MIND, BRAIN, AND EDUCATION

Effects of a Teacher-Designed and Teacher-Led Numerical Board Game Intervention: A Randomized Controlled Study with 4- to 6-Year-Olds

Zachary Hawes[1], Michelle Cain[2], Shelly Jones[2], Nicole Thomson[2], Cristol Bailey[2], Jisoo Seo[3], Beverly Caswell[3], and Joan Moss[3]

ABSTRACT— The purpose of the current pilot study was to examine the effects of a teacher-designed and teacher-led numerical board game intervention. Fifty-four 4- to 6-year-olds were randomly assigned to either the number board game intervention or an active control group. Relative to the control group, children who received the number game intervention demonstrated significant improvements on a numeral identification task. This finding is significant in so far as numeral identification skills play a critical role in more advanced numerical and mathematical reasoning. There was no evidence of training-related improvements on any of the other tasks. In addition to the intervention effects, the present study provides an example of a successful teacher-researcher collaboration, providing new insights into the making of bidirectional relations between research and practice.

Questions of how people learn and develop are of central interest to educators and cognitive scientists alike. Moreover, the disciplines of education and cognitive science bring with them distinct yet complementary knowledge, skill sets, and approaches to questions on human learning. Whereas educators have much to offer in terms of the realities of the "blooming, buzzing confusion" of classroom learning (Brown, 1992, p. 141), researchers have much to offer about learning through carefully controlled laboratory-based studies. These are but some of the reasons that proponents of Mind, Brain, and Educatin (MBE) have argued for and prioritized the need to establish "bidirectional relationships" and "two-way roads" between research and practice (Ansari, Coch, & De Smedt, 2011; De Smedt et al., 2011). Indeed, a central goal of MBE is to improve collaboration between professionals in education and the cognitive, developmental, and brain sciences. As regularly discussed in this journal, the potential benefits of such collaborations are many (e.g., Fischer, 2009); however, the actual reporting of such collaborations are rare (e.g., Samuels, 2009).

The current study addresses this gap in the literature and reports on a collaborative research project between practicing kindergarten teachers, mathematics educators, and developmental cognitive scientists. More specifically, we describe our joint efforts to design, implement, and study the effects of a linear numerical board game on young children's numerical reasoning.

[1]Department of Psychology, Brain and Mind Institute, University of Western Ontario,
[2]Rainy River District School Board,
[3]Department of Applied Psychology and Human Development, Ontario Institute for Studies in Education, University of Toronto,

Address correspondence to Zachary Hawes, Department of Psychology, Brain and Mind Institute, University of Western Ontario, Western Interdisciplinary Research Building, 1151 Richmond Street North, London, ON N6A 5B7, Canada; e-mail: zhawes@gmail.com

Background on the Intervention Design

The numerical board game was created by two kindergarten teachers (Cain and Jones) during an in-service Professional

Fig. 1. Image of educators in process of designing their own numerical boards games for their classrooms.

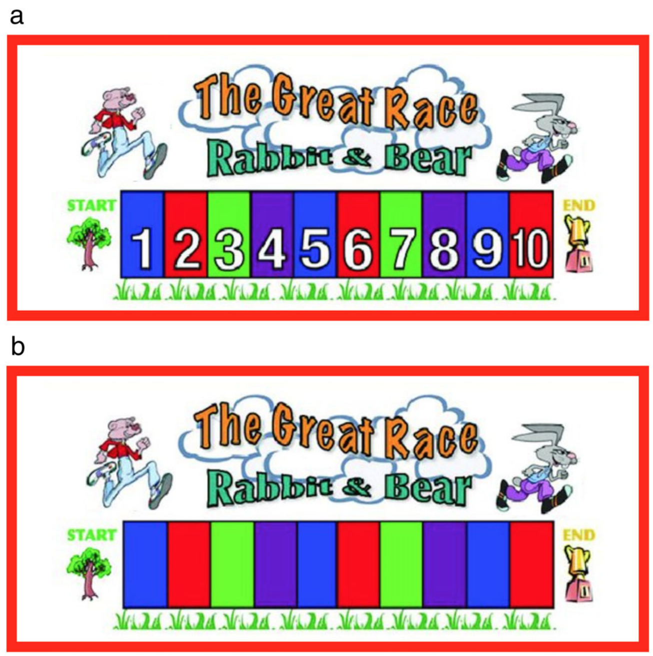

Fig. 2. Example of the linear board games used in the studies by Siegler and Ramani (2008; 2011). (a) Number board game, (b) color board game.

Development (PD) meeting. The idea for the intervention was borne out of the challenge to create an engaging, yet theoretically-based, numerical board game for early years students (K-2). This challenge was presented by researchers and facilitators of the PD (Moss, Caswell, and Hawes) to a group of 17 early years educators (K-2) as part of an ongoing teacher-researcher collaboration. In the previous year, we had worked together to implement a 32-week spatial training intervention (for details of the PD model and the effects of the spatial intervention, see Hawes, Moss, Caswell, Naqvi, & MacKinnon, 2017). Building on the success of this collaboration, the researchers were asked to return for another year of teacher PD (i.e., five full days of PD throughout the school year). In the present iteration of our work, the focus was on improving our collective understanding and the instructional implications of conceptual mappings and relations between numbers and space (aka numerical-spatial associations; e.g., see Lakoff & Núñez, 2000). It was against this background that on the first day of our renewed PD together, the researchers challenged the group to create their own numerical board games. Members were encouraged to work in pairs and provided with a wide assortment of materials; many of which lent themselves to the creation of games emphasizing connections between numbers and space (e.g., square tiles, empty 1×10 and 1×20 partitioned arrays, blank hundred charts, dice, cubes, grid paper, cue cards, etc.: see Figure 1).

