Learning and Individual Differences 119 (2025) 102643

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Does multiplication *always* make bigger? Exploring individual differences in NanoRoboMath game play☆

Tomi Kärki [a,*], Hilma Halme [a], Minna Hannula-Sormunen [a], Erno Lehtinen [a,b], Jake McMullen [a]

[a] *Department of Teacher Education, University of Turku, Finland*
[b] *Education Academy, Vytaytas Magnus University, Lithuania*

ARTICLE INFO

Keywords:
Rational numbers
Game-based learning
Conceptual knowledge
Natural number bias

ABSTRACT

Many students tend to inappropriately apply natural-number-biased reasoning in fraction and decimal tasks, including believing multiplication always makes bigger and division makes smaller. In this study, we examined individual differences in the game play of NanoRoboMath, a digital game designed to improve students' rational number knowledge. Examining the game performance of 90 seventh grade students allowed us to find four player profiles that were connected to learning the effects of multiplicative operations. Students in the High profile used multiplication and division by numbers less than one more frequently and had stronger learning gains with operation knowledge results compared to Long play low gain profile. This suggests that the player profiles reflect some features of students' game play and performance that may be relevant for conceptual change processes of understanding the effects of multiplicative operations.

Educational relevance and implications statement: The present study provides novel insights into individual differences in students' performance while playing a game aimed at promoting conceptual knowledge of rational numbers. It manifests a relation between player profiles and learning gains. The study contributes to the design of game-based learning environments by suggesting that providing students with repeated opportunities to confront their prior misconceptions in game-based learning environments may be beneficial for supporting conceptual development. The study also indicates that there may be multiple ways to productively engage with educational games. This is valuable for educators in understanding how to employ game-based learning environments in their classrooms.

1. Introduction

Rational number learning is a perennial thorn in the side of students, teachers, educators, and parents worldwide (Kieren, 1976; Ni & Zhou, 2005; Van Dooren et al., 2015). Not only are rational numbers difficult to learn, they also are crucial for learning more advanced topics and for future success in mathematics, for example, in middle school algebra and in high school mathematics (Booth & Newton, 2012; Siegler et al., 2012). Many features of rational numbers do not follow the same logic as natural numbers. Yet, after years of instruction and experiences reasoning with natural numbers, some students (as well as adults) misapply natural number logic when dealing with rational numbers (Kainulainen et al., 2017; Obersteiner et al., 2016; Stafylidou & Vosniadou, 2004; Vamvakoussi & Vosniadou, 2004). Overcoming such a natural number bias may require conceptual change, and thus, be onerous, time-consuming, precarious, and difficult (Vamvakoussi & Vosniadou, 2004).

There have been many challenges in designing successful educational interventions to promote conceptual change in a wide range of domains (Aleknavičiūtė et al., 2023; Li et al., 2022; Limón, 2001). In the realm of rational number learning, some game-based learning environments have shown potential for teaching rational number concepts that may require conceptual change, such as the size of fractions for primary school students (Braithwaite & Siegler, 2021; Kiili et al., 2018) or the density of the set of rational numbers for lower secondary school students (Kärki, McMullen, Halme, et al., 2021). However, there is little empirical evidence of the effectiveness of these game-based learning environments for supporting conceptual change and, like any educational setting, these game-based learning interventions are not infallible and students most likely have inter-individual differences in their

☆ This article is part of a special issue entitled: Gamified learning process published in Learning and Individual Differences.

* Corresponding author at: Department of Teacher Education, FI-20014, University of Turku, Finland.
E-mail address: topeka@utu.fi (T. Kärki).

https://doi.org/10.1016/j.lindif.2025.102643
Received 11 March 2024; Received in revised form 7 January 2025; Accepted 29 January 2025
Available online 7 February 2025

learning outcomes. It is possible that how students play the game may have important implications for their learning. Therefore, in the present study, we aim to examine what kind of game play in the NanoRoboMath game supports students' conceptual knowledge about the effects of multiplicative operations with numbers less than one, knowledge that may require substantial conceptual change of the rational number concept to fully understand.

1.1. Conceptual change with rational numbers

Learners often struggle with rational numbers due to a phenomenon called natural number bias, which is the tendency to inappropriately apply natural number features in rational number tasks (Ni & Zhou, 2005; Vamvakoussi & Vosniadou, 2004). Multiple theories of conceptual change argue that prior conceptions can constrain students' ability to learn new content (e.g., Smith III et al., 1994; Vosniadou, 2014). In mathematics, previous research has shown that conceptual change is needed to accommodate learners' initial natural-number-biased ideas to the new aspects of rational numbers that are incompatible with their prior knowledge (e.g., Vamvakoussi & Vosniadou, 2004). Typical examples of natural-number-biased errors relate to rational number size, representations, operations, and density (Van Hoof, Vandewalle, et al., 2015). In the present study, we focus on the effects of arithmetic operations with rational numbers, especially the misconception that multiplication will always lead to a larger and division to a smaller outcome.

Many primary and secondary school students struggle to move past the misconception that "multiplication makes bigger" and "division makes smaller" (Christou, 2015; Fischbein et al., 1985; Van Hoof, Vandewalle, et al., 2015). With multiplication and division by numbers less than one, the effects of the operations are the opposite (e.g. $0.4 \times 0.5 = 0.2$ and $3 \div 0.5 = 6$). Earlier research has shown that students mental models of arithmetic operations, such as repeated addition for multiplication, are in conflict with rational numbers less than one, which causes difficulty in performance (Fischbein et al., 1985). This extends to situations in which missing values are assumed to be natural numbers and therefore results of multiplication and division operations should follow the typical patterns (Christou, 2015; Vamvakoussi et al., 2013). Overall, these studies collectively highlight the persistence of natural number biases in primary and secondary school students and the need for educational strategies to address these misconceptions.

The elements of rational numbers that conflict with natural numbers are seldom clearly articulated in math teaching resources (Van Dooren et al., 2019). Furthermore, instructional interventions designed to stimulate conceptual change have shown limited impact on student learning, for example in secondary and high school students' understanding of density (Vamvakoussi & Vosniadou, 2012). This issue is exacerbated by the fact that teachers may lack comprehensive knowledge and teaching skills related to rational numbers (Depaepe et al., 2015). Without clear instructional materials, the natural number biases interfering with rational number concepts may not be adequately addressed. On the one hand, strategies that directly challenge students' misconceptions have proven effective. For instance, refutational texts that systematically debunk students' incorrect beliefs have been shown to facilitate conceptual change across various scientific and mathematical disciplines in all levels of education from elementary to university students (Aleknavičiūtė et al., 2023; Christou & Prokopou, 2019; Mikkilä-Erdmann, 2001; Tippett, 2010). On the other hand, scholars argue for a more collaborative approach, suggesting that educators should align with students' intuitions and guide their attention towards expert-aligned properties in a supportive manner (Smith III et al., 1994).

The process of gradually shifting students' intuitive number concepts towards rational number reasoning presents challenges for both learners and teachers (Vosniadou et al., 2008). It is clear that the goal of instruction should be to equip intentional learners with the metacognitive skills and diverse perspectives needed to overcome the limitations of their initial understanding (Vosniadou & Verschaffel, 2004). Encouragingly, some research suggests that game-based learning environments could provide valuable support for primary and secondary school students in this process (Ketamo & Kiili, 2010; Koops & Hoevenaar, 2013). Further investigation is required to ascertain the most effective teaching methods in different scenarios (Leonard et al., 2014).

1.2. Game-based learning for conceptual change in mathematics education

The integration of playful learning with technology can enhance mathematical thinking by providing innovative tools for the discovery of mathematical concepts (Brezovszky et al., 2019; Devlin, 2011). This integration can positively impact students' attitudes towards mathematics and improve students' mathematical achievement (Baker et al., 2015; Guerrero et al., 2004). However, technology is often used merely as an alternative content-delivery method, without it providing novel ways to interact with educational content (Bray & Tangney, 2017). Most game-based learning environments focus on drill-and-practice exercises, which aim at enhancing procedural fluency without paying attention to other aspects of mathematical proficiency such as conceptual understanding, strategic competence, or adaptive reasoning (Byun & Joung, 2018; Laato et al., 2020). As a result, the potential benefits of technology for mathematics education are not fully realized.

In recent years, several studies examining the effects of digital game-based learning environments on players' rational number knowledge have yielded positive results both in primary and secondary education (e.g., Kärki et al., 2022; Masek et al., 2017; McMullen et al., 2023; Nejem & Muhanna, 2013; Riconscente, 2013). One of the main mechanisms by which games may be achieving positive learning effects, including facilitating conceptual change, is the use of visual models to support the cognition processes (Hamdan & Gunderson, 2017; Zhang et al., 2020). For example, the area models and the concept of splitting are used for understanding the part-whole meaning of fractions (e.g., Cyr et al., 2019; Gaggi et al., 2018; Martin et al., 2015). Another type of visual model used in digital games to support the understanding of numerical magnitudes is the number line (e.g., Braithwaite & Siegler, 2021; Ninaus et al., 2017; Riconscente, 2013). Thus, the affordances and efficiencies provided by properly designed digital games may also improve students' conceptual understanding of rational numbers, thereby facilitating conceptual change.

