International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

Mathematical game performance as an indicator of deliberate practice

Jake McMullen[1], Phuong Bui[1], Boglárka Brezovszky[1], Erno Lehtinen [1,2]
and Minna M. Hannula-Sormunen[1]

[1]Department of Teacher Education, University of Turku, Turku, Finland; [2] Vytautas Magnus University, Kaunas, Lithuania
{jamcmu,phuong.bui, ernoleh, mimarha}@utu.fi; {brezovszkyboglarka}@gmail.com}

Keywords:
Adaptive number knowledge
Gaming analytics
Game-based learning
Deliberate practice
Mathematics

Received: May 2023
Accepted: November 2023
Published: November
DOI: 10.17083/ijsg.v10i4.634

Abstract

The traditional classroom setting presents challenges when it comes to strengthening adaptive expertise in mathematics education through deliberate practice. This study aimed to investigate whether the Number Navigation Game (NNG) could help promote deliberate practice and whether students' performance in the game was related to their development of Adaptive Number Knowledge, perceived challenge, flow, and math interest. NNG is a game-based learning environment that requires students to progress by solving increasingly complex arithmetic problems, which is crucial for promoting adaptive number knowledge. Game performances of 214 Finnish students were analyzed and compared to the best possible performance for each game level. A growth mixture model based on the students' relative performance levels was used to gain insight into how students' game performance changed throughout the game, and how this related to their knowledge gains, perceived challenge, math motivation, and flow. There were four different profiles of students' game performance. The largest profile consisted of students who steadily improved their performance in the game, despite initially having lower-than-average performance. This group experienced lower levels of flow but achieved larger learning gains than the other groups, suggesting that their engagement may be more aligned with deliberate practice.

1. Introduction

Learning environments that support adaptive, rather than routine, expertise, are rare [1]. This is especially true in typical comprehensive school settings [2]. In contrast with routine expertise, adaptive expertise is typified by an ability to remix and flexibly apply disparate aspects of knowledge in solving atypical tasks[3], [4]; in other words, highly transferable knowledge [5]. The problem of non-transferable knowledge has been documented in school domains from science to mathematics and beyond [6], [7]. One issue is that much of schooling is designed to support more routine expertise, which is typified by sparsely connected knowledge that is unmalleable and only easily applied to well-rehearsed tasks [8], [9].

Game-based learning environments have been shown to be effective in the promotion of school children's knowledge and skills with many topics. However, much of the skills and knowledge that game-based learning has aimed to support aligns more closely with routine expertise [10], [11]. Learning environments that are hypothesized to promote more routine expertise are described by well-structured tasks, which are procedurally oriented, with little meaning-making required [8], [10]. While such routine skills and knowledge are valuable for basic school-like tasks, they run the risk of not being functional in everyday life and future learning [6], [12].

Recent approaches to supporting more adaptive expertise using game-based learning environments have proven promising [10]–[13]. These learning environments are designed with principles that are proposed to support adaptive expertise including less-structured, conceptually rich tasks that allow for meaning-making [8], [10]. In particular, in the domain of mathematics, game-based learning environments that move beyond the typical drill-and-practice approach of a wide swath of digital (game-based) learning environments have shown positive gains in knowledge and skills more aligned with adaptive expertise [10], [12]. One feature of many of these learning environments is the opportunity for players to engage with the game at multiple levels of inquiry [12]. That is, players can take a more surface approach to solving the tasks and progress through the game, whereas others may seek more complexity in their gaming and subsequently have a higher level of performance according to the game mechanics. Given the diverse possibilities for engagement in these games, some students may be more contentious of their performance and aim to engage at these more complex levels. These players may be engaged in a form of deliberate practice, an effortful form of practice that aims to improve skills at the edge of one's competence [14]. Thus, deliberate practice may be valuable for developing adaptive expertise in those learning environments that are designed following adaptive expertise principles. However, to the best of our knowledge, there are few existing examinations of individual differences in how players engage in game-based learning environments that aim to support adaptive expertise.

1.1 Adaptive number knowledge

Adaptive expertise with arithmetic is a highly valued outcome of primary school curricula [15], [16]. Traditionally adaptive expertise with arithmetic has been shown to manifest in students' strategic flexibility, such as being able to employ an advantageous shortcut method to solve a mental arithmetic problem [17]. More recently adaptive number knowledge has been identified as a behavioral manifestation of adaptive expertise in arithmetic, which reflects a well-connected network of knowledge of numerical characteristics and arithmetic relations [18]. Adaptive number knowledge was found to vary even among students with high levels of other aspects of arithmetic knowledge and is a unique predictor of later algebraic knowledge [19]. To support adaptive number knowledge, students must be provided with opportunities for extensive practice solving tasks in which numbers and operations are connected in novel ways. [12]. However, it is usually difficult to achieve such conditions in the larger context of a traditional classroom for all students.

In fact, practice in math classrooms is often connected to the automatization of procedural skills [6] and static routine expertise development [5]. There is a lack of pedagogical models for supporting the development of adaptive expertise in school mathematics[6]. For example, mechanical practice with procedures of rational number arithmetic alone does not necessarily lead to the development of adaptive expertise in all students [20]. In other words, not all practice is equally beneficial. With the affordances of game-based learning, it is possible to overcome these limitations and provide opportunities for the deliberate practice of tasks that should support adaptive expertise [10]. Game mechanics can align with deliberate practice principles while offering an open-ended learning environment that would trigger reflection on different solutions to arithmetic problems [6] and possibly lead to improvements in adaptive number knowledge. However, providing opportunities to engage in deliberate practice does not guarantee all students will do so.

1.2 Deliberate practice in math education

There is a need for more complex forms of practice in mathematics classrooms, especially deliberate practice [6], [14]. Deliberate practice has been argued to be a key factor in explaining extraordinary development in different domains [21]. In the realm of mathematics education, it is especially important to provide students with opportunities to engage in more complex forms of practice that push them to develop their emerging skills and knowledge structures rather than just routine and static practice with their existing skills [6]. Engaging in deliberate practice should support the automatization of progressively more complex activities, freeing cognitive resources, through processes such as chunking, to support even more advanced performance on increasingly complex tasks [14]. Thus, well-designed learning environments that support adaptive expertise should align with principles of deliberate practice.

Due to the constraints of conventional educational settings and the demanding conditions of deliberate practice principles, it is difficult to systematically apply deliberate practice in many math classrooms [6]. However, previous research suggests that it is possible to apply some core aspects of deliberate practice in well-designed game-based learning environments [12]. In the Number Navigation Game, design principles align with some core principles of deliberate practice [12], [22]. For example, players are provided with tasks that are (a) ideally challenging and well-suited to their level of knowledge (e.g., there are multiple avenues for solving any task), (b) offer opportunities to practice their skills at the edge of their competence (i.e. practice with different levels of difficulty embedded in a single task), (c) providing them with continuous feedback to improve their current skills (i.e. game mechanics and the mathematics are one and the same), and (d) the tasks are situated in an open-ended learning environment that can trigger reflection on different solutions to arithmetic problems (i.e. opportunities for improved scores through deeper mathematical engagement) [23]. Previous studies reveal that the NNG was able to support the development of older students' adaptive number knowledge at the same time it supported younger students' basic arithmetic fluency development [12]. In short, the NNG was able to meet students at multiple levels of challenge ideally aligned with their prior knowledge, a core feature of deliberate practice design principles.