This exercise resulted in a handful of promising number interventions, but it was one game in particular that caught the attention of the research team. Two teachers (Cain and Jones) had designed a board game that had a remarkable resemblance to the number game interventions of Siegler and Ramani (2008). Since 2008, Siegler and Ramani's "The Great Race," has been subjected to numerous randomized controlled trials (RCTs) and has emerged as arguably the most successful early number intervention to date (e.g., Ramani & Siegler, 2008, 2011; Ramani, Siegler, & Hitti, 2012; Siegler, 2009; Siegler & Ramani, 2008, 2009, 2011). In brief, the game involves two players and the goal of being the first to reach the finish line. The game is played on a horizontal 1×10 rectangular array, with each partition labeled with the numerals 1–10 from left to right (see Figure 2). Players take turns spinning a "spinner" labeled with "1" or "2" on either half. Depending on the number spun, players take turns moving their game piece from left-to-right and "race" to be the first to reach end of the board game. Critically, on each spin, the player must "count on" from their current position. For example, if a child was at 4 and spun a 2, he/she would say, "5, 6" as the moved their piece. The control version of the game is played with the same materials and has the same goal but does not involve numbers (see Figure 2).

The game designed by Cain and Jones shares many similarities with "The Great Race". In their version, however, the game is played on a 1–20 number board game. Furthermore, rather than moving a token along the board game, each player has his/her own pile of colored tiles labeled 1–20. On each roll, the player must select one or two tiles that correspond to the next one or two number positions on the game board. For example, if a player was at 4 and spun a 2, he/she would select the tiles labeled 5 and 6 and place them on their corresponding positions labeled "5" and "6" on the game board (i.e., symbol-to-symbol mapping). Moreover, identical to Ramani and Siegler's version, players would say "5,6" as they placed down their tiles and moved from left to right along the number sequence (see Methods for further details). As discussed next, despite some differences in game play, both games were designed and theorized to improve students' basic number knowledge through similar mechanisms.

Theoretical Underpinnings

There are a number of mechanisms hypothesized to underlie the success of the "The Great Race." Indeed, roughly an hour of game play has been consistently linked to substantial gains in 4- to 6-year-olds' arithmetic, number line estimation, counting, number identification, and magnitude comparison skills (e.g., see Siegler & Ramani, 2011). The game is purported to target the development and refinement of the "mental number line" (e.g., see Ramani, Siegler, & Hitti, 2012). The mental number line account posits that accurate left-to-right mappings of numbers to space (in Western cultures) is critical to a mature understanding of numbers (e.g., Siegler & Opfer, 2003). According to Okamoto and Case (1996), the mental number line represents the central conceptual structure underlying early numerical understanding. A common way of assessing one's "mental number line" is through number line estimation tasks. Individuals are presented with a horizontal line, marked by "0" at the left end and another number at the right end of the line (e.g., "100"). A target number is presented (e.g., "50") and its correct location is to be indicated on the line. A mass of evidence has demonstrated strong links between number line performance and mathematics achievement (e.g., see Schneider et al., 2018). Findings from the number line task have led some to conclude that better performance is indicative of a more precise "mental number line," which in turn, is thought to underlie higher mathematical reasoning (e.g., Siegler & Ramani, 2008; but see Barth & Paladino, 2011). According to Ramani and Siegler (2008), linear numerical board games, like Snakes and Ladders, "provide a physical realization of the mental number line (p. 377)." Taken together, playing linear numerical board games may promote foundational number skills through helping children learn to accurately map numbers to space; a critical skill, underlying a breadth of mathematical tasks (e.g., see Marghetis, Núñez, & Bergen, 2014).

Siegler and Ramani (2011) further suggest that linear numerical games are effective at improving young children's numerical knowledge because they offer multiple cues to both the order of numbers and the numbers' magnitudes. For example, as children move their tokens along the linear number path, the greater the: "*(a) the distance the child has moved the token; (b) the number of discrete moves of the token the child has made; (c) the number of number names the child has spoken; (d) the number of number names the child has heard; and (e) the amount of time since the game began*" (Siegler & Ramani, 2011, p. 346). Together, these visual–spatial, kinesthetic, auditory, and temporal cues are thought to contribute to a multimodal, embodied linear representation of number (Siegler & Ramani, 2011). In addition, number board games provide young children with practice in number identification and counting. Taken together, linear board games are expected to improve children's understanding and manipulation of numerical magnitudes as well as more basic counting and numeral identification skills (Siegler & Ramani, 2008).