1.3. The NanoRoboMath game-based learning environment

A digital game-based learning environment, NanoRoboMath, was designed to improve students' conceptual knowledge of rational numbers taking into account several natural-number-biased features (Kärki, McMullen, & Lehtinen, 2021). Aligning with suggested design principles of game-based learning environments (Habgood & Ainsworth, 2011), NanoRoboMath integrates rational number concepts and mathematical practices into the core mechanics of the game. It employs the number line metaphor which has been shown to be a core mental representation of numerical magnitude (Göbel et al., 2011; Siegler, 2016) and an effective basis for designing learning environments promoting numerical magnitude knowledge (Riconscente, 2013; Siegler & Ramani, 2009). The number line representation may be particularly powerful for expressing the effects of multiplicative operations with rational numbers, as movement along the number line when using multiplicative operations less than one will be counter to expectations based on whole number properties (e.g., moving to the left, rather than right, when multiplying by a number less than one). Playing NanoRoboMath has been shown to enhance students' conceptual knowledge of rational number representations and operations, mental calculation skills, and adaptive rational number knowledge (Kärki et al., 2022; Kärki, McMullen, Halme, et al., 2021). However, it is unclear what kind of game play of NanoRoboMath supports players understanding of the multiplicative operations with rational numbers.

In the game, players steer a nanorobot along a number line using four basic arithmetic operations and rational numbers. Their task is to clean polluted water, cure pet diseases, or help plants to survive by destroying bacteria, viruses and parasites. The players explore a dynamic, continuous number line, which stretches (zooms in and out) and shifts depending on the players' position relative to the target. The basic idea of the game is to move the nanorobot from the initial position to the target position. In the upper scene of Fig. 1, the nanorobot is in the position 15.5 and the target bacterium is in the position $1\frac{1}{2}$. Unlike in typical equation solving tasks, the player does not need to reach the target directly with one operation. On the contrary, moving freely along the number line by combining several operations in flexible ways is one of the key elements of the game (McMullen et al., 2023). The game also encourages the player to repeat the same tasks in order to improve their performance, as this has been shown to be important for deeper learning in game-based environments (Lee et al., 2022). In the example of Fig. 1, in order to get closer to the target, the player chooses to use division by 10. After pressing the calculate button ("Laske"), the nanorobot moves to the outcome of the chosen operation 15.5/10 = 1.55. The new rescaled number line can be seen in the lower image of Fig. 1. From the new position, the player could reach the target by subtracting 0.05.

The version of NanoRoboMath used in this study contained three playing modes: time, approach and energy. In the *time mode*, the goal is to move the nanorobot to a given interval surrounding the target as quickly as possible. This mode is designed for practicing the estimation of rational number magnitudes, since the player has to make a trade-off between quick approximate and slower precise calculations. In the *approach mode*, the player has to get closer and closer to the target without touching it. This mode is especially designed for understanding the density of the rational number set. In the present study, we focus on the *energy mode* where the players try to reach the target by consuming as little "energy" as possible. The energy consumption of a move is equal to the magnitude of the input number. Players earn medals (bronze, silver, gold, diamond) by saving energy on each level. The consumption limits of the medals and the remaining energy of the nanorobot are indicated on the bar at the top of the game scene (see Fig. 1). The highest medal, a diamond medal, can be obtained only if the player uses multiplication or division by a number less than one in an effective way as a part of their solution. For example, in the example of Fig. 1, to achieve a diamond medal, the player should rather multiply by 0.1 than divide by 10. Together with the subtraction 1.55–0.05 that would lead to an energy consumption of 0.1 + 0.05 = 0.15, which is small enough for obtaining a diamond. After completing each level, the game provides hints on how the players can improve their performance by taking the energy feature better into account. Thus, we argue that this game mode enhances multiplicative thinking and, in particular, challenges the misconception that multiplication makes bigger and division makes smaller.

The elements embedded in the game were not only designed to support multiplicative operations, but also other aspects of rational number knowledge. The zooming feature of the number line should support understanding the density of the rational number set. In addition to routine mental calculation skills needed to figure out suitable input numbers for the moves, the flexibility in moving freely along the

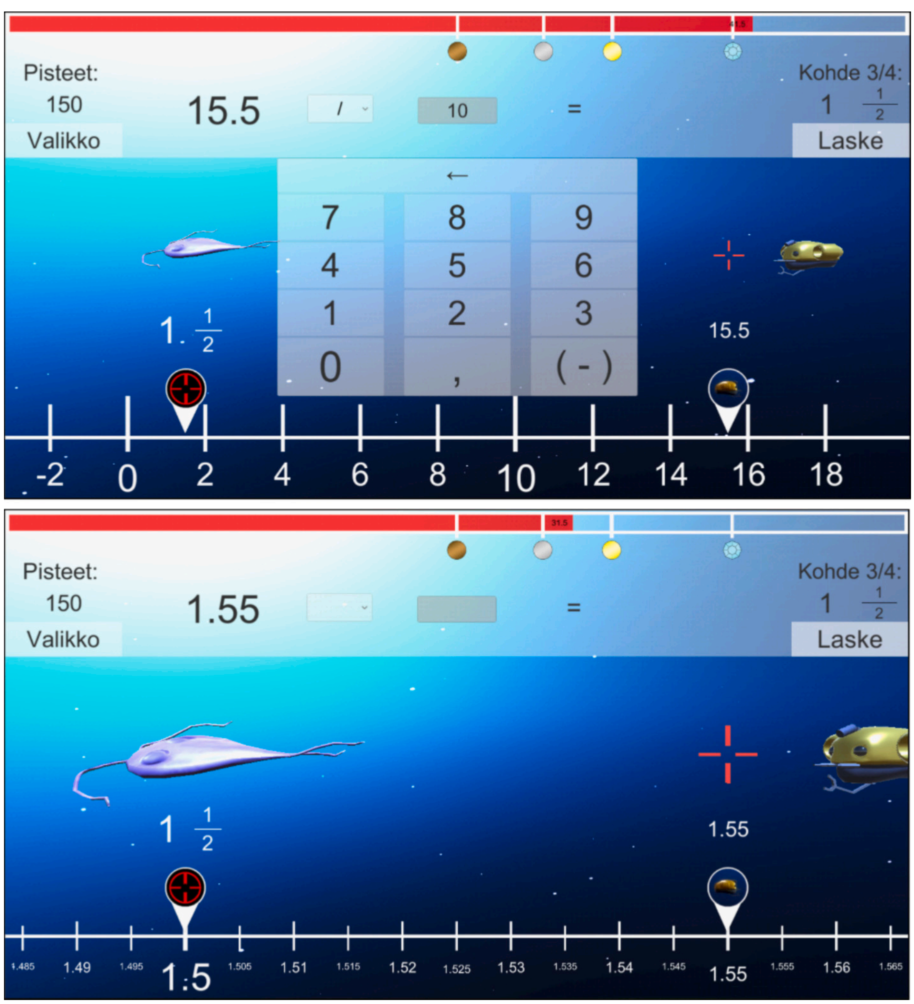

Fig. 1. Two game play images of NanoRoboMath. Translations: Pisteet = Score, Valikko = Menu, Kohde = Target, Laske = Calculate.

number line with different combinations of arithmetic operations should promote players' adaptive rational number knowledge. Adaptive rational number knowledge refers to a rich network of knowledge of numerical characteristics and the arithmetic relations between numbers, which can be flexibly applied in solving novel tasks (McMullen et al., 2020). Moreover, the use of both decimal numbers and fractions and mixing these representations (see, e.g., the location of the nanorobot and the target in Fig. 1) aim to enhance players' cross-notation knowledge of rational numbers.

1.4. Present study

In this study, we use a person-centered approach to examine individual differences in game performance as it can shed light on students' qualitatively distinct approaches to game play. These differences include the quality (e.g. use of multiplicative operations with numbers less than one, success in obtaining diamond medals) and quantity (e.g. length of the game play, number of repetitions of a level) of the students' game play (e.g. Bui et al., 2022; McMullen et al., 2023). We combine these aspects into three indices measuring the amount and persistence of practice, the use of multiplicative strategies and players' level of performance in the game. Furthermore, we examine whether the differences in game play are connected to students' prior knowledge of rational numbers and whether they predict the development of different aspects of rational number knowledge. We pose the following research questions:

1. What kind of player profiles can be found examining the quality and quantity of players' game play from the perspective of engaging with concepts of multiplicative rational number operations?

To address this question, we investigate students' game log data on the NanoRoboMath game to examine differences in how much students practice math skills by playing and replaying the levels, how much they use the beneficial strategies of multiplying and dividing by numbers of magnitude less than one and how well do they succeed in playing the game in the highest performance level (diamond level), which requires a good conceptual understanding of multiplicative operations with rational numbers. Using k-means clustering, we expect to find a diverse set of player profiles that differ in all three aspects of gameplay (practice, strategies, performance). Due to the exploratory and data-driven nature of the analysis, we do not provide specific hypotheses on the exact structure of the profiles.

2. Do the player profiles differ from each other on pre-test measures of rational number knowledge?

Prior knowledge may have an influence on students' game play, as shown in previous studies (Bui et al., 2022). Thus, we examine whether higher prior knowledge is related to the way students play the game.

3. Do the player profiles predict the development of different aspects of students' rational number knowledge?

The NanoRoboMath game performance has been shown to predict later rational number knowledge within 5th and 6th graders (Kärki et al., 2022) and 7th graders (Kärki, McMullen, Halme, et al., 2021). However, previous studies have not differentiated between player profiles that may lead to higher knowledge gains. Thus, we explore how player profiles differ in their post-test rational number knowledge, after controlling for the corresponding pre-test score. We expect a relation between the player profiles and the development of conceptual knowledge of multiplicative operations. Additionally, we examine the development of mental calculation skills and adaptive rational number knowledge due to their close relation to rational number operations.

2. Methods

2.1. Participants

A total of 104 students from nine seventh-grade classes participated in this study. They belonged to four lower secondary schools located in varying socio-economic status areas in a city in southern Finland. To be able to analyze the development of students' knowledge about rational number operations, we excluded students who did not have both pre- and post-test operation knowledge data. We also excluded one player who had played less than one level of NanoRoboMath during the whole intervention. Hence, the final sample size was 90 students (mean age = 13.2 years, SD = 0.41 years).

