

This is a post-peer-review, pre-copyedit version of an article published in *Cognitive Processing*. The final authenticated version is available online at: <http://dx.doi.org/10.1007/s10339-018-0856-7>

**Numbers in Action:
Individual Differences and Interactivity in Mental Arithmetic**

Lisa G. Guthrie and Frédéric Vallée-Tourangeau
Kingston University

Author Note

Address correspondence to Lisa G. Guthrie or Frédéric Vallée-Tourangeau, Department of Psychology, Kingston University, Kingston upon Thames, Surrey, UNITED KINGDOM, KT1 2EE; l.guthrie@bath.ac.uk , f.vallee-tourangeau@kingston.ac.uk. We thank Charlotte Harris for her contributions to the design of the research reported as well as her help recruiting and testing participants. We are grateful for the thoughtful comments from Mark Ashcraft and another anonymous reviewer on a previous version of this manuscript.

Abstract

Previous research indicates that interactive arithmetic tasks may alleviate the deleterious impact of maths anxiety on arithmetic performance. Our aim here was to further test the impact of interactivity on maths anxious individuals and those with poorer numeracy skills. In the experiment reported here participants completed sums in two interactivity contexts. In a low interactivity condition sums were completed with hands down. In a second, high interactivity condition, participants used moveable number tokens. As anticipated, accuracy and efficiency were greater in the high compared to the low interactivity condition. Correlational analyses indicated that maths anxiety, objective numeracy, measures of maths expertise and working memory were stronger predictors of performance in the low than in the high interactivity conditions. Interactivity transformed the deployment of arithmetic skills, improved performance, and reduced the gap between high and low ability individuals. These findings suggest that traditional psychometric efforts that identify the cognitive capacities and dispositions involved in mental arithmetic should take into account the degree of interactivity afforded by the task environment.

Keywords: Interactivity; mental arithmetic; expertise; maths anxiety; working memory

Numbers in Action:
Individual Differences and Interactivity in Mental Arithmetic

Anxiety is typically associated with feelings of panic, tension, uneasiness, helplessness and mental disorganisation when an individual is faced with an uncomfortable situation (Ashcraft, 2002; Núñez-Peña, Suárez-Pellicioni, & Bono, 2013). These feelings can be triggered when maths-anxious individuals are faced with solving maths problems either in the classroom, workplace, or daily life (Ashcraft & Kirk, 2001; Ashcraft & Ridley, 2005). This maths anxiety has the potential to increase the burden on working memory by reducing both storage and processing capabilities, impacting performance on a range of maths tasks (Ashcraft, 2002; Beilock & Carr, 2005). In addition, performance anxiety brought on by time pressure is more likely to affect those with high working memory capacities (Beilock & Carr, 2005; Ramirez, Gunderson, Levine, & Beilock, 2013). Maths anxious people have repeatedly been shown to perform more poorly in problems involving maths than their less anxious counterparts (Hembree, 1990; Lyons & Beilock, 2011; Ma, 1999). In a classroom environment this can lead to feelings of inadequacy and negative attitudes toward maths (Núñez-Peña et al., 2013). Maths anxious students may actively avoid making curriculum and ultimately career choices that involve mathematics (Núñez-Peña et al., 2013).

However, the true mathematical ability of an individual may be masked by the impact of the anxiety experienced when presented with a maths problem (Hoffman, 2010; Ramirez et al. 2013). Previous research indicates that interactive arithmetic tasks may alleviate the deleterious impact of maths anxiety on arithmetic performance (Vallée-Tourangeau, 2013). The aim of the experiment presented here

was to determine how interactivity may also improve performance for participants with low numeracy and lower levels of maths expertise. And while participants in the present experiment were enjoined to complete the mental arithmetic task as quickly and as accurately as possible, time pressure was not explicitly manipulated or combined factorially with other variables (cf. Beilock & Carr, 2005).

Solving a mental arithmetic problem can place demands on limited working memory storage capacity and processing (DeStefano & LeFevre, 2004; Butterworth, 2006). When internal cognitive resources are strained, people naturally mine their external surroundings in order to augment cognition (Kirsh, 2017). During mental arithmetic, individuals, adults and children alike, may use gestures, and fingers to point and count (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Carlson, Avraamides, Cary, & Strasberg, 2007). Gesturing contributes to counting accuracy in children (Alibali & DiRusso, 1999): touching items when counting facilitates more accurate performance than simply pointing to countable items. Children learn to calculate by using their fingers in conjunction with repeating the names of the numbers aloud (Butterworth, 2005). People frequently use and manipulate items around them to complete common maths tasks. For example, when counting a handful of change, it may be useful to lay out the coins, and then grouping common coins together during the tallying. These epistemic actions change the state of the computation: the fashioning of congenial interim sums (e.g., 10 or 50) may cue more efficient calculation strategies that can reduce the load on working memory as some of the limited internal memory storage is unburdened onto the external world (Kirsh, 1995; Vallée-Tourangeau, 2013).