Notably, in presenting the rationale behind their game, Cain and Jones listed many of the same reasons as Siegler and Ramani for why their game may contribute to improved number knowledge in young children. When asked what numerical skills their game targeted, Cain and Jones mentioned: *(a) numeral recognition; (b) symbol-to-symbol mapping; (c) mapping numbers to space/ordinal position; (d) counting on; and (e) estimating/calculating difference and dealing with proportionality (e.g., How many more spaces to the end? Are we half way yet?).* Interestingly, the differences in Cain and Jones' list compared to that of Siegler and Ramani's (2008), reflects the differences in game play. This is most apparent in the symbol-to-symbol mapping present in Cain and Jones' game but not "The Great Race." Players in Cain and Jones' version had to first use the board game to identify which number(s) to select from their pile of square tiles (labeled with numerals 1–20), actively search and subsequently select the correct numbered tile(s), and then place the selected number tile(s) on their corresponding numerals on the board game. It was because of this need to identify and map symbols-to-symbols that we hypothesized the greatest gains would occur in children's performance in numeral identification. It is also notable that although Cain and Jones made no explicit mention of the "mental number line" (an unfamiliar concept at the time), there is emphasis placed on the mapping of numbers to space according to ordinal position. Thus, for the same reasons hypothesized by Siegler and Ramani noted above, we predicted that game play would result in an improved understanding of linear-numerical relations. Through game play, children are afforded multiple cues that the numbers' magnitudes are directly related to their position on the game board. In other words, game play may result in a more refined and accurate mental representation of numbers by way of an improved "mental number line." An improved understanding of the linear structure of numbers, in turn, may confer gains in children's ability to compare magnitudes as well as estimate the relative positions of where numbers belong in relation to other numbers. For example, through repeatedly playing the game, children may come to recognize where the number "5" falls in relation to other numbers, such as "2" and "9." Such recognition and understanding is useful when comparing magnitudes (e.g., 5 vs. 2; 5 vs. 9) as well as placing numbers accurately on an empty number line. In sum, based on the theoretical and empirical work of Siegler and Ramani, along with Cain and Jones' own rationale for why the game would work, it was expected that this newly created game would yield gains across a variety of number knowledge tasks by way of an increased understanding of linear-numerical relations.

The Present Study

Following their presentation, the researchers approached Cain and Jones' and asked whether they were aware of the work of Siegler and Ramani. They were not. The researchers summarized the research that had been carried out with "The Great Race" and in passing mentioned that it would be interesting to conduct a study to examine the effects of their own game. Cain and Jones later followed up on this conversation and expressed interest in carrying out a study on the game. They had also recruited four additional kindergarten teachers who were interested in participating in the study. For the next several months, we—a group of six kindergarten teachers and four researchers—worked together to design and establish the methods reported hereafter.

METHODS

Participants

Fifty-four children aged 4–6 years participated (M_{age} = 5.0 years, *SD* = .53). Children were recruited from two elementary schools located in Northwestern Ontario, Canada. A total of six kindergarten teachers, three from each school, consented to have their classrooms participate. Prior to student recruitment, ethics approval was granted by the University of Toronto and the district school board committee. Information letters and consent forms went home to all students; 91% of parents agreed to have their child participate. Both participating schools serve students from middle SES backgrounds and perform at or above the provincial standard in mathematics. Because of absenteeism during either one of the testing dates ($n = 8$) and unwillingness to participate ($n = 3$), a total of 43 children were included in the final sample (see Table 1 for details).

Study Design

A randomized, controlled pre-post study design was employed. Participants were randomly assigned to either the number board game training condition or the color game control condition (see Table 1). All Junior Kindergarten students (the equivalent of preschool in the U.S.) were randomly assigned to condition. However, because of time constraints and our decision to target a lower age-range, we intentionally randomly selected only half of all eligible Senior Kindergarten students (equivalent of Kindergarten in the U.S.) to participate. All participants took part in identical pretest and post-test measures. Pretests were completed 3–5 days before the intervention began and the post-tests were completed 1–3 days immediately following the intervention. Testing was carried out in a quiet location of the school by a trained test administrator who was blind to group assignment at all times.

Description of Training Procedure and Intervention Games

The participating teachers scheduled and managed children's game play over a 3-week period. Teachers kept detailed logs of game play, including the duration of game play as well as any observational notes. Children in both the experimental and control condition played the games for 10–15 min between three and four times over the 3-week period. As further discussed below, children played the games in pairs and with the guidance of the teacher. Note that children were paired with a different student for each session of game play.

Number Board Game Intervention

The number board game was played in partners and included the following materials: a single linear board game consisting of 20 adjoined squares (see Figure 3), a dice with numbers 1 or 2 on each side, and two sets of different colored square tiles numbered 1–20 (i.e., one set of 20 per player; see Figure 4). Game play involved a "number race," in which each player took turns rolling the dice and placing one or two square tiles on the board. The tiles selected to be placed on the board (number line) were based on the placement of the last tile. For example, if the last tile placed was on number 10 of the number line, then a roll of two would require

Table 1
Participant Information by Group

	Participants	*JK students*	*SK students*	*Mean age; years (SD)*	*Females:males*	*Total duration of training (min)*
Number game group	20	15	5	4.9 (0.57)	11:09	44 (10.65)
Color game group	23	15	8	5.1 (0.55)	10:13	46 (10.19)

Note. JK = Junior Kindergarten (equivalent of preschool in the U.S.); SK = Senior Kindergarten (equivalent of preschool in the U.S.).

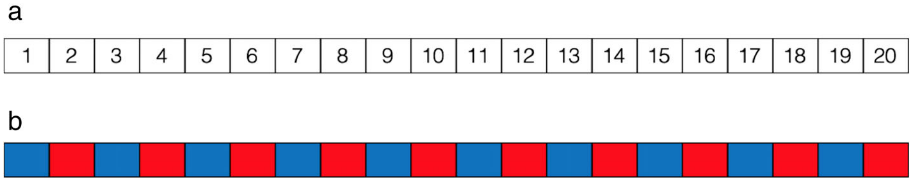

Fig. 3. Linear board games used in the experimental condition (a) and control condition (b).