Participation was voluntary and students were informed of their right to stop taking part in the study at any point. Permission to take part in the study was also asked in written form from the legal guardians of the students. The permission to conduct the study was received from the municipality, school principals, and participating teachers. Ethical permission was granted from the University of Helsinki ethical board and the ethical guidelines of the Finnish national board on research integrity (TENK) were followed.

2.2. Design and procedure

As part of the Growing Mind project, we conducted a pretest-posttest classroom intervention in which the digital game-based learning environment NanoRoboMath was used for teaching and learning rational numbers. This study was a part of a larger intervention testing the effectiveness of the Number Line Elaboration and Exploration (NLEE) learning environment with participants divided into a gaming and a control condition. Students in the gaming condition had the following procedure: pre-test, Number Trace game (see, for example, Kiili et al., 2021), mid-test, NanoRoboMath game, and post-test. The overall effectiveness of the NLEE learning environment has been reported in McMullen et al. (2023). In this article, we concentrate on examining the NanoRoboMath part of the learning environment, as the version of Number Trace did not target supporting rational number operations conceptual knowledge. Prior to playing NanoRoboMath, the students took a computer-based rational number test during one mathematics lesson, and the teachers participated in a 1.5-h training session led by the research team. In addition to giving instructions about the basic mechanics of the game, we encouraged teachers to support struggling students, to elicit classroom discussions of alternative strategies to solve a particular task in the game, and to challenge students to replay the levels and explore the different strategies to improve their scores. The student test in the beginning of the NanoRoboMath intervention is the mid-test of the larger intervention, but for the sake of clarity, it is called the pre-test of this study from now on.

The teachers were asked to have their students play the NanoRoboMath game for five mathematics lessons over a two-week period. The lessons were 45 min in length, but the actual playing time was typically much shorter. According to the game log data, the students played on average 1 h 37 min (SD = 56 min, Range = 18–268 min) in total. Note that although the students were only asked to play during five math lessons, the students had also access to the online game in their free time. Based on the time stamps of the game log data, some students appeared to play NanoRoboMath outside the school hours, which explains why the maximum playing time exceeds the length of five lessons. At the same time, there were students whose playing time was low. We interpret the differences in the playing time as a natural variation in school contexts. Moreover, players with a very small playing time may have still benefited from the intervention, for example, by watching their peers' playing and taking part in the classroom conversations. As we can take the amount of playing into account in our analysis, we do not exclude these students from this study.

Since in this study we examine only energy mode levels, the length of the students' game play was measured using the number of played targets in energy mode levels including replays (Mean = 57.6, SD = 34.2, Range = 10–153). We note that the students could not freely choose which playing mode to play. A sequence of energy (E) and time mode (T) levels was designed by the investigators for the beginning part of the game (ETEEETTTTE repeated four times), and the approach mode tasks were the last ten levels of the NanoRoboMath game. We restrict our consideration into the energy mode levels, because only in this mode were multiplication and division by numbers less than one truly beneficial strategies from the perspective of the game mechanics. The Pearson correlation between NanoRoboMath total playing time and played targets in energy mode was $r = 0.81$. After the final game session, students completed a computer-based post-test, which is used as the outcome measure of the intervention.

2.3. Measures

2.3.1. Pre- and post-test mathematical measures

The computer-based testing sessions lasted approximately 40 min and were led by a member of the research team. The students were not allowed to use calculators in the tests. The pre- and post-tests contained similar but not identical items. Both tests consisted of eight different task types in the following order: arithmetic sentence production, number line estimation, rational number sets, mental arithmetic calculation, mental cross-notation conversion, magnitude ordering, density of rational numbers and knowledge of rational number operations. This study is targeted to examine the conceptual development of students' knowledge of multiplicative rational number operations. In addition to the test measure of operation knowledge, we examine the arithmetic sentence production task as a measure of adaptive rational number knowledge and the mental arithmetic calculation tasks as a measure of routine knowledge. These two additional measures are closely related to rational number operations, unlike the other tasks measured in the computer-based testing.

Conceptual knowledge of multiplicative rational number operations was measured using three multiple choice items adapted from Van Hoof, Janssen, et al. (2015). The items tested students' knowledge of the effects of arithmetic operations using questions like "Is the outcome of 49 × 1/2 smaller or larger than 49?" In addition to "smaller" and "larger" options, there was also the option to choose "I don't know". The items in both tests concerned operations where an integer was either multiplied or divided by a decimal number or a fraction of magnitude less than one. In these items, natural-number-biased reasoning would lead to an incorrect answer. One point was given for each correct answer and the sum of the items with maximum of 3 points was interpreted as the measure of rational number *operation knowledge*. The internal consistency of the operation items was low, the Cronbach's α being 0.66 in the pre-test and 0.62 in the post-test. This might reflect the conceptual differences between the multiplication and division operations and between the different representations of rational numbers. Although Van Hoof, Janssen, et al. (2015) showed that the inhibition of natural-number-biased reasoning is a one-dimensional construct, they also observed that the difficulty level of rational number operation items varied in their test instrument.

Mental calculation skills were measured with 12 arithmetic problems. These included six fraction items (e.g. $\frac{1}{2}+\frac{1}{2}$; $\frac{3}{4}-\frac{1}{4}$; $\frac{1}{2}\div 2$), five decimal items (e.g. 0.25×4; 0.75–0.25; $4 \div 0.5$) and one item with mixed representations (e.g. $\frac{1}{2}\times 0.5$). One point was given for each correct answer with a maximum of 12 points. We interpreted this measure as an example of students' routine rational number knowledge. The Cronbach's α was 0.77 in the pre-test and 0.75 in the post-test.

The arithmetic sentence production task (McMullen et al., 2017) aims to capture students' ability to recognize and use different numerical characteristics and relations of numbers in a multifaceted and flexible way. The test included four items, where the participants had two minutes to form as many valid arithmetic sentences as they could, using a set of five given numbers, which produced a target number. In order to ensure that the respondents had understood the aim of the task, a practice item with whole numbers and a possibility to ask questions after completing the practice item was provided. In each test item, the given numbers included two pairs of equivalent fractions and decimals (e.g. $\frac{1}{2}$ and 0.5, $\frac{1}{4}$ and 0.25), a single whole number (e.g. 4), and a target number (e.g. 1). The participants could use four basic arithmetic operations and each given number as many times as they wanted in each arithmetic sentence they produced. Answers were counted as correct if they were mathematically valid representations of the target number, including equivalent solutions (e.g. $\frac{1}{2}+\frac{1}{2}$ and $0.5 + 0.5$) and commutative equivalents (e.g. $\frac{1}{4}+\frac{1}{2}$ and $\frac{1}{2}+\frac{1}{4}$) but not literal repetitions of previous answers. The average number of correct responses across the four items was interpreted as the measure of *adaptive rational number knowledge* (see McMullen et al., 2020). The reliability was good: Cronbach's α was 0.85 at both time points.

2.3.2. Game performance related measures

A web-browser version of NanoRoboMath with individual player IDs was used to collect detailed information about students' actions in the game. The log data of each player was recorded to the game server. The log data contained the sequence of the NanoRoboMath moves (target number, position of the nanorobot, chosen operation, input number, position of the target, remaining amount of energy) of the player with time stamps. When the player reached a target, his or her performance was evaluated based on the energy consumption on a five-point scale: 1 = pass (high energy consumption), 2 = bronze (additive strategy), 3 = silver (mixed strategy), 4 = gold (multiplicative strategy), 5 = diamond (multiplication or division using a number with magnitude less than one). The scale reflects player's ability to use rational number operations according to the game mechanic.

From the game log data of each player, we formed game performance variables. The names of these variables are emphasized. The length of the game play was measured by the number of played *targets*, including the replayed targets. The replaying of the targets was interpreted as an indication of deliberate practice. A repetition index was calculated for each played target as the number of times the player had played the target before and, then the average of the repetition indices was used as a measure of *replay*. We also measured the effectiveness of replaying by calculating an *improvement* variable indicating how many times the player obtained a better result on the five-point energy consumption scale by replaying a target.

To get detailed information about the use of advanced multiplicative strategies, we separately counted how many times a player used multiplication (*multiplication* < *1*) and division (*division* < *1*) by numbers of magnitude less than one. We were especially interested in the moments of game play where the target was reached with the best level (diamond level) of the five-point energy consumption scale, since it requires the use of multiplication and division by numbers of less than one in a beneficial way for reaching the target. We calculated the number of *diamonds* the player obtained in the game including replayed targets and the percentage of played targets that the player was able to reach in diamond level (*diamonds%*). Moreover, we were interested in knowing at which point in the game the use of advanced strategies led to the first success in reaching a target with a diamond level energy consumption. The position of the first diamond score (*1st diamond*) was calculated relative to the total number of played targets and a value 1 was given for those players who did not obtain any diamonds in the game.

2.4. Analysis

First, we used descriptive statistics to examine the distributions of the game performance variables and investigated correlations between the variables to find out potential key indicators of game performance by using SPSS version 27. In order to handle the complexity of eight game performance variables, we formed the following three sum variables: *Practice index*, *Strategy index* and *Performance index*. The *Practice index* is the sum of the standardized values of *targets*, *replay* and *improvement* (Cronbach's $\alpha = 0.87$). These variables describe the quantity and quality of deliberately trying to make progress in the game. The *Strategy index* is the sum of the standardized values of the *multiplication* < *1* and *division* < *1* variables (Cronbach's $\alpha = 0.83$). This index measures the amount of using both strategies that are beneficial for obtaining the highest performance level in the energy mode (without taking into account the potential success in acquiring diamonds). The *Performance index* is the sum of the standardized values of *diamond*, *diamond%* and reversed *1st diamond* (calculated as 100 % - *1st diamond* so that the higher values indicate that the first diamond was obtained earlier and zero indicates no diamond obtained; Cronbach's $\alpha = 0.89$). It describes how well the player was able to use the beneficial multiplicative strategies for obtaining diamonds.