Promoting students' adaptive number knowledge is challenging and often demands a great quantity of variable practice [19]. This practice has to be complex, with varying numbers and operations, while offering multiple strategies for solving arithmetic problems, and opportunities for students to contemplate their solutions and the underlying arithmetic relations [3]. Teaching in many classrooms usually involves teaching various problem-solving strategies, which does not lead to strong adaptive number knowledge in most students [5]. Ideal training environments need to incorporate numerous attempts to strengthen students' underlying knowledge of numerical relations while offering a large amount of practice with open-ended arithmetic problems [6]. The NNG aims at improving primary school students' adaptive number knowledge by providing different opportunities to engage in strategic work with various arithmetic combinations and operations with the use of the 100-square as the external representation of whole numbers (1-100) [23]. This is made possible through various design choices (see Section 2.2 for a full game description), especially by including so-called Energy maps. In Energy maps, students must minimize the numerical inputs they use in their arithmetic solutions to navigate around the 100-square. High-level performance on Energy maps requires using complex arithmetic relations and strategic thinking [12]. Thus, these maps are expected to be crucial for the game in supporting adaptive expertise. However, it is unclear exactly if and how students engaged in deliberate practice while playing the game, nor how that influenced their learning outcomes and perception of the game.

Research on gaming has emphasized the importance of "flow" as a desirable experience during gameplay [24], which happens when one directs all of their attention to the gameplay [25]. Both flow and deliberate practice refer to strong engagement with a task [21], [25]. However, there are fundamental differences between these phenomena. When developing the flow concept, Csikszentmihalyi [25] analyzed the experiences of high-level experts (e.g. chess players and rock climbers) while they pursued enjoyable activities at an optimal level of challenge. During a flow state, the "self" often seems to disappear from one's awareness [26] and time is often distorted as one's experience can pass really fast or feel really slow. In contrast, Ericsson et al. [14] showed that the type of deliberate practice that is needed for enhancing expertise is a highly conscious, and sometimes unpleasant, practice. During deliberate practice, individuals aim to go beyond their current skill level. In contrast to being in the flow state, one needs to be fully aware and present during such deliberate practice. As such, progress is often slow, but persistent [27]. Thus, we can assume that in a mathematics serious game, deliberate practice may be related to persistent but slow progress, low flow experience, and a high feeling of challenge.

2. Present Study and Methods

2.1 Present study

Previously, NNG was found to be effective in promoting adaptive number knowledge. As well, students' game performance, as measured by the number of maps completed, was shown to predict their learning gains [12]. This suggests that playing the game was directly related to the learning outcomes of students. In principle, the design of NNG has been argued to offer opportunities for students to engage in deliberate practice [23]. However, there is not any direct evidence of students engaging in deliberate practice while playing the game. Thus, in the present study, we aim to better understand the different ways that students engage in playing the NNG. In doing so, we examine if there is evidence of gameplay behaviors that suggest deliberate practice-like activities. Thus, we ask:

(1) What are the main patterns of game performance?

Previous evidence suggests that there are important individual differences in students' performance in the NNG, which predict their learning outcomes [12]. Using a growth mixture model [28], we aim to examine if there are individual differences in the patterns of game performance across the different maps of the game. We expect that there are substantial and relevant individual differences in students' game performance while playing the Number Navigation Game. However, it is not clear what the particular profiles of performance may look like. Thus, we consider this analysis to be exploratory in nature.

(2) How is game performance related to prior mathematical knowledge and math interest?

In order to better understand and validate the resulting profiles, we will examine how profile membership is related to prior mathematical knowledge and motivation. Prior evidence suggests that students' initial levels of knowledge and motivation impact their in-game performance [12], [29]. Thus, we expect that higher prior knowledge and math interest would be related to more advanced initial performance within the game. Likewise, lower prior knowledge and mathematical interest should predict lower initial performance in the game.

(3) How is game performance related to mathematics learning outcomes?

As per previous research [12], we expect that better performance within the game would also support larger learning gains from pre- to post-test.

(4) How is game performance related to experiences of flow and challenge in the game?

Presumably, flow experiences should be higher for those who experience an appropriate level of challenge [26]. However, it is not clear how this relation will manifest with the profiles of gaming performance. Thus, these analyses will be treated as exploratory.

2.2 Method

2.2.1 Participants

Participants were 214 fourth to sixth-grade students (M age= 11.37, SD= 7.13) selected from a larger sample of 642 students who took part in a large-scale randomized control trial over ten weeks [12]. A pre-test, carried out by trained research assistants, was followed by the intervention. The students played the NNG during their normal mathematics lessons as dictated by their teacher. Teachers were asked to have their students play the game for around 10 hours total, a few times a week, with sessions lasting at least 30 minutes. The intervention period was followed by a post-test. In order to capture sufficient developmental changes in gameplay, we included in the present study only those participants who finished at least 5 Energy maps (see 2.2. for description). Participation was voluntary and informed consent forms were gathered from the participants' parents. Ethical guidelines of the University of Turku were followed strictly.

2.2.2 Description of the Number Navigation Game

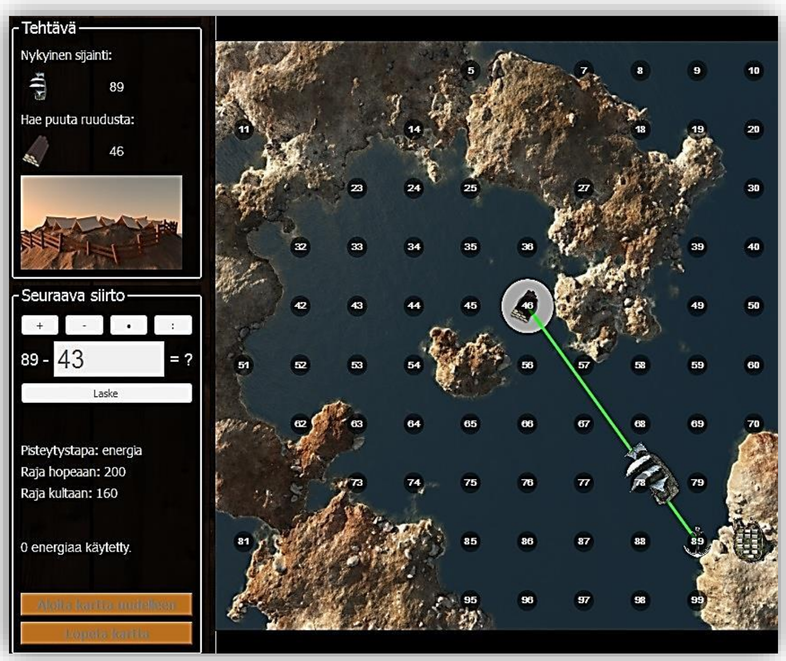

Figure 1. Example of a map in the Number Navigation Game (version 1) in the Energy scoring mode.

NNG has an *intrinsically integrated* design [30] in which the gaming mechanism is integrated directly into the mathematical content. The gaming interface is a 100-square superimposed on different maps of land and sea. In the game, the players' task is to collect raw materials to build settlements. To progress, players need to navigate a ship from a starting number (the harbor) to retrieve material from a given point and return it to the harbor. Four sets of materials should be gathered to complete a map. Players move about the map using arithmetic operations to navigate between different numbers on the map. For instance, in Fig. 1, the player starts from number 89 and has to collect wood located at number 46. To move, the player needs to input a mathematical equation in the calculation box on the left side of the screen. The moves have to take the ship to the targeted material (number 46) while avoiding the land masses. A map is completed when all four of the materials are retrieved.

The game has two scoring modes: Moves and Energy scoring. In the Moves mode, players need to retrieve the materials and return to the harbor using the least number of moves (operations). In this mode, the main aim was to have players use more complex addition and subtraction, in contrast to simply moving piecemeal using 1 and 10 (e.g., subtracting 10 four consecutive times, then 1 three times, to move from 89 to 46).