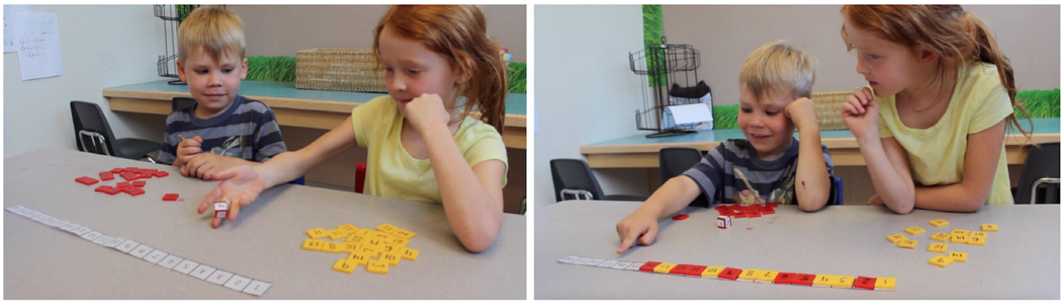

Fig. 4. An illustration of game play in action. The image on the left shows the very first roll of the game. The image on the right shows how the game has progressed. Both children have paused here to count how many more spaces remain.

the player to select the tiles labeled 11 and 12 from his/her pile of tiles. He/she would then "count on" from tile number 10 and say aloud "eleven, twelve." The game continued in this way until one of the partners eventually landed on 20, at which point he/she would be considered the winner. To view a video example of game play, visit: https://osf.io/yt7zk/.

Color Game Intervention

The color game was also played in partners and included the following materials: a single linear board game consisting of 20 adjoined squares, alternating between blue and red (see Figure 3), a dice with either blue or red on each side, and two sets of different colored square tiles (one set of 20 per player; see Figure 4). Game play involved a "color race," in which each player took turns rolling the dice and placing either one or two tiles on the board. For example, if a player was positioned on a red space and rolled a red, he/she would place two tiles along the path, saying "blue," "red," as they laid their tiles. If a player was positioned on red and rolled a blue, he/she would place one blue tile on the board and say "blue" as the laid their tile. Game play progressed in this way until one of the partners eventually landed on the last position of the game.

Pretest and Post-Test Measures

Participants took part in identical pretests and post-tests. The measures were based on the ones used by Ramani and Siegler (2008) and presented in the following order at both pre and post: Counting, Numeral Identification, Number Line Estimation (0–10), Symbolic Number Comparison, Non-Symbolic Number Comparison, Number Line Estimation (0–20). Note that the symbolic and non-symbolic number comparison task were part of the same assessment tool (see http://www.numeracyscreener.org/). There are two versions of this assessment; in one version the symbolic task is presented first, followed by the non-symbolic task, and in the other version the order of presentation is reversed. Because we used both versions, children were randomly presented with either the symbolic or non-symbolic comparison task first.

Counting Task

Children were asked to count from 1 to 20 and were given a score based on the highest number reached before their first error. The maximum possible score was 20.

Numeral Identification

Children were presented with 20 randomly ordered cards (i.e., cards with numerals 1–20 on them). The child was asked to name the number presented on each card. Children were awarded a total score based on the number of correctly identified numerals.

Number Line Estimation (0–10)

Children were presented with a 25-cm horizontal line with "0" marked at the left end of the line and "10" marked at the right end of the line. A target number ranging from 1 to 9 was printed above the center of the line. After completing one practice trial with the target number "5," children were randomly presented with the remaining trials (i.e., separate empty number lines for each of the following numbers, 1–4 and 6–9). For each trial, children were asked to indicate the number's position on the line with a pencil.

Number Line Estimation (0–20)

This task was identical to the one above but used a 0–20 empty number line. Instead of using "5" for a practice trial the number "10" was used. Children were randomly presented with 18 sheets of paper as they were asked to indicate where numbers 1–9 and 11–20 belong on the line. For both the 0–10 and the 0–20 number line tasks, children were given a score based on the child's overall accuracy of their estimates. To do this, we calculated each child's percent absolute error (PAE) using the following formula:

$$\text{PAE} = \left| \frac{\text{Estimate} - \text{estimated quantity}}{\text{Scale of estimates}} \right| \times 100.$$

For example, if a child was asked to estimate the location of 3 on the 0–10 number line and placed his/her response at the location that corresponded to 5, the PAE would be 20%: $[(5-3)/10] \times 100$.

Symbolic Number Comparison

Children were administered the Numeracy Screener (see Hawes, Nosworthy, Archibald, & Ansari, 2019; Nosworthy, Bugden, Archibald, Evans, & Ansari, 2013). This assessment includes both symbolic and non-symbolic number comparison tasks. For the symbolic comparison tasks, participants were presented with pairs of numerals ranging from 1–9. For each pair (e.g., 2 | 7), children were asked to indicate the larger of two numerals using a pencil. Children were given a 2-min time limit to complete as many items as possible (56 items total). Each child is given a score based on their total number of correct responses.

Non-Symbolic Number Comparison

The test format and administration was the same as above but instead of indicating the larger of two numerals, children indicated the larger of two dot arrays (e.g., ∷· | ·⋰). In both formats, numerical magnitude was counterbalanced for the side of presentation (i.e., 4 | 9, 9 | 4). In the non-symbolic tasks, dot stimuli were controlled for area and density (for details, see Nosworthy et al., 2013).