The player profiles were formed by K-means clustering using the three sum variables described above. Three students were identified as outliers (extremely long game, high performance) and were excluded from the cluster analysis. The final cluster solution was chosen using distance-based criteria (Schubert, 2023) and ensuring that the clusters can be characterized in terms of the game performance variables, i.e., the clusters describe meaningful and different player profiles. A one-way ANOVA and post-hoc analysis with the Bonferroni method for adjusting multiple comparisons were used to analyze how the profiles differed from each other on the pre-test measures.

For each measure of rational number knowledge, a linear mixed model was used to examine if the player profiles predicted learning outcomes when the corresponding group-mean-centered pre-test score was used as a covariant and student's class membership was treated as a level 2 random factor. The interclass correlation coefficients (ICC) of the no predictor models suggested that multi-level modeling of individual and classroom effects may be appropriate in the present study (see Heck et al., 2014): operation knowledge ICC = 0.10, mental calculation skills ICC = 0.18 and adaptive rational number knowledge ICC = 0.11. In each model, we estimated the effect of player profiles and pre-test scores on post-test scores. Hence, level 1 included pre-test scores of the corresponding dependent variable for individual students nested within the classes (level 2). For estimating the classroom effects, we compared a model with a random intercept to a model with a random intercept and a slope using a likelihood-ratio test (West, 2009). We note that due to a technical error, there was some data missing in the pre-test (4 values) and post-test (2 values) of mental calculation skills.

3. Results

3.1. Descriptives for the measures of game performance

To obtain an overall picture of students' gaming processes and the potential key indicators of game performance, we first examined the distributions and relations between the game performance variables. The descriptive statistics of these variables are given in Table 1 and the correlations between the game performance variables and post-test measures are given in Table 2.

The length of the game play (*targets*) varied considerably in our study, the largest amount of targets (with replays) played being >15 times the number of targets in the shortest game play. The mean of the *replay* variable was 0.91. We note that if each target were replayed once, the replay value would be 0.50. Hence, the measured mean value indicates that either replaying was frequent across targets or some targets were replayed quite many times. The average amount of *improvements* per student was about 6.5. The *replay* and *improvement* variables were strongly correlated with each other ($r = 0.61$, $p < .001$) and moderately or strongly correlated with *multiplication* < *1* and *division* < *1* (see Table 2).

The variables *multiplication* < *1* and *division* < *1* correlated strongly with each other ($r = 0.70$, $p < .001$). Both variables also correlated positively with *diamonds%*, but *division* < *1* appeared to have a stronger relation with *diamonds%* ($r = 0.64$, $p < .001$) than *multiplication* < *1* ($r = 0.48$, $p < .001$). Using quantpsy.org online-calculator (Lee & Preacher, 2013), the difference between these correlations was confirmed statistically significant ($Z = 2.49$, $p = .01$). Moreover, *division* < *1* was significantly correlated with all post-test rational number measures. Multiplication by a number of magnitude less than one was somewhat more frequent than division. There were students who did not use these advanced multiplication and division strategies at all, and about a quarter of students used the strategies at most 1 or 2 times. The variables *multiplication* < *1* and *division* < *1* indicate the overall use of advanced multiplicative strategies, also in the case when the player did not succeed in obtaining a diamond. The unsuccessful explorations might be equally important for the conceptual change processes of understanding multiplicative operations since the direction of the movement of the nanorobot enables the player to contemplate the effects of multiplication and division also in this case.

In addition, the number of *diamonds* and the relative amount *diamonds%* varied heavily. Both of them correlated significantly with all post-test rational number measures. The number of *diamonds* was strongly correlated with *targets* ($r = 0.63$, $p < .001$) whereas *diamonds%* was only weakly correlated ($r = 0.23$, $p = .03$). However, we still find it useful to consider both the absolute and the relative amount of diamonds since, for example, obtaining just a few occasional diamonds in a short game might indicate a high relative amount although diamond level performance was not frequent. On average, the players reached the diamond level in almost one fifth of the targets that they played. The first diamond (*1st diamond*) was typically obtained around the midpoint of the game play. If the player discovers the use of the advanced strategy late in the game, it is natural that the relative number of diamonds is low. Indeed, there was a strong negative correlation ($r = -0.80$, $p < .001$) between *1st diamond* and *diamonds%*, and *1st diamond* was also significantly negatively correlated with the use of advanced multiplicative strategies. We note that if the first diamond is obtained late but the relative number of diamonds is high, this indicates very purposeful and efficient use of the advanced strategies at the final part of the game play.

Table 1
Descriptive statistics of the game performance variables.

Game performance variable	*Mean*	*SD*	*Min*	*Max*	*Skewness*	*Kurtosis*	*Q1*	*Q2*	*Q3*
Targets	57.60	34.23	10	153	0.48	−0.49	25	59	82
Replay	0.91	0.80	0	3.69	1.22	1.14	0.30	0.60	1.39
Improvements	6.48	5.01	0	23	1.02	0.69	3	5.5	9
Multiplication<1	12.24	12.38	0	56	1.33	1.93	1.75	10.5	20
Division<1	11.23	12.15	0	61	1.60	1.99	1	6	14.25
Diamonds	9.36	10.65	0	45	1.61	2.44	1	6	13.25
Diamonds% [%]	17.96	17.53	0	71.43	1.13	0.81	2.72	14.12	26.35
1st diamond [%]	57.58	34.44	1.09	100	−0.18	−1.43	22.22	61.64	95.22

Note. SD = standard deviation, Q1 = 1st quartile, Q2 = 2nd quartile, Q3 = 3rd quartile.

Table 2
Pearson correlations of the game performance variables and post-test rational number measures.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Targets										
2. Replay	0.61***									
3. Improvements	0.85***	0.61***								
4. Mult<1	0.63***	0.51***	0.58***							
5. Division<1	0.61***	0.50***	0.61***	0.70***						
6. Diamonds	0.63***	0.39***	0.55***	0.74***	0.83***					
7. Diamonds%	0.23*	0.21*	0.30**	0.48***	0.64***	0.74***				
8. 1st diamond	−0.23*	−0.21*	−0.31**	−0.43***	−0.58***	−0.64***	−0.80***			
9. Oper. knowl.	0.05	0.02	0.14	0.18	0.36***	0.36***	0.43***	−0.39***		
10. Mental calc.	0.11	0.06	0.08	0.16	0.36***	0.34**	0.46***	−0.40***	0.40***	
11. ArNK	0.23*	0.13	0.24*	0.28**	0.51***	0.49***	0.53***	−0.41***	0.45***	0.60***

Note. Mult<1 = Multiplication<1, Oper. knowl. = operation knowledge, Mental calc. = mental calculation skills, ArNK = adaptive rational number knowledge.
* $p < .05$
** $p < .01$
*** $p < .001$

3.2. Player profiles

Based on Within-Cluster Sum of Squares (WCSS), Dunn index (Schubert, 2023) and meaningful interpretation of the clusters in terms of game performance variables, we decided to use the four clusters solution for the player profiles. The player profiles were named according to their game play characteristics as 1) High, 2) Short but strong, 3) Long play low gain, and 4) Low. The final cluster centers of the player profiles are given in Table 3 and the standardized mean values with standard errors for each game play indicator are depicted in Fig. 2 (see Appendix A for a more detailed visualization of each profile members' scores for all the indicators).

Players in the *High* profile had mainly high values for all game performance variables. They played long games (54–128 targets), multiplied and divided many times with numbers of magnitude less than one, and they were able to obtain lots of diamond medals. In particular, these players used efficiently the division strategy (all values of division<1 belonged to the highest quartile). In fact, division by numbers less than one was more common than multiplication by numbers less than one. The first diamond of a player was usually obtained early in the game. In this profile, the replay values were above the average and players were also able to improve their scores by replaying the targets.

The *Short but strong* profile consists of players who managed to use beneficial multiplicative strategies so that the percentage of targets that were awarded diamonds was fairly high, even though their game was quite short. More precisely, the number of targets they played belonged quite evenly to the three lowest quartiles (see Table 1), whereas diamond% values were all above the median. There was more multiplication than division by numbers of magnitude less than one. Although replying was infrequent (below the average), some improvements were obtained (most values in the second and third quartile).

Players in the *Long play low gain* profile had a high practice index but a low performance index. They played long games (targets variable values mainly in the two highest quartiles). About 40 % of the players in this profile obtained a high number of diamonds (in two highest quartiles), but the diamonds% values were relatively low (mainly in the two lowest quartiles). Hence, this means that obtaining diamonds was sparse. Typically, they got their first diamond during the final quarter of the game play. The use of beneficial multiplicative strategies was generally quite modest. Three players never used any of these strategies. However, four players in this profile used a lot of multiplication and one player a lot of division by numbers of magnitude less than one, but nonetheless they were not able to obtain many diamonds. The values of the replay variable in this profile (0.31–2.86) were divided quite evenly across the three highest quartiles and the players made improvements more than average. In other words, the players in this profile used a lot of effort to play and improve in the game, but their performance did not reach the highest level.