In the Energy scoring mode, players use up energy points from the numbers they input into the calculation box. The idea is to use the least amount of "energy" possible to achieve better scores. A high level of performance in the Energy mode requires more complex arithmetic relations and players need to use all four arithmetic operations, including multiplication and division. Hence, the Energy maps are considered to be particularly important for promoting adaptive number knowledge. For more information about the developmental process of NNG and the role of game features in enhancing adaptive number knowledge see [23], [31].

2.2.3 Pre and Post-test Measures

Before and after the intervention students completed measures of adaptive number knowledge, arithmetic procedural fluency (Woodcock-Johnson Math Fluency sub-test), and math interest. The Post-test also included a questionnaire measuring core dimensions of gaming experience including perceived Challenge and Flow.

Adaptive Number Knowledge was measured with the Arithmetic Sentence Production Task [18]. The task is a timed, paper-and-pencil instrument that measures students' ability to recognize and use different numerical characteristics and relations during their problem-solving. An item includes four to five given numbers (e.g., 2, 4, 8, 12, 32) and the four basic arithmetic operations, and the aim is for students to produce as many arithmetic sentences that equal a target number (e.g. 16) as they can in 90 seconds. The task included 4 items at the Pre- and Post-test and Cronbach's α reliability value for the total number of correct solutions across the four items was .70 at the Pre-test.

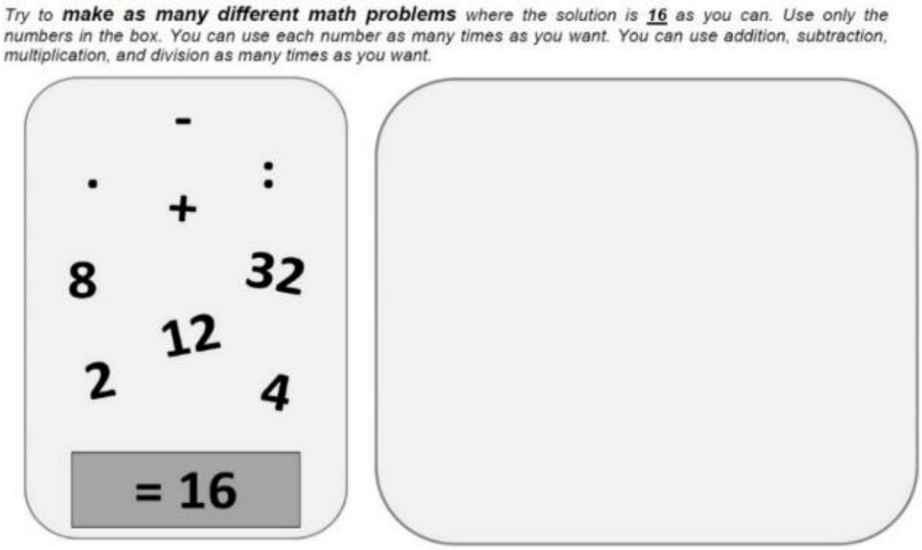

Figure 2. Example of an Arithmetic Production Task

Arithmetic Procedural Fluency was measured by the Woodcock-Johnson Math Fluency sub-test [32] which includes a total of 160 items. Students were asked to complete as many arithmetic problems as possible in 3 minutes.

Math Interest was measured as part of the Expectancy Value Math Motivation questionnaire [33] with three items (e.g. " I like Math"). *Perceived Challenge* and *Flow* were measured as part of the Gaming Experience Questionnaire which aims to measure eight core dimensions of game experience [34]. The questionnaire was translated into Finnish [35] and simplified in language and in length to be more suitable for the age of the participants (see [29]).

2.2.4 In-game Measures

Game performance was measured with a Relative energy score. This was a map-neutral measure of performance that was calculated by comparing the individual's score with the threshold for a Gold medal for the map. Players complete a map by retrieving all four materials, depending on how optimal their arithmetic solutions are, their performances (or scores) will earn them either Bronze, Silver, or Gold coins, in which Gold coins solutions are considered close-to-optimal solutions. Players completed a different number of Energy maps, although those included in the present study completed at least five. Thus, in order to ensure comparability, players' scores for the five measurement points used in the study were calculated by taking their performance on the first two maps they completed as T1 and T2, the last two maps as T4 and T5, and the average of the intervening maps as T3. The total amount of maps completed and Time Spent (in minutes) on completing these maps are also included as a point of comparison between different profiles.

2.2.5 Analysis

In order to capture patterns of performance across the Energy maps, a series of growth mixture models were estimated using Mplus version 8.4. Growth mixture models are a clustering procedure that identifies coherent profiles based on individuals' estimated growth curves. Individuals' initial scores of Energy map performance and their linear and quadradic slopes were estimated and used as indicators for defining a categorical variable of profile membership (e.g., a cluster membership). We set the number of profiles in the model starting with one profile and subsequently increasing the number of profiles by one. Each subsequent model is compared using a set of statistical indicators [36] and theoretical considerations [37]. Once the most appropriate number of profiles is determined, we used a 3-step approach [38] to examine the relation between profile membership and external variables (i.e. adaptive number knowledge, math interest, etc.). The 3-step approach first estimates the profile structures of the model, assigns individuals probabilities of being members in each profile, and then examines the relation between the external variable and profile membership.

3. Results

3.1 Profiles of game performance

Table 1 details the statistical indicators of the estimated growth mixture models. Based on these, we considered the four-profile growth mixture model to be the most appropriate. This was based on it having the lowest BIC value of the estimated models and a significant BLRT in comparison with the three-profile model, which the 5-profile model did not have a significant BLRT [36]. As well, the additional profile introduced in the 5-profile model was extremely small, around 1% of the sample, and theoretically not well-founded.

Table 1. Statistical indicators for 2 to 7 profile models

Number of Profiles	AIC	BIC	Entropy	LMR-RT	BLRT
2	2472	2529	.97	<.001	<.001
3	2401	2471	.97	.01	<.001
4	2385	2469	.88	.45	<.001
5	2375	2472	.96	.99	.11
6	2362	2473	.92	.99	.99
7	2370	2494	.93	.99	.99

Figure 3 displays the mean scores and standard errors for each profile across the five game maps. As can be seen, there are substantial differences between all profiles, except the two generally advanced performing groups. The Low Stable Performers profile (68 students) had lower than average performance on all energy maps (Mean initial level = -0.64, SE = .04; Slope = -.12, SE = .08; quadratic = 0.01, SE = .02). In general, students in this profile appear to mainly play the game without attaining high scores in the energy maps. The Improving Performers (72 students) profile had lower than average initial performance but improved their performance on every Energy map (Mean initial map = -0.69, SE = .02; Slope = 0.51, SE = .12; Quadratic = -0.04, SE = .03). These students were the only group who appeared to have substantial changes in their performance across the game. In particular, this group appears to make substantial gains in their performance levels as they progress in the game. Given the systemic positive changes in performance, we argue these students may have been engaged in practice behaviors that have similarities with deliberate practice. The High Stable Performers profile (38 students) had higher than average and stable performance on all energy maps (Mean initial map = 0.81, SE = .09; slope = -0.01, SE = .15; quadratic = -0.03, SE = .04). They generally perform well across all maps in the game with little change in how they perform in the energy maps. The High Not-Maintained Performers profile (35 students) initially had very strong performance, with an immediate drop to stable above average; it was the only profile with a non-linear slope (Mean initial map = 1.80, SE = .08; slope = -0.78, SE = .14; Quadratic = 0.11, SE = .04). This group appeared to perform extremely well on the first Energy map they completed, but then regressed to more typical levels of high performance, akin to the High Stable Performers profile, on subsequent maps. The significant quadradic term of the profile suggests this inconsistency in changes in performance between the first two maps and the rest of the maps.