Analytical Approach

Analyses were carried out using JASP (V. 0.9.0.1). This software package allowed us to conduct a combination of both frequentist and Bayesian statistics. Bayes factors were calculated to quantify the amount of evidence in support of training-related gains as well as the amount of evidence in favor of the null (i.e., no training gains). More specifically, for all preliminary analyses, we report on Bayes factors as they correspond to evidence in favor of the alternative hypothesis (i.e., presence of group differences at pre) compared to the null hypothesis (i.e., absence of group differences at pre). For these analyses, the symbol BF_{10} is used to signify the strength of evidence for the alternative hypothesis over the null. For all main analyses, we used Bayesian repeated measures analysis of variance (ANOVA) and report on the statistic referred to as Bayes factor inclusion (hereafter BF_{incl}). The BF_{incl} provides a means to quantify the amount by which the prior odds of including an effect term in the model (e.g., presence of a group × time interaction) is updated after observing the data. For example, a BF_{incl} of 3 indicates that the observed data have increased the odds of an interaction by a factor of 3. As described next, in cases where the Bayes factor is 3 or above, this is considered evidence in support of an interaction. In short, the higher the Bayes factor, the higher the odds of there being a group difference from pre-to-post. In cases where the reported Bayes factors are below 1, this is an indication that there is more support for a model that does not include an interaction factor.

Although Bayes factors are expressed on a continuum, the following recommendations have been outlined as a general guideline for the interpretation of Bayes factors: Bayes factors of 0–3 offer weak support for the H1 (i.e., alternative hypothesis), 3–20 positive support for the H1, 20–150 strong support for H1, and values greater than 150 as very strong evidence in favor of the H1 (Andraszewicz et al., 2015; Jarosz & Wiley, 2014; Raftery, 1995). We adhere to these guidelines in interpreting the results of the present article.

RESULTS

Preliminary Analyses

Our first analysis involved testing for group differences at pretest. Results revealed no statistically significant differences on any of the following measures: Age, $t(41) = -1.20$, $p = .24$, $BF_{10} = 0.54$; Counting Task, $t(41) = -0.38$, $p = .70$, $BF_{10} = 0.32$; Numeral Identification, $t(41) = -0.56$, $p = .58$, $BF_{10} = 0.34$; Symbolic Number Comparison, $t(41) = -1.79$, $p = .82$, $BF_{10} = 1.06$; Non-Symbolic Number Comparison, $t(40) = -0.54$, $p = .59$, $BF_{10} = 0.34$; Number Line Estimation (0–20), $t(40) = 1.13$, $p = .26$, $BF_{10} = 0.51$. A group difference was observed on the Number Line Estimation 0–10 task, $t(40) = 2.32$, $p = .03$, $BF_{10} = 2.42$; the number board group performed significantly worse than the color group. For this reason, on this particular task, we analyzed the effects of the intervention by conducting an analysis of covariance (ANCOVA) on the post-test scores, while controlling for pretest scores. All other effects where analyzed

Table 2
Mean Scores at Pretest and Post-Test for Each Group

	Number game group		*Color game group*	
Measures	*Pretest mean (SD)*	*Post-test mean (SD)*	*Pretest mean (SD)*	*Post-test mean (SD)*
Counting	16.60 (3.55)	17.75 (2.43)	17.00 (3.29)	17.09 (3.45)
Numeral identification	8.95 (6.19)	10.95 (6.64)	9.74 (2.51)	10.30 (2.57)
Symbolic number comparison	22.95 (11.07)	27.65 (13.72)	29.00 (11.10)	32.96 (11.72)
Non-symbolic number comparison	30.90 (10.75)	36.95 (8.00)	32.57 (9.22)	39.22 (8.15)
Number line (0–10)	0.31 (0.13)	0.25 (0.10)	0.22 (0.11)	0.21 (0.08)
Number line (0–20)	0.25 (0.06)	0.23 (0.10)	0.24 (0.09)	0.19 (0.07)

using repeated measure ANOVAs. Table 2 shows the mean scores at pre and post for each group.

Intervention Effects

Figure 5 provides an overall summary of the results. As predicted, there was evidence to suggest that the number game was effective in improving children's numeral identification skills. Compared to the control group, children in the number game condition made statistically significant improvements from pre to post, $F(1,41) = 5.03$, $p = .03$, $BF_{incl} = 3.01$. Averaged across all candidate models, the Bayes factor of 3.01 provides positive support for the inclusion a group × time interaction. There was no evidence of training associated gains on any of the other measures: Counting, $F(1,41) = 2.50$, $p = .12$, $BF_{incl} = 0.47$; Symbolic Number Comparison, $F(1,41) = 0.13$, $p = .72$, $BF_{incl} = 0.62$; Non-Symbolic Number Comparison, $F(1,40) = 0.08$, $p = .78$, $BF_{incl} = 0.44$; Number Line Estimation (0–10), $F(1,39) = 0.03$, $p = .87$, $BF_{incl} = 0.34$; Number Line Estimation (0–20), $F(1,37) = 0.38$, $p = .54$, $BF_{incl} = 0.47$. Bayesian analyses indicated that more data are needed to determine sufficient evidence for *or* against the inclusion of an interaction term. That is, with the exception of the numeral identification, more evidence is needed to claim support for or against training-related group differences from pre-to-post.