The members of the *Low* profile had mainly low values for each of the game performance variables. The number of played targets varied between 10 and 67 and half of the values belonged to the lowest quartile. About half of the players did not obtain any diamonds in the game, and nine players got just one or two. Moreover, this profile contains ten players who did not use any advanced multiplicative strategies during the game. In general, the replay values were low in this profile and the players did not make many improvements. If the players obtained diamonds, it happened usually quite late in the game.

Table 3
Final cluster centers of the player profiles.

	High	*Short but strong*	*Long play low gain*	*Low*
Number of students	16	18	21	32
Final Cluster Centers				
Practice index	2.51	−1.18	1.69	−2.34
Strategy index	2.26	−0.20	−0.45	−1.24
Performance index	3.29	1.68	−1.58	−2.06

3.3. Relation of profiles to pre-knowledge and learning outcomes

Next, we consider the relation between NanoRoboMath game play and player's rational number knowledge. The descriptive statistics of the pre- and post-test measures of rational number operation knowledge, mental calculation skills and adaptive rational number knowledge are given in Table 4.

A one-way ANOVA and post-hoc tests were performed to evaluate whether the player profiles differed from each other with respect to the pre-test measures of rational number knowledge. Table 5 shows the means and standard deviations of the pre-test measures for the profiles. The profiles did not differ on the pre-test means of operation knowledge ($F(3,83) = 2.02, p = .12$). However, the profiles differed from each other on the measures of mental calculation skills ($F(3,79) = 2.97, p = .04, \eta^2 = 0.10$) and adaptive rational number knowledge ($F(3,83) = 3.98, p = .01, \eta^2 = 0.13$). Post-hoc tests indicated that the Short but strong profile had a significantly higher mean than the Low profile for mental calculation skills ($p = .04$) and the High profile had a significantly higher mean than the Long play low gain profile for adaptive rational number knowledge ($p = .04$).

To address our third research question, we used linear mixed models to determine statistically significant differences between the profiles on post-test measures of mathematical knowledge, while controlling for the corresponding pre-test score and taking into account classroom membership as a level 2 random factor. For each post-test measure, we compared two models: a random intercept model and a model with a random intercept and a random slope of the pre-test score. The inclusion of the random slope improved the model fit only for mental calculation

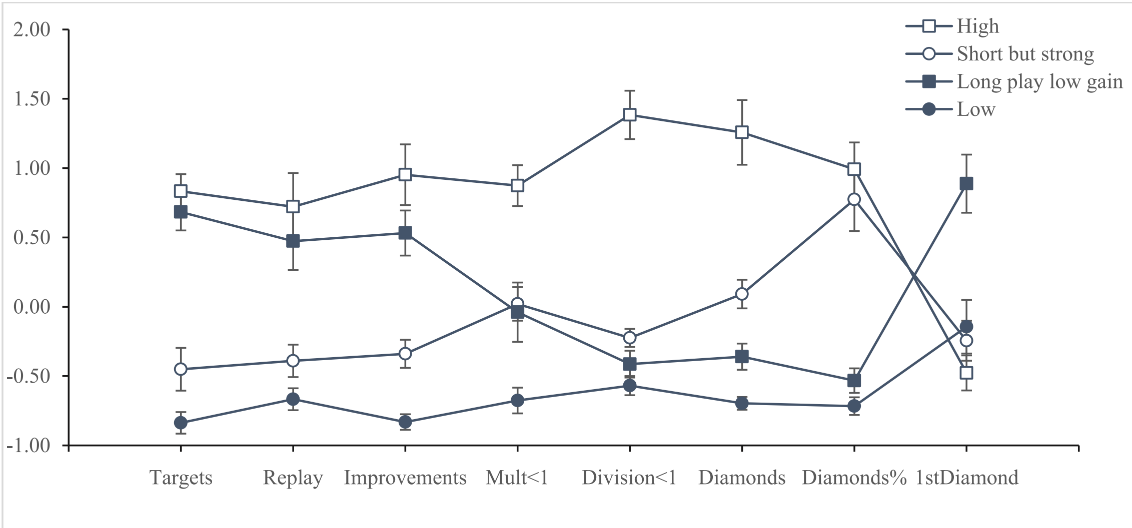

Fig. 2. Standardized mean values for each game play indicator by profile. Error bars = ±1 S.E.

Table 4
Descriptive statistics of the operation knowledge, mental calculation skills and adaptive rational number knowledge in pre- and post-tests.

	n	*Mean*	*SD*	*Min*	*Max*	*Skewness*	*Kurtosis*
Operation knowledge							
Pre-test	90	1.23	1.10	0.00	3.00	0.24	−1.32
Post-test	90	1.70	1.10	0.00	3.00	−0.27	−1.24
Mental calculation							
Pre-test	86	7.21	2.53	1.00	12.00	−0.26	−0.49
Post-test	88	8.65	2.37	1.00	12.00	−0.89	0.98
ArNK							
Pre-test	90	3.48	2.31	0.00	8.75	0.62	−0.48
Post-test	90	5.40	2.55	0.50	12.50	0.66	0.29

Note. SD = standard deviation, ArNK = adaptive rational number knowledge.

skills (see Table 6). Hence, we used the random intercept model for operation knowledge and adaptive rational number knowledge and the random intercept and random slope model for mental calculation skills. All models had the group mean centered standardized pre-test score as a covariant and profile as a fixed factor in level 1. We used the “variance components” covariance structure for the level 2 variables.

The estimates of the fixed effects and covariance parameters of the linear mixed models are reported in Table 7. There was a significant effect of the corresponding pre-test knowledge on the post-test operation knowledge ($\beta = 0.41$, SE = 0.11, $t(73.84) = 3.69$, $p < .001$), mental calculation skills ($\beta = 1.04$, SE = 0.31, $t(10{,}21) = 3.31$, $p = .008$) and adaptive rational number knowledge ($\beta = 1.81$, SE = 0.23, $t(74.03) = 8.00$, $p < .001$). The omnibus tests (Type III tests) of fixed effects indicated that profile was a significant predictor, when taking the other predictors into account, in the models for operation knowledge ($F(3, 80.21) = 3.47$, $p = .02$) and mental calculation skills ($F(3, 70.92) = 3.39$, $p = .02$), but not for adaptive rational number knowledge ($F(3, 79.88) = 0.36$, $p = .78$). The estimated marginal means and standard errors of the profiles for the dependent variables are depicted in Fig. 3. A post-hoc test indicated that the High profile obtained a significantly higher post-test operation knowledge compared to the Long game low gain profile (Mean difference = 0.94, SE = 0.31, $p = .02$). Other profiles did not differ from each other significantly on operation knowledge. Moreover, there were no statistically significant differences between the profiles on mental calculation skills and adaptive rational number knowledge. However, we note that the differences in Mental calculation skills between the Low and the High (Mean difference = −1.33, SE = 0.51, $p = .07$) and between the Low and the Short but strong (Mean difference = −1.33, SE = 0.53, $p = .09$) were close to being significant.

Table 6
-2 times log-likelihood for linear mixed models with random intercept and models with random intercept and slope for each dependent variable.

	Random intercept	Random intercept and slope	Difference in -2 REML LL
Operation knowledge	239.15	239.03	0.12
Mental calculation skills	312.29	304.52	7.77
Adaptive rational number knowledge	340.61	340.61[†]	0.00

Note. -2 REML LL = −2 times log-likelihood calculated using restricted maximum likelihood estimation. Critical value of chi-squared distribution ($p < .05$) for $\Delta df = 1$ is 3.84. [†]Hessian matrix not positive definite.

Table 5
Means and standard deviations of the pre- and post-test measures for the player profiles.

	Operation knowledge		*Mental calculation*		*ArNK*	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Group	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
High	1.63 (1.20)	2.38 (0.81)	7.53 (2.36)	9.56 (1.59)	4.41 (2.33)	6.69 (2.02)
Short but strong	1.44 (1.04)	1.94 (1.11)	8.63 (2.22)	9.47 (2.03)	4.21 (2.49)	5.58 (2.57)
Long play low gain	0.86 (1.06)	1.14 (1.11)	6.76 (2.61)	8.00 (2.85)	2.40 (2.07)	4.51 (2.20)
Low	1.06 (1.05)	1.50 (1.02)	6.48 (2.53)	7.97 (2.29)	2.98 (1.83)	4.78 (2.38)

Note. ArNK = Adaptive rational number knowledge.

Table 7
Estimates of fixed effects (coefficients) and covariance parameters (variances) for the linear mixed models. Standard errors for all values are reported in parentheses.

	Operation knowledge	Mental calculation skills	Adaptive rational number knowledge
Intercept	1.49 (0.21)***	7.81 (0.51)***	5.07 (0.44)***
Pre-test	0.41 (0.11)***	1.04 (0.31)**	1.81 (0.23)***
Profile = High	0.73 (0.31)*	1.33 (0.51)*	0.45 (0.58)
Profile = Short but strong	0.33 (0.28)	1.33 (0.53)*	−0.17 (0.53)
Profile = Long play low gain	−0.21 (0.27)	0.76 (0.47)	−0.04 (0.50)
Residual variance	0.81 (0.13)***	1.94 (0.35)***	2.68 (0.44)***
Random intercept variance	0.15 (0.14)	1.65 (0.99)*	0.94 (0.66)
Random slope variance	–	0.46 (0.35)	–

Note. The significance was tested using *t*-test for fixed effects and the Wald *Z* -test *p*-value for estimating the significance of covariance parameters was split in half as suggested by Heck et al. (2014). The estimate of the coefficient for Low profile was set to zero and is therefore redundant.
* $p < .05$
** $p < .01$
*** $p < .001$

4. Discussion

The aim of the current study was to investigate how individual differences in the NanoRoboMath game play are related to the development of conceptual knowledge of rational number operations. The game was designed to confront students with their misconceptions about multiplicative operations with rational numbers, such as multiplication makes bigger and division makes smaller. Within students' game play, we focused on examining the players' use of advanced multiplicative strategies (i.e., multiplying and dividing by numbers of magnitude less than one), length of the game play and replay indices as well as players ability to obtain diamond medals (highest performance level) in the game. Based on practice index, strategy index and performance index, we were able to find four distinct player profiles. The High profile had high values and the Low profile had low values for all indices. The Short but strong profile players had a low practice index, but a high performance index. On the contrary, players in the Long play low gain group made a lot of effort playing the game, but were not able to obtain many diamonds. The profiles could be characterized by the game performance measures, which suggests qualitative differences in game play.