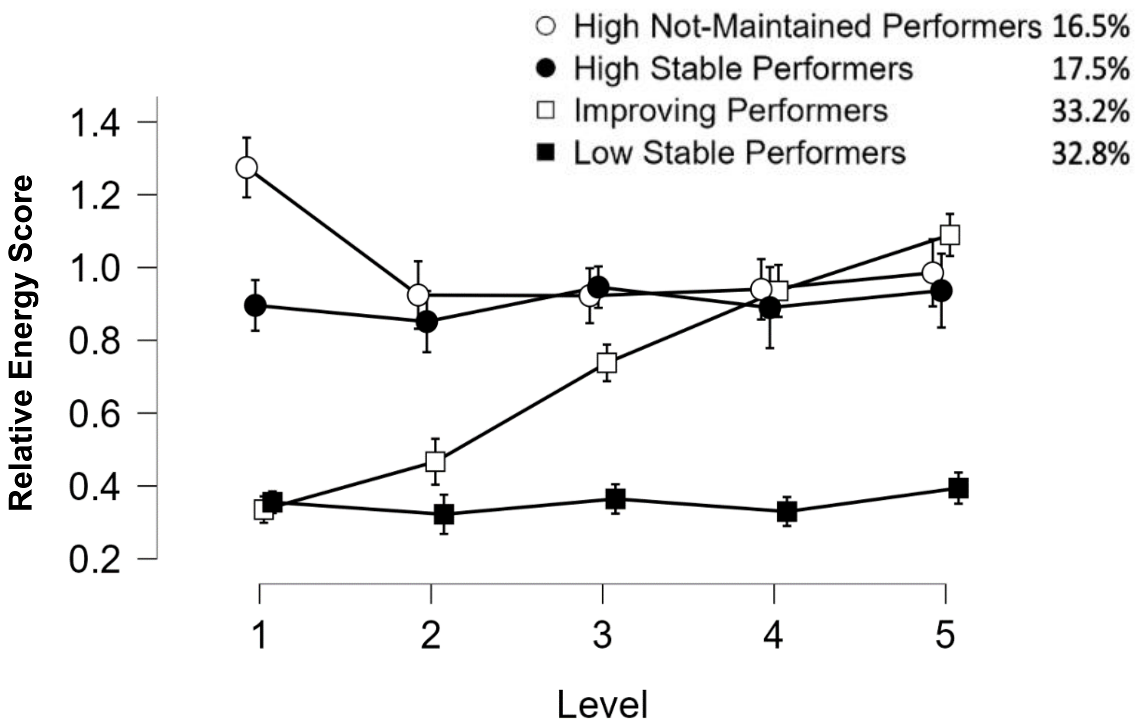

Figure 3. Mean scores of relative game performance scores by growth mixture model profile for each game map. Error bars represent 95% confidence intervals.

In order to examine the validity of the profiles identified in the growth mixture model, we examined the relation between profile membership and pre-test measures using the 3-step approach. First, we examined if there were significant differences in prior mathematical knowledge, by examining if there were differences in adaptive number knowledge between the groups. Results of the 3-step comparison revealed that Low Stable Performers had lower adaptive number knowledge at the pretest than Improving Performers (B = .46; SE = .14; p = .001), High Stable Performers (B = .35; SE = .15; p = .02), and High Not-Maintained Performers (B = .51; SE = .16; p = .001) profiles. None of the other three profiles differed from each other. These results suggest that the Low Stable Performers had the weakest pre-test adaptive number knowledge. Interestingly, there were no differences between the Improving Performers and the two High Performers groups, suggesting these students came into playing the game with fairly high levels of adaptive number knowledge.

As well, we examined if there were significant differences in prior mathematical knowledge by examining if there were differences in arithmetic fluency between the groups at the pre-test. Results of the 3-step comparison revealed that Low Stable Performers had lower adaptive number knowledge at the pretest than Improving Performers (B = .038; SE = .015; p = .003), High Stable Performers (B = .05; SE = .016; p = .001), and High Not-Maintained Performers (B = .037; SE = .015; p = .013) profiles. None of the other three profiles differed from each other. These results suggest that the Low Stable Performers had the weakest pre-test arithmetic fluency. Again, there were no differences between the Improving Performers and the two High Performers groups.

Table 2. Means, standard deviation (in parentheses) of pre and post-test of adaptive number knowledge (ANK) and Woodcock-Johnson arithmetic fluency (AF), pre-test math interest, post-test of gaming experience, time spent on tasks and maps completed.

	Low Stable Performers	Improving Performers	High Stable Performers	High Not-Maintained Performers
Pretest ANK	3.34 (1.31)	4.24 (1.84)	4.0 (1.47)	4.44 (1.69)
Posttest ANK	3.89 (1.68)	5.2 (2.07)	4.9 (1.44)	5.52 (2.36)
Pretest Arith Flu	68.81 (17.0)	77.22 (17.04)	80.66 (16.56)	77.76 (14.7)
Posttest Arith Flu	79.74 (15.67)	89.53 (18.32)	88.72 (16.91)	88.29 (19.61)
Math interest	2.90 (1.0)	3.4 (0.87)	3.46 (0.96)	3.3 (0.99)
Flow	2.06 (0.93)	1.94 (0.7)	2.41 (0.98)	2.17 (0.9)
Perceived challenge	2.43 (0.7)	2.35 (0.66)	2.72 (0.80)	2.45 (0.68)
Time on tasks	389.3 (142.64)	464.93 (186.93)	419.92 (149.96)	393.09 (117.07)
Maps completed	31.78 (10.08)	36.14 (10.38)	36.74 (14.61)	33.83 (9.97)

Finally, we examined, if there were between profile differences in prior mathematical motivation, by examining group differences in Math Interest measured at the pre-test. Results of the 3-step comparison revealed that the Low Stable Performers profile had lower math interest than the Improving Performers (beta = .68; SE = .24; p = .004)and the High Stable Performers (beta = .68; SE = .27; p = .01) profiles, but not the High Not-Maintained profile. None of the other three profiles differed in their initial Math Interest. This suggests that the Low Performers have lower prior motivation for mathematics in general. However, the Improving Performers appear more similar to the two higher-performing groups.

3.2 Learning outcomes

In order to examine the relation between game performance profiles and learning outcomes from the game, we employed the 3-step approach to examine if there were differences in the four profiles on their post-test scores for adaptive number knowledge and arithmetic fluency after controlling for pre-test knowledge.

Based on the multinomial logistic regressions from the 3-step procedure, there were differences between profiles in adaptive number knowledge at the post-test, after controlling for pre-test adaptive number knowledge (See Figure 4 for estimated marginal means). Students in the Improving Performers profile were shown to have higher adaptive number knowledge at the post-test compared to the Low Stable Performers profile, after controlling for pre-test scores (beta = -0.48, SE = 0.22; Odds ratio = 1.58; p = .03). As well, students in the Low Stable Performers profile had lower posttest adaptive number knowledge, after controlling for pretest, than the High Stable Performers (beta = .46; SE = .22; p = .03) and High Not-Maintained Performers profiles (beta = .55, SE = .23, p = .02) . There were no other differences in post-test adaptive number knowledge between the other profiles. These results suggest that game performance was related to learning gains in adaptive number knowledge.