DISCUSSION

The purpose of the present study was to examine the effects of a teacher-designed number board game intervention. More specifically, a group of teachers and researchers collaborated to implement a RCT of the game in the participating teachers' kindergarten classrooms as part of regular instruction. Relative to an active control group, children who received the number game intervention demonstrated significant improvements on a numeral identification task. There was no evidence of training related improvements on any of the other number knowledge measures. In the following sections, we address why the intervention may have led to improvements in children's numeral identification skills, but not any of the other measures, and conclude by discussing ways in which this research achieves several goals of Mind, Bran, and Education.

Gains in Numeral Identification

The number board game was associated with gains in children's ability to identify numerals. This was a predicted outcome of game play given that a key feature of the game involved symbol-to-symbol mapping. Children had to first use the game board to identify which number(s) they had to select from amongst their pile, search through their pile to actually find the correct number(s), and then place the selected number tiles on their corresponding position on the game board. In fact, the teachers who led the intervention remarked that considerable more time was spent on this aspect of game play than any other. Taken together, the saliency of this particular aspect of the game may have led to improvements in children's numeral identification. Although we had expected more widespread improvements in children's number skills, the finding of improved numeral identification should not be overlooked. Numeral identification skills are a necessary and fundamental bridge to other more advanced numerical skills (e.g., number comparison, arithmetic, number line estimation; e.g., see Purpura, Baroody, & Lonigan, 2013). As such, "numeral knowledge may act as gatekeeper (or barrier) in the development of formal mathematical knowledge" (Purpura et al., 2013, p. 460). Future research efforts are needed to further investigate the role that numeral identification plays in the learning of higher level mathematics. A more extensive and high-powered study using the current game may be one way of approaching this goal.

Why the Game Did Not Work as well as We Had Predicted

We propose two reasons why the game may not have worked as well as we had predicted. First, the saliency of the symbol-to-symbol mapping component of the game

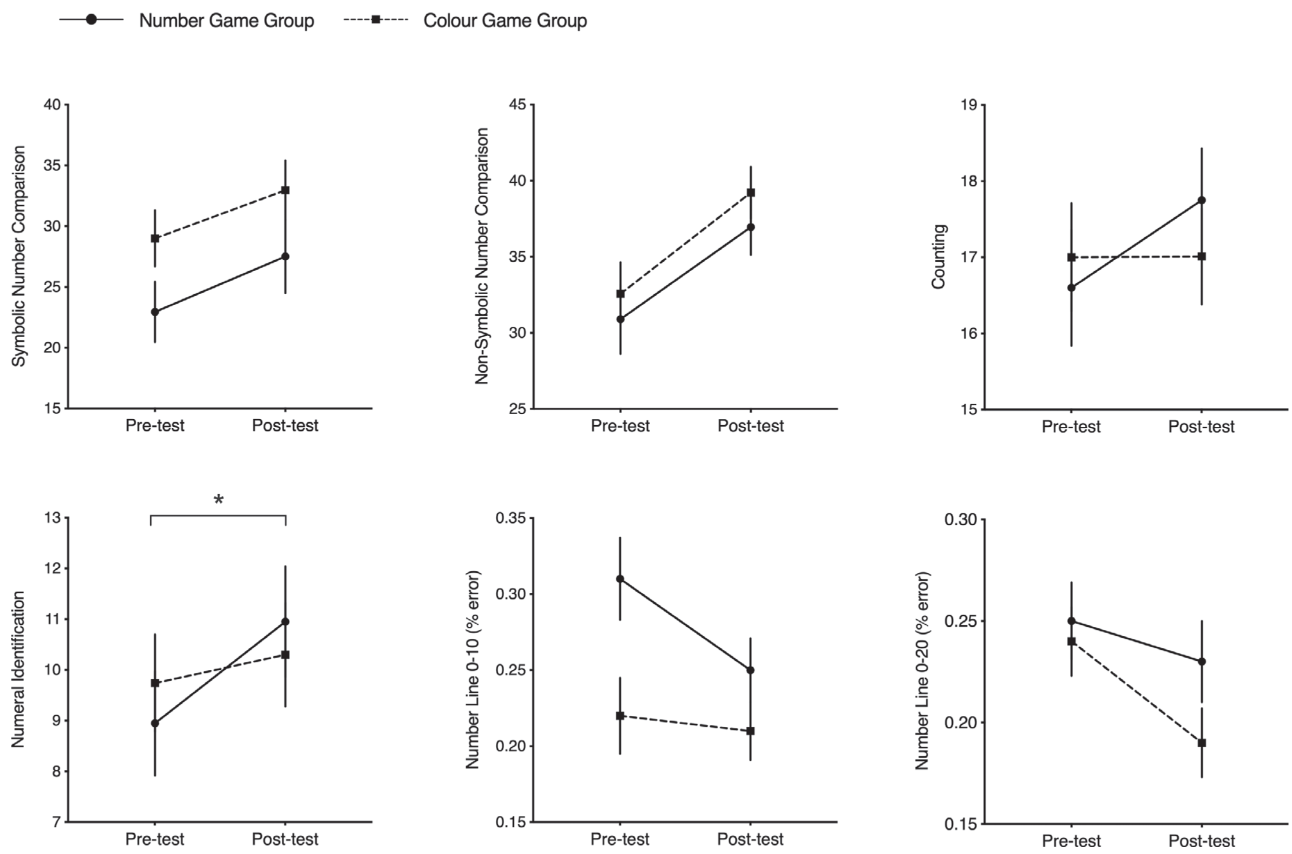

Fig. 5. Training gains by group. Error bars represent standard error of the mean. $^*p < .05$.