Actions and the manipulation of artefacts during problem solving may provide new information and unveil new affordances; the dynamic agent-environment

coupling configures a cognitive system that functionally enhances working memory resources. In a study investigating maths anxiety, working memory and interactivity, Vallée-Tourangeau, Sirota, and Villejoubert (2013) asked participants to complete sums in two conditions: a low interactivity hands down condition, or a high interactivity condition by moving numbered tokens. They found maths anxiety to be highly correlated with calculation error in a low interactivity condition where participants could not modify the problem presentation nor use their hands to point at numbers; however, in a high interactivity condition where participants could shape and reshape the problem presentation, maths anxiety was no longer a predictor of performance.

Vallée-Tourangeau, Sirota, and Vallée-Tourangeau (2016), using similar low and high interactivity conditions, tested the impact of articulatory suppression on the calculation of 11-digit sums. Their results indicated that an increased level of interactivity reduced the impact of maths anxiety on performance. In addition, any deterioration in performance as a result of articulatory suppression was mitigated when participants moved the tokens in comparison to when they were prevented from doing so. In the low interactivity condition in these two experiments numbers for each addition were presented to participants on paper, all at once, in a random pattern. Calculation of these additions requires the retrieval of maths knowledge from long-term memory to evaluate which numbers might be added to create favourable subtotals. Interim totals are calculated and stored for short-term recall, while attention must be paid to numbers that have been added and those that have not. Therefore, even for these simple additions, considerable demands are placed on executive function skills and draw on storage capacity. Vallée-Tourangeau et al. (2013, 2016) argued that in the higher interactivity condition, a dynamic problem

presentation wrought through action transforms working memory capacity, not only in terms of storage but also executive function skills, mitigating the impact of performance anxiety.

Engagement and Expertise

The experience of learning and achievement are potentially influenced by active engagement in the performance of academic tasks in the classroom (Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003). A greater sense of engagement may result from a perception of control and relevance to the real world (Shernoff et al., 2003; Newmann, Wehlage, & Lamborn, 1992). Affective variables such as enjoyment, interest and challenge have been associated with academic success, thus positive emotions elicited by the task experience contribute to increased problem-solving capacities and improved mathematical performance (Hembree, 1990; Shernoff et al., 2003). In turn, difficulty in performing tasks may be experienced as a result of negative affect (Storbeck & Clore, 2007). Increasing the level of interactivity when solving a maths problem has been shown to positively impact the level of engagement (Guthrie & Vallée-Tourangeau, 2015). This implies that giving participants control over their environment, through a higher degree of interactivity, may directly increase affect and engagement in the task compared to the level of engagement in a low interactivity environment.

A number of factors have been identified as contributors to exceptional mathematical performance including working memory, deliberate practice, and intrinsic reward (Butterworth, 2006; Ericsson & Charness, 1994). Greater working memory capacity may explain the proficiency of an individual to manipulate a greater amount of information, for example, in the ability to retain interim totals for additions while completing a maths task. High levels of performance and numeracy may also

reflect the ongoing acquisition and consolidation of skills through prolonged exposure and deliberate practice (Ericsson & Charness, 1994; Sternberg, 1999). It has also been suggested that expertise may be acquired through the enjoyment and emergent reward experienced when unraveling solutions to mathematical problems. This intrinsic motivation may encourage progression to increasingly more challenging tasks, resulting in greater knowledge and expertise (Butterworth, 2006; Csikszentmihalyi, 1978). While these factors may contribute to increased expertise in maths, it may be the case that avoidance of maths and maths-related subjects in school and in subsequent careers by maths anxious individuals may lead to a reduction in maths expertise and general numeracy (Ramirez et al., 2013).

The Current Experiment

A task ecology that promotes interactivity with a physical presentation of a mental arithmetic problem has been shown to increase efficiency and accuracy, while reducing absolute calculation error (Vallée-Tourangeau, 2013). In the current experiment, participants were invited to complete simple sums. The maximum sum length in a similar experiment by Vallée-Tourangeau et al. (2013) was 11 digits. Here the maximum sum length was increased, with additions composed of either 11 or 17 single-digit numbers. While this task should not challenge the arithmetic knowledge and skills of university-educated participants; in the absence of pen and paper, calculations would require good working memory capacity and executive function skills especially when dealing with the longer sums.

The mental arithmetic task was presented in two different contexts. In one, the low-interactivity context, participants were shown a random configuration of numbers, and were asked to calculate the sum with hands on the table. In the second, a high-interactivity context, the same configurations were presented with

numbered tokens, and participants were free to move them and re-arrange the problem presentation as they calculated an answer. The participants were also profiled in terms of a number of individual differences implicated in mental arithmetic including maths anxiety, working memory, numeracy, maths expertise and engagement in the task. All participants completed the short and long sums in both interactivity conditions. Therefore this within subjects design reduced the attribution of any changes in performance to between-subjects individual differences.