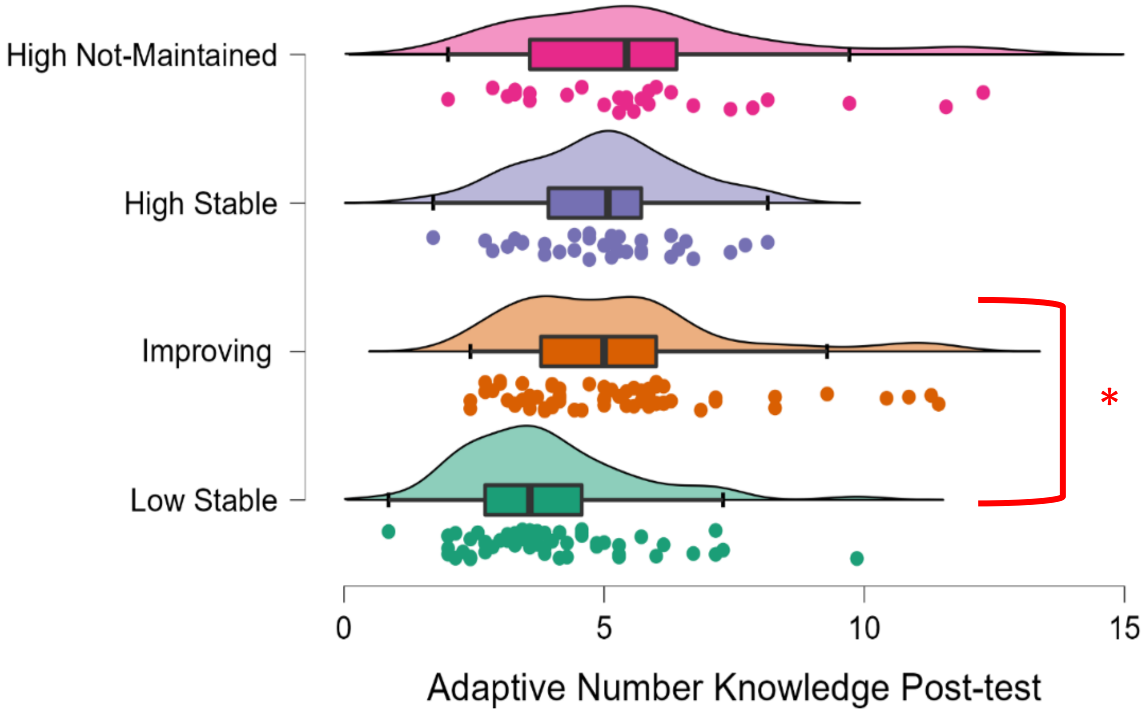

Figure 4. Estimated marginal means for Adaptive Number Knowledge at post-test, after controlling for pre-test scores. Note: * $p < .05$

Based on the multinomial logistic regressions from the 3-step procedure, there were no differences between profiles in arithmetic fluency at the post-test, after controlling for pre-test arithmetic fluency (See Figure 5 for estimated marginal means).

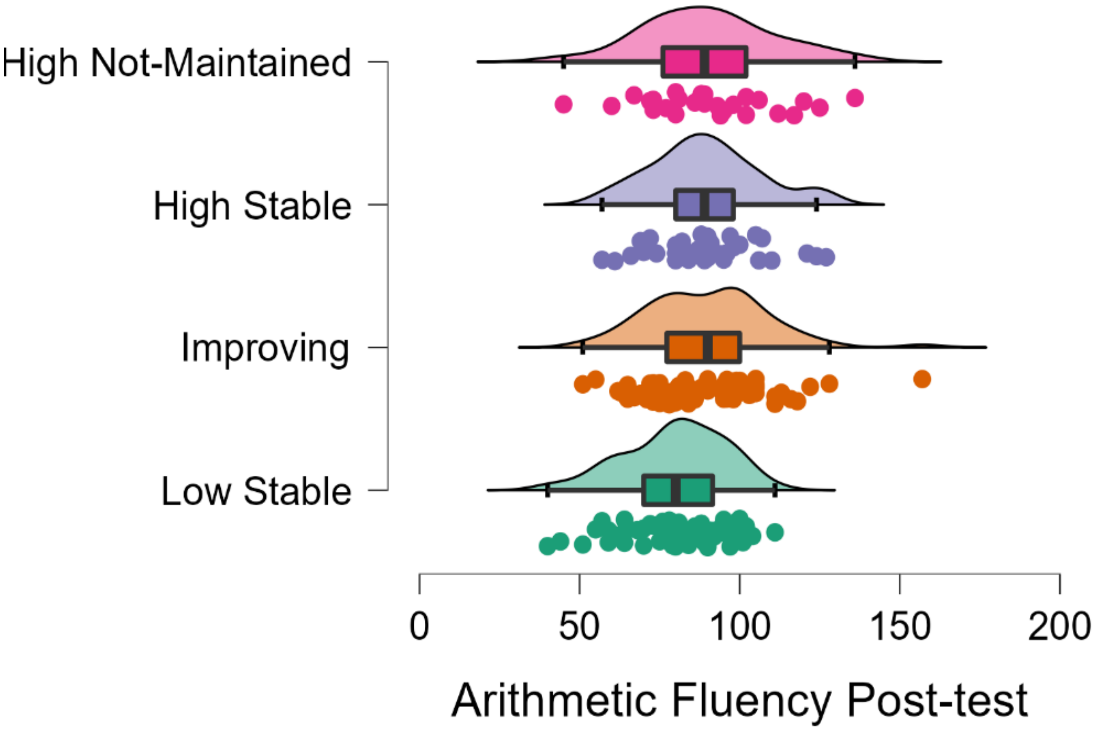

Figure 5. Estimated marginal means for Arithmetic Fluency at post-test, after controlling for pre-test scores.

3.3 Game experience and game play

In order to examine the relation between game performance profiles and game experiences, we employed the 3-step approach to examine if there were differences in the four profiles on their perceived flow experiences and challenge.

Based on the multinomial logistic regressions from the 3-step procedure, there were differences between profiles in perceived flow experience (See Figure 6a). Students in the Improving Performers profile were shown to have lower perceived flow experience than students in the High Stable profile (beta = .63; SE = .25; p = .01). There were no other differences in perceived flow experience. These results suggest that those students who seemed to improve the most in their performance had lower flow experiences than those whose performance was high, but stable.

Based on the multinomial logistic regressions from the 3-step procedure, there were also differences between profiles in perceived challenge (See Figure 6b). Students in the Improving Performers profile were shown to have lower perceived challenge than students in the High Stable profile (beta = .81; SE = .38; p = .03). There were no other differences in perceived challenge. These results suggest that those students who seemed to improve the most in their performance had lower perceived challenge than those whose performance was high, but stable.