may have made other aspects of the game less salient. In other words, the time and effort spent searching, identifying, and mapping numbers symbols may have served as a double-edge sword: Providing ample practice at numeral identification skills, but detracting attention away from other learning cues (e.g., counting on, numerical-spatial relations, etc.). Second, the intervention was relatively short-lived (~45 min total) and carried out in the teachers' classrooms as part of regular math instruction. The classroom environment stands in stark contrast to the quiet and controlled environments of the lab (e.g., see Brown, 1992). Although the present intervention is similar to other successful board game interventions—in both game design and time spent training—longer training interventions may be required when carried out in ecologically valid contexts (Elofsson, Gustafson, Samuelsson, & Träff, 2016; Ramani & Siegler, 2008; Whyte & Bull, 2008). Future research efforts are needed to better understand the extent to which successful lab-based interventions also work in authentic learning contexts. The opposite is also true: What are the effects of taking classroom interventions, like the present one, and implementing them in the lab?

For the reasons just mentioned, we may not have achieved the same level of gains in students' numerical reasoning as are typically achieved with the "The Great Race" (e.g., see Ramani & Siegler, 2008). Additionally, another critical difference exists between our game and Siegler and Ramani's. While our game was played by pairs of children with the teacher playing the role of facilitator, "The Great Race" is typically played between an experimenter and an individual child (e.g., see Ramani & Siegler, 2008). Thus, it is possible that children learn more from playing numerical board games when played with an attentive adult compared to a peer (but see Ramani et al. (2012) for evidence that "The Great Race" is also widely effective when played in small-group classroom settings). Lastly, a major limitation of the current study, and another reason we may not have observed the same gains as "The Great Race," has to do with the small sample size employed. With the exception of the improvements on the numeral identification task, Bayesian analyses indicated that more data are needed to determine sufficient evidence for or against the effectiveness of the intervention. Thus, the present results and comparisons to "The Great Race" must be interpreted with caution. A high powered follow-up study is needed to more extensively and conclusively evaluate the effects of the intervention.

Taken together, it is not immediately clear why our game may not have been as effective as other very similar number board games (Elofsson et al., 2016; Ramani & Siegler, 2008; Whyte & Bull, 2008). Indeed, from a theoretical standpoint, it remains unclear why the mechanisms underlying the success of “The Great Race” would not also underlie the performance of the present game. Moving forward, it may be of benefit to directly compare our game to Siegler and Ramani’s (2008) in an effort to potentially reveal key differences in the learning processes and outcomes between the two games. This approach may also provide further insight into the success of “The Great Race,” which has important educational implications for the design of effective number interventions moving forward.

Lessons Learned: Building Bi-Directional Relations

A central goal of MBE is to improve collaboration and knowledge exchange between professionals in education and the cognitive, developmental, and brain sciences. Although the advantages of such collaborations are many (e.g., see Ansari & Coch, 2006; Fischer, 2009; Fischer et al., 2007), examples of such collaborations are rare (e.g., see Samuels, 2009). One major obstacle to this effort are problems related to both perceived and structural hierarchies that currently exist between educational practice and research. For example, teacher-led RCTs have frequently been mentioned as one way of establishing improved contact between research and practice (e.g., see Gorard, See, & Siddiqui, 2017). However, the way RCTs are currently employed may be doing more harm than good, further contributing to hierarchical dominance of research over practice. Thomas, Ansari, and Knowland (2018) caution that instructional techniques derived from RCTs tend to be prescriptive—“*to be delivered by teachers as designed by researchers*” (Thomas et al., 2018, p. 487). However, the current teacher-led RCT (small as it may be) appears to be a notable exception. In reflecting on how the current RCT came to be, that is, bottom-up as opposed to top-down, we believe there are some valuable insights to be shared.

As researchers, our most fruitful teacher-researcher collaborations have emerged from models of PD that have built-in flexibility and teacher “degrees of freedom.” The opposite has been true when we have begun work with teachers with a more scientifically rigorous but rigid research agenda. Models of PD that honor and promote teacher agency and feedback, including the approaches of Design-Research or Japanese Lesson Study, stand to benefit all members (e.g., see Brown, 1992; Bruce & Hawes, 2015; Hawes et al., 2017; Moss, Hawes, Naqvi, & Caswell, 2015). Educators are given autonomy and contribute to the group’s learning as they fulfill the role of teacher-researchers. Researchers gain valuable insight into the realities of classroom learning and are given multiple opportunities to observe and collect data on student learning (Brown, 1992; Hawes et al., 2017). Most importantly, this approach allows relationships to form between practitioners and researchers. Once mutually respectful and beneficial relationships have been established, more rigorous scientific collaboration is possible. The current study provides evidence of this: By prioritizing relationships first and science second, our shared goal of improving children’s learning outcomes was realized through increasingly higher standards for conducting high quality scientific research. Paradoxically, over the long run, this approach may offer more sustainable and ongoing contributions to the interdisciplinary study of learning. As evidence of this, we are now in our fifth year of collaboration with one another. In sum, high-quality collaborations and “gold standard” research (RCTs), like the one reported on here, may be emergent properties of flexible models of teacher-researcher collaboration.