Interestingly, the profiles differed statistically significantly from each other on the pre-test measures of mental calculation skills and adaptive rational number knowledge, but not on the pre-test operation knowledge although the formation of the profiles was mainly based on the advanced use of multiplicative operations in the game. Nevertheless, the profiles did not predict the development of mental calculation skills or adaptive rational number knowledge. Instead, we observed differences between the profiles in the development of conceptual knowledge of multiplicative operations, indicating effects of game play on learning outcomes. The students in the High profile improved their understanding of multiplicative operations more than the students in the Long play low gain profile. On one hand, this result could be interpreted as further evidence that changing natural-number-biased understanding about rational number operations is a long-standing process and that being able to occasionally obtain a diamond in the game does not necessarily indicate a firm understanding of multiplicative operations. On the other hand, frequent and successful use of multiplication and division by numbers with magnitude less than one within the NanoRoboMath game appears to improve students' understanding of the effects of these operations.

4.1. Implications for conceptual change research

According to previous research (see, e.g., McMullen et al., 2018; Van Hoof et al., 2018), conceptual change is required to fully understand the new aspects of rational numbers that are incompatible with learner's

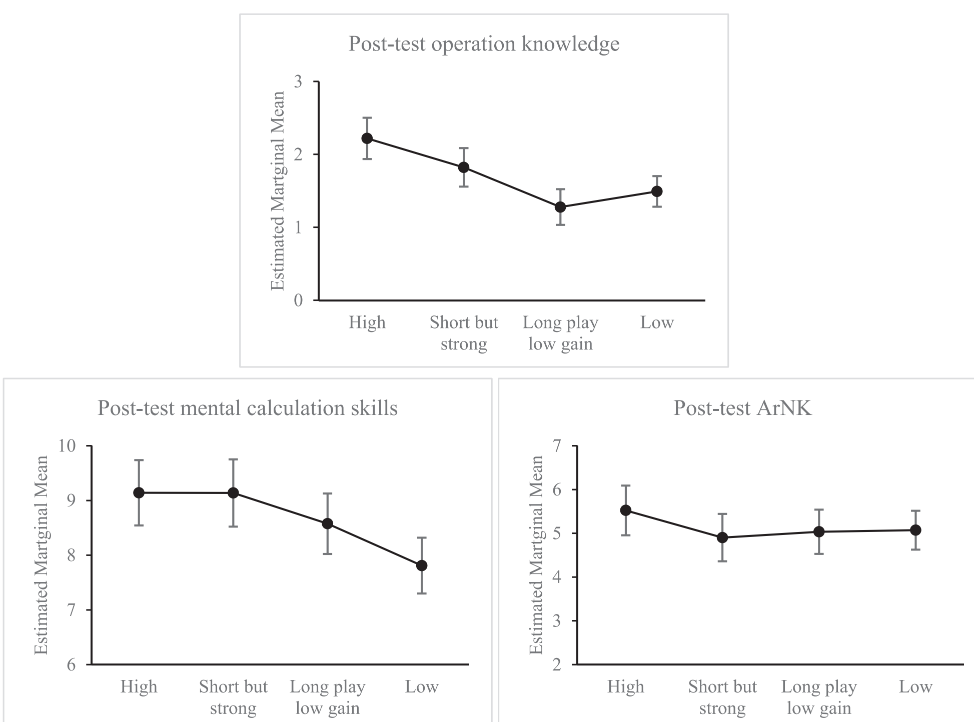

Fig. 3. Estimated marginal means of the post-test operation knowledge, mental calculation skills and adaptive rational number knowledge for the profiles in the linear mixed models. Covariants appearing in the model are evaluated at 0.00 for operation knowledge and adaptive rational number knowledge and at 0.003 for mental calculation skills, where the standardized value differs slightly from zero due to two missing values in the post-test.

prior knowledge. The NanoRoboMath game has been developed as an instructional tool to improve students rational number knowledge. It enables students to confront the misconception that multiplication always makes bigger and division smaller and thus provides experiences that support students' conceptual change processes in developing correct understanding about multiplicative operations with rational numbers.

The present study suggests that not only is the game successful in providing these experiences, but it appears that a high concentration of such experiences may be crucial for conceptual development. In other words, for improving students' knowledge about multiplicative operations, it appears to be important that the students frequently use multiplications and divisions with numbers less than one (either successfully or not) in the game and consistently get immediate visual feedback about the effects of the operations through the movements of the nanorobot. This is apparent from the behavior of the students in the High profile. When these opportunities to experience the counter-intuitive effects of multiplicative operations are less concentrated (i.e. spaced out over longer stretches of game play), it appears that they may have a weaker effect on potential conceptual change. In line with notions of deliberate practice (Lehtinen et al., 2017), a high number of attempts to improve performance by replaying levels appears to be advantageous. However, if the replaying and improvements in the game do not frequently lead to the successful use of multiplication and division with numbers less than one, it might indicate that the conceptual change process is still not complete.

Although there were only a few statistically significant differences between the profiles on the pre-test rational number knowledge, considering the small group sizes and the descriptive statistics, we could hypothesize that at least some prior knowledge may be beneficial in engaging with the game at the most fruitful level. We also noticed that players in the High profile used especially division by numbers of magnitude less than one more than the players in the other profiles. It is possible that understanding division by numbers of less than one is conceptually more demanding than understanding multiplication and therefore the ability to successfully use division in NanoRoboMath might be a more important indicator of the correct understanding of rational number operations. This should be further investigated in future research.

Notably, consistent experiences with the counter-intuitive effects may not necessarily be enough for all students to gain a mathematical correct understanding of the effects of multiplicative operations with numbers less than one. The results of the current study show that some students clearly had inconsistent application of their conceptual knowledge of multiplicative operations across the testing and gaming situations. This is in line with previous research (Schneider & Hardy, 2013) suggesting that conceptual knowledge and related misconceptions may be, at times, situated in nature. In the present study, this most clearly manifests in students who consistently use multiplicative operations with numbers less than one in the game, but struggle to do so on the post-test. One possibility is that students are able to reason about the effects of the operations when they can see the visual manifestation of the operation on the number line in the form of the robot moving in the opposite direction of what they intuitively expect, but they cannot extrapolate this phenomenon when faced with the straightforward and decontextualized question regarding the effects of multiplicative operations. Notably, there are also students who perform well on the tests, but fail to recognize and use this strategy within the game. Given that the game and the tests use both fractions and decimal representations, another possibility is that the inconsistencies are due to lack of a coherent framework around the effects of rational number operations as a whole. Prior research shows that students have inconsistent conceptual knowledge about fractions and decimals (Vamvakoussi & Vosniadou, 2004). Nevertheless, the current results align with previous research on conceptual change which suggests that cultural tools (in this case, the number line) may mediate students' ability to reason about concepts that require conceptual change (Schoultz et al., 2001).

4.2. Implications for game-based learning research

Previous research highlights that, within games with integrated content and game play, game performance and learning outcomes should be related to each other (Brezovszky, 2019; Devlin, 2011). In our novel approach to examine individual differences in game performance, we expand upon this work by investigating a wider range of game performance measures, including both the quality and the quantity of gameplay. Based on these measures, we were able to find player profiles which differed in the amount of practice, the use of advanced strategies, and the amount and frequency of high scores in the game. Although some players in the Long play low gain profile obtained many high scores, these occurrences were sparse in the game play. Hence, based on the differences in learning outcomes between the High and Long play low gain profiles, the total number of high scores appears less relevant than the proportion of game tasks that achieved these scores. This suggests that creating more opportunities for the concentrated and consistent use of advanced multiplicative strategies, and thereby gaining more potent exposure to and experience with their counter-intuitive effects, may increase the likelihood of conceptual change for more students.

Overall, the profiles appear to have validity for capturing important individual differences in learning processes, indicating that there are multiple avenues to achieving positive learning outcomes in the game. The players in the High profile appeared to improve their conceptual knowledge of operations more than the players in the Long play low gain profile, but this was not the case with other measures of rational number knowledge. This indicates discriminant validity of the profiles, which were designed specifically to capture game play performances that differed in the use of advanced multiplicative strategies. In addition, it highlights that the alignment of the examined measures of game performance and learning outcomes is important for understanding the effects of game play on learning outcomes. This may be especially significant in light of the aim of promoting conceptual change in rational number conceptual understanding within the present study. It was particularly the engagement with the effects of multiplicative operations that we expected would be most fruitful in facilitating conceptual change. It should be acknowledged that overall game performance measures (i.e., number of points, number of completed targets) might not be the most appropriate metrics for determining how game play and learning outcomes are related. This is in line with previous research that suggests that the relation between learning outcomes and game play should be considered from multiple angles (Brezovszky et al., 2019).

4.3. Limitations and future directions

Despite promising findings, the present study contains limitations that should be considered and provides suggestions for new directions in research.