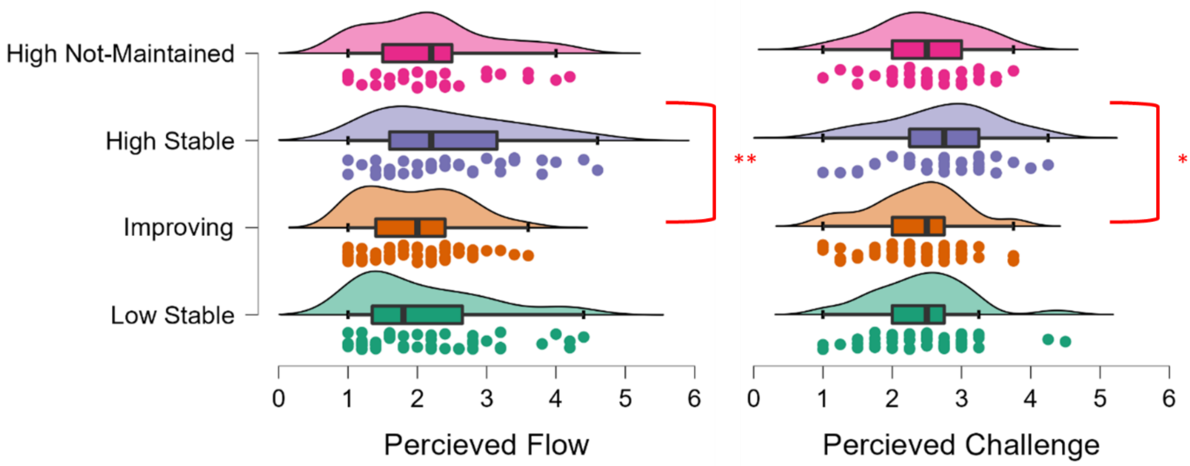

Figure 6. (**a&b**) Estimated marginal means for Perceived Flow & Challenge at post-test. Note: * $p < .05$, ** $p < .01$.

In order to examine the relation between game performance profiles and game play, we employed the 3-step approach to examine if there were differences in the four profiles on their time on task and the number of maps completed.

Based on the multinomial logistic regressions from the 3-step procedure, there were differences between profiles in time on task (See Figure 7a). Students in the Improving Performers profile were shown to have played longer than students in the Low Stable profile (beta = .012; SE = .005; p = .03). There were no other differences in the time on task between the groups.

Based on the multinomial logistic regressions from the 3-step procedure, there were differences between profiles in the number of maps completed (See Figure 7b). Students in the Improving Performers profile were shown to have completed more maps than students in the Low Stable profile (beta = .04; SE = .02; p = .02), who also completed fewer maps than the High Stable profile (beta = .05, SE = .02, p = .04). There were no other differences in the number of maps completed between the groups.

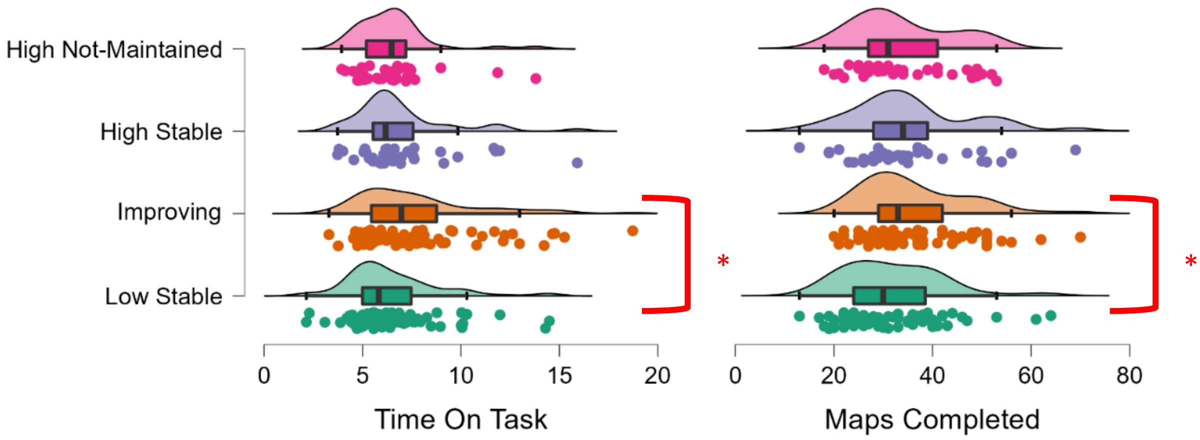

Figure 7. (**a&b**) Estimated marginal means for Time on Task & Maps Completed. Note: $p < .05$.

4. Discussion

The present study aimed to examine in more detail the different approaches to playing the Number Navigation Game among those students who completed a substantial number of energy maps, which emphasized using complex arithmetic relations. Overall results revealed substantial individual differences in how students performed across the energy maps they completed while playing the game during a classroom intervention. Notably, the majority of participants did not exhibit substantial changes in their performance while playing the game. Indeed, most participants' performance in the game could be clearly connected with their prior knowledge of arithmetic, especially their adaptive number knowledge. However, profile membership was related to learning gains, even after taking into account prior knowledge. These results are in line with previous research that indicates game performance in the NNG is related to students' learning outcomes [12].

We also aimed to examine if it was possible to identify students who engaged in practice within the game that resembled features of deliberate practice. In line with this aim, the Improving Performance Profile exhibited signs of engaging in deliberate practice-like behavior while playing the game. First, this group of students was the only group who consistently improved their performance on the most mathematically challenging maps of the NNG - the energy mode. One core characteristic of engaging in deliberate practice is practicing skills that are at the edge of one's competencies [6], [14]. In line with this characteristic, students in the Improving Performance profile showed indications that they were constantly aiming to improve their performance in the game.

The Improving Performance Profile provided strong evidence of significant learning gains specifically related to adaptive number knowledge, which is the main focus of the energy maps. Although this evidence alone may not be conclusive in demonstrating deliberate practice by the players, it is consistent with the theories of deliberate practice. According to these theories, deliberate practice involves focused and intentional practice of well-defined skills in order to improve performance [6], [14]. These results are further supported by the fact that students in the Improving Performance profile did not show an advantage in the development of their arithmetic fluency, which is less aligned with the targeted skills in the energy maps. Thus, the learning outcomes of the Improving Performance profile aligned with what would be expected if they were engaged in deliberate practice-like behavior while playing these energy maps.

Third, students in the Improving Performance profile showed signs of having less experience of flow while playing the game. The nature of deliberate practice and flow experiences suggests they should not necessarily exist simultaneously [21], [24]. In particular, deliberate practice requires full sustained concentration and effort, which is often uncomfortable as it requires practicing skills that are just beyond one's current abilities [14]. In contrast, flow should be a more immersive experience that is felt when one's skills match the requirements of the task, and when one often loses one's sense of self [26]. Thus, the lower flow experiences of the Improved Performance profile are in line with what would be expected if they were engaged in deliberate practice.

Additionally, these students appeared to be more likely to complete a higher number of maps in the game while spending more time playing compared with Low Stable profile, who had a similar initial game performance. Again, while not directly evidencing deliberate practice-like behaviors, this result is in line with what would be expected. In particular, the need to engage with content in a systematic and persistent way to improve performance would most likely necessitate repeating maps in attempts to improve scores [31]. This result is in agreement with previous research which found that the slow-and-steady progressors – students who repeatedly replayed and reattempted to improve their solutions in an online mathematical game (*Materials: From Here to There!*) – showed the largest absolute learning gains compared to other groups [27]. Thus, those students who engage in deliberate practice could be expected to spend more time playing the NNG and also complete more maps. In total, these results present further evidence that the students in the Improving Performance profile engaged in deliberate practice-like behaviors when they played the game.

Finally, another reason for us to believe this group engaged in deliberate practice-like behavior in the Energy maps is due to the design of the game version that they played. As this was the first complete working version, the usability and clarity of the game interface were not yet optimal. A previous study confirmed that students' gaming experience in a later version of the NNG, namely NNG2, was higher when compared to those students who played this version of the NNG [31]. Hence, the Improving Performance profile students were able to improve their game performance despite these usability shortcomings. Given these shortcomings, especially in the more complex maps in the Energy mode, we expect that these students would have to consciously decide to play the game more times than needed. In other words, they may have deliberately chosen to practice and look for different alternatives to find more efficient solutions. This kind of practice is possible because the flexible and open nature of the game is supported with fixed and clear rules (i.e., game modes, materials retrieving to and from the harbor, etc.) and novel contexts (i.e., map layouts, positions of harbor, ship, and materials, etc.), which allow students to practice and improve their game performances [23].