Conclusion

Our findings suggest that gains in children’s numeral identification skills can be achieved after a relatively short (45 min) number board game intervention. This finding is significant in so far as numeral identification skills play a critical role in more advanced numerical and mathematical reasoning. In addition to the intervention effects, the present study provides an example of successful teacher-researcher collaboration. We provide proof of concept that bidirectional relations are possible, demonstrating teacher contributions to psychological research and theory as well as evidence for the translation and application of research to practice (e.g., the work of Siegler; the methodological approach of RCTs). We credit this achievement to a model of teacher PD that prioritizes relationships first and science second. Once relationships are formed, opportunities for higher scientific rigor may follow.

ACKNOWLEDGMENTS—This research was funded in part by the Robertson Program for Inquiry-Based Teaching in Mathematics and Science as well as a Social Sciences and Social Sciences and Humanities Research Council of Canada (SSHRC) Doctoral Fellowship to ZH. We also would like to thank the Dr. Eric Jackman Institute of Child Study for their ongoing support in our partnership with the Rainy River District School Board. Relatedly, this research would not be possible without the support and commitment to engage in collaborative research by the educators and administration of the Rainy River District School Board.

REFERENCES

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E. J. (2015). An introduction to Bayesian hypothesis testing for management research. *Journal of Management*, *41*(2), 521–543.

Ansari, D., & Coch, D. (2006). Bridges over troubled waters: Education and cognitive neuroscience. *Trends in Cognitive Sciences*, *10*(4), 146–151.

Ansari, D., Coch, D., & De Smedt, B. (2011). Connecting education and cognitive neuroscience: Where will the journey take us? *Educational Philosophy and Theory*, *43*(1), 37–42.

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science*, *14*(1), 125–135.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Bruce, C. D., & Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: What is it? Why does it matter? And what can we do about it? *ZDM*, *47*(3), 331–343.

De Smedt, B., Ansari, D., Grabner, R. H., Hannula-Sormunen, M., Schneider, M., & Verschaffel, L. (2011). Cognitive neuroscience meets mathematics education: It takes two to tango. *Educational Research Review*, *6*(3), 232–237.

Elofsson, J., Gustafson, S., Samuelsson, J., & Träff, U. (2016). Playing number board games supports 5-year-old children's early mathematical development. *The Journal of Mathematical Behavior*, *43*, 134–147.

Fischer, K. W. (2009). Mind, brain, and education: Building a scientific groundwork for learning and teaching. *Mind, Brain, and Education*, *3*(1), 3–16.

Fischer, K. W., Daniel, D. B., Immordino-Yang, M. H., Stern, E., Battro, A., & Koizumi, H. (2007). Why mind, brain, and education? Why now? *Mind, Brain, and Education*, *1*(1), 1–2.

Gorard, S., See, B. H., & Siddiqui, N. (2017) *The trials of evidence-based education: The promises, opportunities and problems of trials in education*. New York, NY: Routledge.

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. *Cognition and Instruction*, *35*(3), 236–264.

Hawes, Z., Nosworthy, N., Archibald, L., & Ansari, D. (2019). Kindergarten children's symbolic number comparison skills predict 1st grade mathematics achievement: Evidence from a two-minute paper-and-pencil test. *Learning and Instruction*, *59*, 21–33.

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1), 2–9.

Lakoff, G., & Núñez, R. E. (2000) *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

Marghetis, T., Núñez, R., & Bergen, B. K. (2014). Doing arithmetic by hand: Hand movements during exact arithmetic reveal systematic, dynamic spatial processing. *The Quarterly Journal of Experimental Psychology*, *67*(8), 1579–1596.

Moss, J., Hawes, Z., Naqvi, S., & Caswell, B. (2015). Adapting Japanese lesson study to enhance the teaching and learning of geometry and spatial reasoning in early years classrooms: A case study. *ZDM*, *47*(3), 377–390.

Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A two-minute paper- and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS One*, *8*(7), e67918.

Okamoto, Y., & Case, R. (1996). II. Exploring the microstructure of children's central conceptual structures in the domain of number. *Monographs of the Society for Research in Child Development*, *61*(1–2), 27–58.

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology*, *105*(2), 453–464.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development*, *79*(2), 375–394.

Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology*, *32*(3), 146–159.

Ramani, G. B., Siegler, R. S., & Hitti, A. (2012). Taking it to the classroom: Number board games as a small group learning activity. *Journal of Educational Psychology*, *104*(3), 661–672.

Samuels, B. M. (2009). Can the differences between education and neuroscience be overcome by mind, brain, and education? *Mind, Brain, and Education*, *3*(1), 45–55.

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta- analysis. *Child Development*, *89*(5), 1467–1484.

Siegler, R. S. (2009). Improving the numerical understanding of children from low-income families. *Child Development Perspectives*, *3*(2), 118–124.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250.

Siegler, R. S., & Ramani, G. (2011). Improving low-income children's number sense. In S. Dehaene, & E. Brannon (Eds.), *Space, time, and number in the brain: Searching for the foundations of mathematical thought*. Attention and performance series (Vol. *XXIII*, pp. 343–354). London, England: Academic Press.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low- income children's numerical development. *Developmental Science*, *11*(5), 655–661.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—But not circular ones—Improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*(3), 545–572.

Thomas, M. S., Ansari, D., & Knowland, V. C. (2018). Annual research review: Educational neuroscience: Progress and prospects. *Journal of Child Psychology and Psychiatry*, *60*(4), 477–492.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology*, *44*(2), 588–596.