First, while the process of defining the playing profiles was built upon theory, the profiles and subsequent results are exploratory in nature. Moreover, even though our sample size was not small, conclusions linked to the smallest clusters should be replicated in larger samples. However, we argue that although our multi-faceted approach did not allow us to test a specific hypothesis, it generated the grounding for future hypotheses. For example, future studies could test whether the concentration of experiences that are in conflict with student's intuitive preconceptions is more important than the total number of such experiences. Nevertheless, an important finding of the present study is the clear existence of multiple profiles of game play and that these profiles can be related to learning outcomes.

Second, we acknowledge that using an educational game for mathematics teaching and learning is not a straightforward task. Earlier research (see, for example, Sun et al., 2023) indicates the crucial role of the teacher in scaffolding game-based learning. The teacher directs the learning process and acts as a facilitator by probing questions, discussing, explaining, and providing hints so that the students can build a link between mathematical knowledge and game content. While the NanoRoboMath game provided hints towards the benefits of using multiplicative strategies and lots of immediate feedback as the nanorobot moved according to the results of the arithmetic operations, these may have been insufficient scaffolds for some students without adequate guidance provided by the teacher. This may explain some of the inconsistencies between the test results and game performance. Moreover, in our analysis we ignored the time and approach mode tasks where the students also sometimes used multiplication and division by numbers of magnitude less than one. Although these multiplicative strategies were not especially beneficial from the game mechanic point of view in the non-energy modes, these might have influenced their learning. In addition, there was variation in the implementations of the NanoRoboMath experiment in different classrooms. In some classes, majority of the students progressed through most of the levels, whereas in some classes, none of the students progressed beyond the first decimal number levels and their playing time was very limited. Hence, in addition to individual skill and motivation differences, the game behavior of the students may have been influenced by the different ways of implementing game-based learning within the classroom. However, this is to be expected in naturalistic intervention settings and it was taken into account in the statistical analysis by using multilevel modeling.

Third, our measure of conceptual operation knowledge is admittedly quite narrow and the Cronbach's alpha value for the operation items was low (0.66 in pre-test and 0.62 in post-test). A set of three multiple choice questions does not provide much opportunity for students to display their conceptual knowledge, nor does it provide much variation in students' performance. This may be one reason for the discrepancies between performance in the game and on the tests. However, we emphasize the discriminant validity of the profiles for operation knowledge in comparison to the other test performance measures that were not aligned with our profiling approach. Nonetheless, future studies should use more comprehensive test measures of conceptual knowledge, and provide more clarity regarding students' strategy use in the game (for example, with think-aloud protocols).

5. Conclusions

The present study represents an attempt to bridge research in conceptual change with game-based learning by examining how game performance metrics can be utilized to examine conceptual development. We find that those students who had a more concentrated dosage of experiences that involved exposure of the potential inconsistencies in their prior knowledge and mathematically correct concepts appeared more likely to improve in their conceptual knowledge. These results provide further evidence that game-based learning environments may be useful tools for supporting conceptual change. We also suggest that for capturing specific learning processes, it is necessary to identify the most relevant fine-grained and well-calibrated measures of game performance.

CRediT authorship contribution statement

Tomi Kärki: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Hilma Halme:** Writing – review & editing, Methodology, Conceptualization. **Minna Hannula-Sormunen:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Erno Lehtinen:** Writing – review & editing, Methodology, Conceptualization. **Jake McMullen:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Funding

This work was supported by the Academy of Finland Strategic Research Council [grant number 336068] awarded to the third author; Academy of Finland [grant number 331080] and the Jacobs Foundation Fellowship awarded to the last author.

Declaration of competing interest

None.

Acknowledgements

We gratefully acknowledge the support of the Academy of Finland (Grant 331080), the Strategic Research Council of the Academy of Finland (Grant 336068), the Jacobs Foundation Fellowship, and all participants and teachers involved in this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lindif.2025.102643.

References

Aleknavičiūtė, V., Lehtinen, E., & Södervik, I. (2023). Thirty years of conceptual change research in biology – A systematic review and meta-analysis of intervention studies. *Educational Research Review, 41*, Article 100556. https://doi.org/10.1016/j.edurev.2023.100556

Baker, J. M., Martin, T., Aghababyan, A., Armaghanyan, A., & Gillam, R. (2015). Cortical activations during a computer-based fraction learning game: Preliminary results from a pilot study. *Technology, Knowledge and Learning, 20*(3), 339–355.

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

Braithwaite, D. W., & Siegler, R. S. (2021). Putting fractions together. *Journal of Educational Psychology, 113*(3), 556–571. https://doi.org/10.1037/edu0000477

Bray, A., & Tangney, B. (2017). Technology usage in mathematics education research – A systematic review of recent trends. *Computers and Education, 114*, 255–273. https://doi.org/10.1016/j.compedu.2017.07.004

Brezovszky, B. (2019). *Using game-based learning to enhance adaptive number knowledge* [University of Turku]. http://urn.fi/URN:ISBN:978-951-29-7622-5.

Brezovszky, B., McMullen, J., Veermans, K., Hannula-Sormunen, M. M., Rodríguez-Aflecht, G., Pongsakdi, N., … Lehtinen, E. (2019). Effects of a mathematics game-based learning environment on primary school students' adaptive number knowledge. *Computers & Education, 128*, 63–74. https://doi.org/10.1016/j.compedu.2018.09.011

Bui, P., Hannula-Sormunen, M. M., Brezovszky, B., Lehtinen, E., & McMullen, J. (2022). Promoting adaptive number knowledge through deliberate practice in the number navigation game. In K. Kiili, K. Antti, F. de Rosa, M. Dindar, M. Kickmeier-Rust, & F. Bellotti (Eds.), *Games and learning Alliance* (pp. 127–136). Springer International Publishing. https://doi.org/10.1007/978-3-031-22124-8_13.

Byun, J., & Joung, E. (2018). Digital game-based learning for K-12 mathematics education: A meta-analysis. *School Science and Mathematics, 118*(3–4), 113–126. https://doi.org/10.1111/ssm.12271

Christou, K. (2015). Natural number bias in operations with missing numbers. *ZDM Mathematics Education, 47*(5), 747–758. https://doi.org/10.1007/s11858-015-0675-6

Christou, K., & Prokopou, A. (2019). *Using refutational text to remedy the multiplication makes bigger misconception*. Aachen, Germany: Paper presented at EARLI bi-annual meeting.

Cyr, S., Charland, P., Riopel, M., & Bruyère, M.-H. (2019). Integrating a game design model in a serious video game for learning fractions in mathematics. *Journal of Computers in Mathematics and Science Teaching, 38*(1), 5–29.

Depaepe, F., Torbeyns, J., Vermeersch, N., Janssens, D., Janssen, R., Kelchtermans, G., … Van Dooren, W. (2015). Teachers' content and pedagogical content knowledge on rational numbers: A comparison of prospective elementary and lower secondary school teachers. *Teaching and Teacher Education, 47*, 82–92. https://doi.org/10.1016/j.tate.2014.12.009

Devlin, K. (2011). *Mathematics education for a new era: Video games as a medium for learning*. A K Peters.

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education, 16*(1), 3–17. https://doi.org/10.2307/748969

Gaggi, O., Ciraulo, F., & Casagrande, M. (2018). Eating pizza to learn fractions. In *In proceedings of the 4th EAI international conference on smart objects and Technologies for Social Good (Goodtechs '18)* (pp. 220–225). https://doi.org/10.1145/3284869.3284921

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. *Journal of Cross-Cultural Psychology, 42*(4), 543–565. https://doi.org/10.1177/0022022111406251

Guerrero, S., Walker, N., & Dugdale, S. (2004). Technology in support of middle grade mathematics: What have we learned? *The Journal of Computers in Mathematics and Science and Teaching, 23*(1), 5–20.

Habgood, M. P. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences, 20*(2), 169–206. https://doi.org/10.1080/10508406.2010.508029

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental Psychology, 53*(3), 587–596. https://doi.org/10.1037/dev0000252

Heck, R. H., Thomas, S. L. T., & Tabata, L. N. (2014). *Multilevel and longitudinal modeling with IBM SPSS*. Routledge.

Kainulainen, M., McMullen, J., & Lehtinen, E. (2017). Early developmental trajectories toward concepts of rational numbers. *Cognition and Instruction, 35*(1), 4–19. https://doi.org/10.1080/07370008.2016.1251287

Kärki, T., McMullen, J., Halme, H., Määttä, S., Lehtinen, E., & Hannula-Sormunen, M. M. (2021). Pelaamalla kohti joustavaa rationaalilukukäsitettä. [Towards adaptive rational number knowledge by playing games]. *Psykologia, 56*(6), 567–583.

Kärki, T., McMullen, J., & Lehtinen, E. (2021). Designing a game-based environment for enhancing rational number knowledge. *Nordic Studies in Mathematics Education, 26*(2), 25–46.

Kärki, T., McMullen, J., & Lehtinen, E. (2022). Improving rational number knowledge using the NanoRoboMath digital game. *Educational Studies in Mathematics, 110*(1), 101–123. https://doi.org/10.1007/s10649-021-10120-6

Ketamo, H., & Kiili, K. (2010). Conceptual change takes time: Game based learning cannot be only supplementary amusement. *Journal of Educational Multimedia & Hypermedia, 19*(4), 399–419.

Kieren, T. (1976). On the mathematical, cognitive, and instructional foundations of rational numbers. In R. Lesh, & D. Bradbard (Eds.), *Number and measurement. Papers from a research workshop* (pp. 101–144). ERIC Information Analysis Center for Science, Mathematics, and Environmental Education.