4.1 Theoretical implications

These results suggest that game-based learning environments may be designed in a way to elicit deliberate practice-like behaviors during gameplay. Previously, the NNG has been proposed as possibly eliciting deliberate practice, yet, up until now there was no direct empirical evidence of this [6], [12]. As well, our results reveal that slow and sustained progress in success in the NNG is just as beneficial for learning as consistently high levels of performance. The fact that those students who started out performing poorly on the early maps, but gradually improved their performance, ended up making similar gains in their adaptive number knowledge as those who consistently were performing well suggests that engagement in continuous development in line with features of deliberate practice is a powerful learning strategy.

Additionally, these results are the first we know of to combine investigations of flow theory in game-based learning with examinations of deliberate practice in students' gameplay. Flow experiences and deliberate practice have been argued to potentially conflict with each other [39]. Deliberate practice is described as a carefully planned and effortful behavior that involves working where challenge exceeds skill [14]. In contrast, flow is an enjoyable experience characterized by a balance between challenge and skill, which is described as spontaneous and effortless [25]. In line with these descriptions, our results show that those students who were able to make persistent gains in their performance appeared to have lower flow experiences.

This indicates they were practicing skills that were not well-calibrated with their existing abilities, suggesting their gameplay was more in line with deliberate practice than flow. However, it should be noted that the Improving Performing profile students also had lower retrospective perceived challenge compared to those in the High Stable profile. One possibility is that the students who improved throughout the game could view this improvement as the game being easier to master than those students who did not show any substantial improvement throughout the game (as was the case with the High Stable profile students). In any case, the incongruency between flow experiences and challenge being lower in the Improving Performance profile suggests further examination of this phenomenon is valuable.

4.2 Practical implications

One important practical implication of our study is that it serves as an example for future applications of deliberate practice in the promotion of adaptive expertise via game-based learning environments in mathematics. Previously, traditional practices that aim to strengthen understanding and representation of numerical characteristics and relations would require one-on-one teacher support and scaffolding [40]. With the affordances of technology and game-based learning, it is possible to provide opportunities for students to engage with non-routine mathematical tasks that are challenging and individually well-suited [11], [22]. These tasks require students to practice at the edge of their competence, and continuous feedback and learning progress are made visible to help improve students' current levels. The arithmetic tasks in NNG are positioned in an open-ended learning environment that triggers mental reflections on possible alternative solutions to the tasks that are not "one dimensional" as there are varying possible solution paths. Furthermore, because there are no ready-made mathematical tasks, the design of the NNG offers great "replay value" with alternative options to practice [41]). As opposed to traditional textbooks, the NNG allows for multiple trials, errors, and reattempts as students can decide to restart a map or replay a previously finished map to get higher scores with new mathematical solutions compared to their previous problem-solving strategies based on clear visual feedback (e.g., Gold coins earned).

5. Limitations and Future Directions

There are several limitations to our study. Firstly, while using game logs to determine the four-profile memberships allows us to identify different profiles of game performance and their development trajectories, we know little about other external factors such as classroom conditions or teachers' support. Also, future studies should include other types of game analytics to measure students' "in-game experience" [24] and explore the roles of teachers with NNG in the classrooms [31]. Next, we argued that one group engaged in deliberate practice-like behavior based on their game performance, learning gains, lower flow experience, and game design. However, we are uninformed about their patterns of practice, or why the game only enabled deliberate practice for one group. How to trigger deliberate practice and maintain such vigorous activity in game-based learning environments in as many students as possible are questions that remain unexplored. Also, prior knowledge of the Low Stable group is lower than Improving Performers, therefore it might limit their capacities in self-initiating more intense practice. Thus, more attention is needed in future studies on students' prior knowledge. Lastly, it is also worth examining whether the NNG triggers deliberate practice behaviors in Low Stable's game performance in easier game maps (for instance when the game requires only addition and subtraction operations).

Despite these limitations, this study answers unaddressed questions about the relations between game performance and learning development; it sheds light on the little-known application of deliberate practice in mathematics education via game-based learning platforms.

The results indicated that while both adaptive expertise in mathematics and deliberate practice are very demanding and challenging concepts to directly apply in authentic settings, well-designed game-based learning environments offer unprecedented opportunities to overcome those limitations and provide unmatched advantages to "cultivate mathematical minds" in the future [6].

Acknowledgments

This study was funded by the Jacobs Foundation, awarded to JM, and grant 336068 from the Strategic Research Council of Finland to MH-S.

Conflicts of interest

All authors declare that they have no conflicts of interest.

References

[1] L. Verschaffel, K. Luwel, J. Torbeyns, and W. Van Dooren, "Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education," *Eur. J. Psychol. Educ.*, vol. 24, pp. 335–359, 2009, doi: 10.1007/BF03174765.

[2] H. Sievert, A. K. van den Ham, I. Niedermeyer, and A. Heinze, "Effects of mathematics textbooks on the development of primary school children's adaptive expertise in arithmetic," *Learn. Individ. Differ.*, vol. 74, no. January, p. 101716, 2019, doi: 10.1016/j.lindif.2019.02.006.

[3] A. J. Baroody, "The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge," in *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise*, A. J. Baroody and A. Dowker, Eds., Lawrence Erlbaum Associates, 2003, pp. 1–33. doi: 10.4324/9781410607218.

[4] G. Hatano and K. Inagaki, "Two courses of expertise," in *Child Development and Education in Japan*, H. Stevenson, J. Azuma, and K. Hakuta, Eds., New York: W. H. Freeman & Co, 1986, pp. 262–272.

[5] J. McMullen, M. M. Hannula-Sormunen, E. Lehtinen, and R. S. Siegler, "Distinguishing adaptive from routine expertise with rational number arithmetic," *Learn. Instr.*, vol. 68, no. June 2019, p. 101347, 2020, doi: 10.1016/j.learninstruc.2020.101347.

[6] E. Lehtinen, M. M. Hannula-Sormunen, J. McMullen, and H. Gruber, "Cultivating mathematical skills: From drill-and-practice to deliberate practice," *ZDM - Math. Educ.*, vol. 49, no. 4, 2017, doi: 10.1007/s11858-017-0856-6.

[7] S. Van den Eynde, M. Goedhart, J. Deprez, and M. De Cock, "Role of Graphs in Blending Physical and Mathematical Meaning of Partial Derivatives in the Context of the Heat Equation," *Int. J. Sci. Math. Educ.*, Jan. 2022, doi: 10.1007/s10763-021-10237-3.

[8] P. J. Feltovich, R. J. Spiro, and R. L. Coulson, "Issues of expert flexibility in contexts characterized by complexity and change," in *Expertise in context: Human and machine*, P. J. Feltovich, K. M. Ford, and R. R. Hoffman, Eds., Menlo Park, CA: AAAI/MIT Press, 1997, pp. 125–146.

[9] G. Hatano, "Foreword," in *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise*, A. J. Baroody and A. Dowker, Eds., 2003, pp. xi–xii.

[10] J. McMullen *et al.*, "A game-based approach to promoting adaptive rational number knowledge," *Math. Think. Learn.*, vol. 0, no. 0, pp. 1–17, Feb. 2023, doi: 10.1080/10986065.2023.2177818.

[11] K. Devlin, *Mathematics education for a new era: Video games as a medium for learning*. Natick, Massachusetts: A K Peters, 2011.

