12-4.2. URL: https://CRAN.R-project.org/package=BayesFactor.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195.

Peeters, D., Degrande, T., Ebersbach, M., Verschaffel, L., & Luwel, K. (2016). Children's use of number line estimation strategies. *European Journal of Psychology of Education, 31*, 117–134.

R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.R-project.org.

Ramani, G. B., & Siegler, R. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*, 375–394.

Ramani, G. B., Siegler, R. S., & Hitti, A. (2012). Taking it to the classroom: Number board games as a small group learning activity. *Journal of Educational Psychology, 104*, 661–672.

Rinne, L. F., Ye, A., & Jordan, N. C. (2017). Development of fraction comparison strategies: A latent transition analysis. *Developmental Psychology, 53*, 713–730.

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*, 346–362.

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018a). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development, 89*, 1467–1484.

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1227–1238.

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology, 46*, 178–192.

Schneider, M., Thompson, C. A., & Rittle-Johnson, B. (2018b). Associations of magnitude comparison and number line estimation with mathematical competence: A comparative review. In P. Lemaire (Ed.), *Cognitive development from a strategy perspective: A Festschrift for Robert S. Siegler* (pp. 100–119). Milton Park, Oxfordshire, UK: Routledge/Taylor & Francis Group.

Schumacher, R. F., Jayanthi, M., Gersten, R., Dimino, J., Spallone, S., & Haymond, K. S. (2018). Using the number line to promote understanding of fractions for struggling fifth graders: A formative pilot study. *Learning Disabilities Research & Practice, 33*, 192–206.

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science, 19*, 341–361.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*, 691–697.

Siegler, R. S., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives, 8*, 144–150.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, 11*, 655–661.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*, 545–560.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*, 273–296.

Tian, J., Bartek, V., Rahman, M. Z., & Gunderson, E. A. (2021). Learning improper fractions with the number line and the area model. *Journal of Cognition and Development, 22*, 305–327.

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.

Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction, 28*, 181–209.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*, 249–265.