Kiili, K., Moeller, K., & Ninaus, M. (2018). Evaluating the effectiveness of a game-based rational number training—In-game metrics as learning indicators. *Computers and Education, 120*, 13–28. https://doi.org/10.1016/j.compedu.2018.01.012

Kiili, K. J. M., Lindstedt, A., Koskinen, A., Halme, H., Ninaus, M., & McMullen, J. (2021). Flow experience and situational interest in game-based learning: Cousins or identical twins. *International Journal of Serious Games, 8(3), Article 3*. https://doi.org/10.17083/ijsg.v8i3.462

Koops, M., & Hoevenaar, M. (2013). Conceptual change during a serious game: Using a lemniscate model to compare strategies in a physics game. *Simulation and Gaming, 44*(4), 544–561. https://doi.org/10.1177/1046878112459261

Laato, S., Lindberg, R., Laine, T. H., Bui, P., Brezovszky, B., Koivunen, L., … Lehtinen, E. (2020). *Evaluation of the pedagogical quality of mobile math games in app marketplaces, 2020 IEEE international conference on engineering, technology and innovation (ICE/ITMC), 1–8*. https://doi.org/10.1109/ICE/ITMC49519.2020.9198621

Lee, I. A., & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org.

Lee, J.-E., Chan, J. Y.-C., Botelho, A., & Ottmar, E. (2022). Does slow and steady win the race?: Clustering patterns of students' behaviors in an interactive online mathematics game. *Educational Technology Research and Development, 70*(5), 1575–1599. https://doi.org/10.1007/s11423-022-10138-4

Lehtinen, E., Hannula-Sormunen, M. M., McMullen, J., & Gruber, H. (2017). Cultivating mathematical skills: From drill-and-practice to deliberate practice. *ZDM Mathematics Education, 49*(4), 625–636. https://doi.org/10.1007/s11858-017-0856-6

Leonard, M. J., Kalinowski, S. T., & Andrews, T. C. (2014). Misconceptions yesterday, today, and tomorrow. *CBE Life Sciences Education, 13*(2), 179–186. https://doi.org/10.1187/cbe.13-12-0244

Li, X., Wang, W., & Li, Y. (2022). Systematically reviewing the potential of scientific argumentation to promote multidimensional conceptual change in science education. *International Journal of Science Education, 44*(7), 1165–1185. https://doi.org/10.1080/09500693.2022.2070787

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: a critical appraisal. *Learning and Instruction, 11*(4–5), 357–380. https://doi.org/10.1016/S0959-4752(00)00037-2

Martin, T., Smith, C. P., Forsgren, N., Aghababyan, A., Janisiewicz, P., & Stephanie, B. (2015). Learning fractions by splitting: Using learning analytics to illuminate the development of mathematical understanding. *Journal of the Learning Sciences, 24*(4), 593–637. https://doi.org/10.1080/10508406.2015.1078244

Masek, M., Boston, J., Lam, C. P., & Corcoran, S. (2017). Improving mastery of fractions by blending video games into the math classroom. *Journal of Computer Assisted Learning, 33*, 486–499. https://doi.org/10.1111/jcal.12194

McMullen, J., Brezovszky, B., Hannula-Sormunen, M. M., Veermans, K., Rodríguez-Aflecht, G., Pongsakdi, N., & Lehtinen, E. (2017). Adaptive number knowledge and its relation to arithmetic and pre-algebra knowledge. *Learning and Instruction, 49*, 178–187. https://doi.org/10.1016/j.learninstruc.2017.02.001

McMullen, J., Hannula-Sormunen, M. M., Lehtinen, E., & Siegler, R. S. (2020). Distinguishing adaptive from routine expertise with rational number arithmetic. *Learning and Instruction, 68*, Article 101347. https://doi.org/10.1016/j.learninstruc.2020.101347

McMullen, J., Koskinen, A., Kärki, T., Lindstedt, A., Määttä, S., Halme, H., … Kiili, K. (2023). A game-based approach to promoting adaptive rational number knowledge. *Mathematical Thinking and Learning, 26*(4), 411–427. https://doi.org/10.1080/10986065.2023.2177818

McMullen, J., Van Hoof, J., Degrande, T., Verschaffel, L., & Van Dooren, W. (2018). Profiles of rational number knowledge in Finnish and Flemish students – A multigroup latent class analysis. *Learning and Individual Differences, 66*, 70–77. https://doi.org/10.1016/j.lindif.2018.02.005

Mikkilä-Erdmann, M. (2001). Improving conceptual change concerning photosynthesis through text design. *Learning and Instruction, 11*(3), 241–257. https://doi.org/10.1016/S0959-4752(00)00041-4

Nejem, K. M., & Muhanna, W. (2013). The effect of using computer games in teaching mathematics on developing the number sense of fourth grade students. *Educational Research and Reviews, 8*(16), 1477–1482.

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

Ninaus, M., Kiili, K., McMullen, J., & Moeller, K. (2017). Assessing fraction knowledge by a digital game. *Computers in Human Behavior, 70*, 197–206. https://doi.org/10.1016/j.chb.2017.01.004

Obersteiner, A., Hoof, J. V., Verschaffel, L., & Dooren, W. V. (2016). Who can escape the natural number bias in rational number tasks? A study involving students and experts. *British Journal of Psychology, 107*(3), 537–555. https://doi.org/10.1111/bjop.12161

Riconscente, M. M. (2013). Results from a controlled study of the iPad fractions game Motion Math. *Games and Culture, 8*(4), 186–214. https://doi.org/10.1177/1555412013496894

Schneider, M., & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology, 49*(9), 1639–1649. https://doi.org/10.1037/a0030976

Schoultz, J., Säljö, R., & Wyndhamn, J. (2001). Heavenly talk: Discourse, artifacts, and children's understanding of elementary astronomy. *Human Development, 44*(2–3), 103–118. https://doi.org/10.1159/000057050

Schubert, E. (2023). Stop using the elbow criterion for k-means and how to choose the number of clusters instead. *SIGKDD Explorations Newsletter, 25*(1), 36–42. https://doi.org/10.1145/3606274.3606278

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science, 3*, 341–361. https://doi.org/10.1111/desc.12395

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. https://doi.org/10.1177/0956797612440101

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*(3), 545–560. https://doi.org/10.1037/a0014239

Smith III, J. P., diSessa, A.a., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences, 3*(2), 115–163. https://doi.org/10.1207/s15327809jls0302_1

Stafylidou, S., & Vosniadou, S. (2004). The development of students' understanding of the numerical value of fractions. *Learning and Instruction, 14*(5), 503–518. https://doi.org/10.1016/j.learninstruc.2004.06.015

Sun, L., Kangas, M., Ruokamo, H., & Siklander, S. (2023). A systematic literature review of teacher scaffolding in game-based learning in primary education. *Educational Research Review, 40*, Article 100546. https://doi.org/10.1016/j.edurev.2023.100546

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education, 8*(6), 951–970. https://doi.org/10.1007/s10763-010-9203-x

Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2013). Educated adults are still affected by intuitions about the effect of arithmetical operations: Evidence from a reaction-time study. *Educational Studies in Mathematics, 82*(2), 323–330.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*(5), 453–467. https://doi.org/10.1016/j.learninstruc.2004.06.013

Vamvakoussi, X., & Vosniadou, S. (2012). Bridging the gap between the dense and the discrete: The number line and the "rubber line" bridging analogy. *Mathematical Thinking and Learning, 14*(4), 265–284. https://doi.org/10.1080/10986065.2012.717378

Van Dooren, W., Christou, K., Depaepe, F., Inglis, M., Määttä, S., McMullen, J., Obersteiner, A., Heck Ribeiras, P., Van Hoof, J., Triandafyllou, M., Vamvakoussi, X., Verschaffel, L., Wittmann, G., & Woollacott, B. (2019). *Tackling the natural number bias – A comparative textbook analysis*. Aachen, Germany: Paper presented at EARLI bi-annual meeting.

Van Dooren, W., Lehtinen, E., & Verschaffel, L. (2015). Unraveling the gap between natural and rational numbers. *Learning and Instruction, 37*, 1–4. https://doi.org/10.1016/j.learninstruc.2015.01.001

Van Hoof, J., Degrande, T., Ceulemans, E., Verschaffel, L., & Van Dooren, W. (2018). Towards a mathematically more correct understanding of rational numbers: A longitudinal study with upper elementary school learners. *Learning and Individual Differences, 61*, 99–108. https://doi.org/10.1016/j.lindif.2017.11.010

Van Hoof, J., Janssen, R., Verschaffel, L., & Van Dooren, W. (2015). Inhibiting natural knowledge in fourth graders: Towards a comprehensive test instrument. *ZDM, 47*(5), 849–857. https://doi.org/10.1007/s11858-014-0650-7

Van Hoof, J., Vandewalle, J., Verschaffel, L., & Van Dooren, W. (2015). In search for the natural number bias in secondary school students' interpretation of the effect of arithmetical operations. *Learning and Instruction, 37*, 30–38. https://doi.org/10.1016/j.learninstruc.2014.03.004

Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology, 11*(6), 645–661. https://doi.org/10.1080/17405629.2014.921153

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3–34). Lawrence Erlbaum.

Vosniadou, S., & Verschaffel, L. (2004). Extending the conceptual change approach to mathematics learning and teaching. *Learning and Instruction, 14*(5), 445–451. https://doi.org/10.1016/J.LEARNINSTRUC.2004.06.014

West, B. T. (2009). Analysing longitudinal data with the linear mixed models procedure in SPSS. *Evaluation & the Health Professions, 32*(3), 207–228. https://doi.org/10.1177/0163278709338554

Zhang, L., Shang, J., Pelton, T., & Pelton, L. F. (2020). Supporting primary students' learning of fraction conceptual knowledge through digital games. *Journal of Computer Assisted Learning, 36*(4), 540–548. https://doi.org/10.1111/jcal.12422