[12] B. Brezovszky *et al.*, "Effects of a mathematics game-based learning environment on primary school students' adaptive number knowledge," *Comput. Educ.*, vol. 128, pp. 63–74, Jan. 2019, doi: 10.1016/j.compedu.2018.09.011.

[13] M. P. J. Habgood, “The Effective Integration of Digital Games and Learning Content.” (Unpublished doctoral dissertation). Nottingham, UK: University of Nottingham. Retrieved from: https://eprints.nottingham.ac.uk/10385/1/Habgood_2007_Final.pdf, 2007.
[14] K. A. Ericsson, R. Th. Krampe, and C. Tesch-Römer, “The role of deliberate practice in the acquisition of expert performance,” *Psychol. Rev.*, vol. 1, no. 3, 1993. doi:10.1037/0033-295X.100.3.363
[15] Common Core State Standards Initiative, “Common Core State Standards for Mathematics,” Washington, DC, 2010. [Online]. Available: http://www.corestandards.org/
[16] Finnish National Agency for Education, *Core curriculum for basic Education 2014.* 2014.
[17] M. Hickendorff, “Flexibility and Adaptivity in Arithmetic Strategy Use: What Children Know and What They Show,” *J. Numer. Cogn.*, vol. 8, no. 3, pp. 367–381, Nov. 2022, doi: 10.5964/jnc.7277.
[18] J. McMullen, B. Brezovszky, G. Rodríguez-Aflecht, N. Pongsakdi, M. M. Hannula-Sormunen, and E. Lehtinen, “Adaptive number knowledge: Exploring the foundations of adaptivity with whole-number arithmetic,” *Learn. Individ. Differ.*, vol. 47, pp. 172–181, 2016, doi: 10.1016/j.lindif.2016.02.007.
[19] J. McMullen *et al.*, “Adaptive number knowledge and its relation to arithmetic and pre-algebra knowledge,” *Learn. Instr.*, vol. 49, pp. 178–187, Jun. 2017, doi: 10.1016/j.learninstruc.2017.02.001.
[20] J. Moss and R. Case, “Developing Children’s Understanding of the Rational Numbers: Curriculum A New Model and an Experimental,” *J. Res. Math. Educ.*, vol. 30, no. 2, pp. 122–147, 1999, doi: 10.2307/749607.
[21] K. A. Ericsson, “The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance,” in *The Cambridge Handbook of Expertise and Expert Performance*, 2006, pp. 685–705. doi: 10.1017/CBO9780511816796.038.
[22] B. Brezovszky, “Using Game-Based Learning to Enhance Adaptive Number Knowledge,” University of Turku, 2019. [Online]. Available: http://urn.fi/URN:ISBN:978-951-29-7622-5
[23] E. Lehtinen *et al.*, “Number Navigation Game (NNG): Design Principles and Game Description,” in *Describing and Studying Domain-Specific Serious Games*, J. Torbeyns, E. Lehtinen, and J. Elen, Eds., Cham: Springer International Publishing, 2015, pp. 45–61. doi: 10.1007/978-3-319-20276-1_4.
[24] K. J. M. Kiili, A. Lindstedt, A. Koskinen, H. Halme, M. Ninaus, and J. McMullen, “Flow Experience and Situational Interest in Game-Based Learning: Cousins or Identical Twins,” *Int. J. Serious Games*, vol. 8, no. 3, Art. no. 3, Sep. 2021, doi: 10.17083/ijsg.v8i3.462.
[25] M. Csikszentmihalyi, *Flow and the Foundations of Positive Psychology*. Dordrecht: Springer Netherlands, 2014. doi: 10.1007/978-94-017-9088-8
[26] K. Kiili, T. Lainema, S. de Freitas, and S. Arnab, “Flow framework for analyzing the quality of educational games,” *Entertain. Comput.*, vol. 5, no. 4, pp. 367–377, 2014, doi: 10.1016/j.entcom.2014.08.002.
[27] J.-E. Lee, J. Y.-C. Chan, A. Botelho, and E. Ottmar, “Does slow and steady win the race?: Clustering patterns of students’ behaviors in an interactive online mathematics game,” *Educ. Technol. Res. Dev.*, vol. 70, no. 5, pp. 1575–1599, Oct. 2022, doi: 10.1007/s11423-022-10138-4.
[28] B. O. Muthen, “The potential of growth mixture modeling,” *Infant Child Dev.*, vol. 3, no. December 2007, pp. 3–5, 2006, doi: 10.1002/icd.
[29] G. Rodríguez-Aflecht, T. Jaakkola, N. Pongsakdi, M. Hannula-Sormunen, B. Brezovszky, and E. Lehtinen, “The development of situational interest during a digital mathematics game,” *J. Comput. Assist. Learn.*, vol. 34, no. 3, pp. 259–268, 2018, doi: 10.1111/jcal.12239.
[30] C. Walkington, “Intrinsic integration in learning games and virtual instruction,” *Educ. Technol. Res. Dev.*, no. November, pp. 1–4, 2020, doi: 10.1007/s11423-020-09886-y.
[31] P. Bui, G. Rodríguez-Aflecht, B. Brezovszky, M. M. Hannula-Sormunen, S. Laato, and E. Lehtinen, “Understanding students’ game experiences throughout the developmental process of the number navigation game,” *Educ. Technol. Res. Dev.*, vol. 68, no. 5, pp. 2395–2421, Oct. 2020, doi: 10.1007/s11423-020-09755-8.
[32] R. W. Woodcock, K. S. McGrew, and N. Mather, *Woodcock-Johnson III Tests of Achievement*, vol. 2001. 2001.
[33] J. L. Berger and S. A. Karabenick, “Motivation and students’ use of learning strategies: Evidence of unidirectional effects in mathematics classrooms,” *Learn. Instr.*, vol. 21, no. 3, pp. 416–428, 2011, doi: 10.1016/j.learninstruc.2010.06.002.

[34] K. Poels, Y. A. W. de Kort, and W. A. IJsselsteijn, “D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games.” Technische Universiteit Eindhoven, 2007.
[35] K. Oksanen, “Subjective Experience and Sociability in a Collaborative Serious Game,” *Simul. Gaming*, vol. 44, no. 6, pp. 767–793, Dec. 2013, doi: 10.1177/1046878113513079.
[36] K. L. Nylund, T. Asparouhov, and B. O. Muthén, “Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study,” *Struct. Equ. Model. Multidiscip. J.*, vol. 14, no. 4, pp. 535–569, Oct. 2007, doi: 10.1080/10705510701575396.
[37] M. Hickendorff, P. Edelsbrunner, J. McMullen, M. Schneider, and K. Trezise, “Informative Tools for Characterizing Individual Differences in Learning: Latent Class, Latent Profile, and Latent Transition Analysis,” *Learn. Individ. Differ.*, vol. 66, 2018, doi: 10.17605/OSF.IO/JMTCV.
[38] T. Asparouhov and B. Muthén, “Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using M plus,” *Struct. Equ. Model. Multidiscip. J.*, vol. 0, no. June, pp. 1–13, 2014, doi: 10.1080/10705511.2014.915181.
[39] A. Duckworth, *Grit: The power of passion and perseverance*. in Grit: The power of passion and perseverance. New York, NY, US: Scribner/Simon & Schuster, 2016, pp. xv, 335.
[40] A. Heirdsfield, “Teaching mental computation strategies in early mathematics,” *YC Young Child.*, vol. 66, no. 2, pp. 96–102, 2011.
[41] D. A. Phoenix, “How to Add Replay Value to Your Educational Game,” *J. Appl. Learn. Technol.*, vol. 4, no. 1, pp. 20–23, Winter 2014.