A dynamic, high interactivity environment using artefacts as opposed to a low interactivity one may encourage more efficient calculations, the enactment of better arithmetic skills, through the dynamic reconfiguration of the problem, and hence lead to improved mental arithmetic performance. The high-interactivity condition enables participants to configure the problem in a manner that facilitates computation reducing the load on working memory, which in turn should free up internal resources facilitating an improvement in performance for maths anxious participants (Ashcraft & Kirk, 2001; Vallée-Tourangeau, 2013). Thus, the spatial arrangement of the numbers that compose the problem may be re-shaped to segregate numbers that have been added, clearly identifying those that remain to be processed. In addition, the dynamic reconfiguration of the problem potentially cues arithmetic knowledge facilitating the recognition and composition of congenial groupings (e.g., $7 + 8 = 15$) that may conveniently complement the interim total, enhancing accuracy and efficiency. As a result, a higher degree of interactivity may functionally extend the storage capacity of a reasoner, as well as enable her to express more efficiently her arithmetic knowledge.

The predicted improvement in performance in a high interactivity task environment should therefore narrow the gap between reasoners with lower

cognitive capacities, numeracy and maths expertise and those who score higher on these dimensions. As a result, while measures of individual differences should correlate strongly with performance in a low interactivity condition, we predicted that they should do so to a much lesser degree in the condition that afforded a higher degree of interactivity.

Method

Participants

Sixty participants (38 women, $M_{\text{age}} = 21.3$, $SD = 2.37$) were recruited from various academic backgrounds. Thirty-two psychology undergraduates participated in exchange for credits. Twenty-one undergraduates from other disciplines and seven additional participants either working in a highly numerical field (e.g., accounting) or recently graduated with a maths-related degree participated voluntarily.

Materials and Measures

Arithmetic Task. Each participant was presented with two sets of additions, one set for each interactivity condition. Both sets were composed of ten sums – five sums of 11 single-digit numbers and five sums of 17 single-digit numbers. In the low interactivity condition, participants were given a sheet of A4 paper, with numbers to be summed distributed randomly on the page (see Figure 1, left panel). While adding the numbers, participants were instructed to keep their hands flat on the table. In the high interactivity condition, participants were given the same sets of sums in a similar configuration but presented as moveable numbered wooden tokens (1.2 cm in diameter; see Figure 1, right panel). On completing each sum participants announced the answer to the researcher.

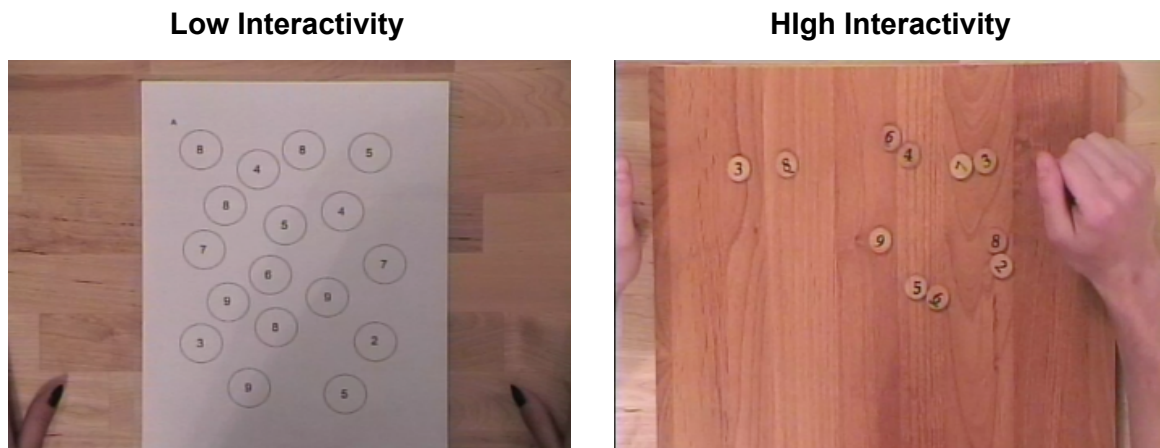


Figure 1. Each sum was presented on a sheet of A4 as a random configuration of digits in the low interactivity condition (left panel); participants in that condition kept their hands flat on the tabletop. In the high interactivity condition, the same sums were presented with movable wooden tokens (right panel) which participants touched, moved, grouped, as they saw fit.

Maths Anxiety. Participants completed a 25-item Mathematics Anxiety Scale-UK (MAS-UK; Hunt, Clark-Carter & Sheffield, 2011). They were asked to indicate, on a 5-point Likert scale (from 1 = “not at all” to 5 = “very much”) how anxious they would feel in certain situations. Items included statements such as “Adding up a pile of change” or “Taking a math exam” ($M = 41.6$, $SD = 9.50$). The MAS scores were normally distributed, $D(60) = .117$, $p = .200$; the scale showed excellent reliability, Cronbach’s $\alpha = .93$.

Objective Numeracy. A basic arithmetic scale (BAS) was used to test participants’ objective numeracy. It consisted of 60 simple arithmetic problems (e.g., $7 \times 8 = ?$). Participants wrote the answers next to the problem on the paper provided, in the order presented, completing as many as possible in 60 seconds.

Working Memory. Participants completed two working memory tasks. The computation-span task tested both processing and storage of numbers, while a non-numerical visuo-spatial task, the Corsi block task, tested the temporary storage of visuospatial information.

Computation-span task. The computation-span task (adapted from Ashcraft & Kirk, 2001) required participants to answer simple arithmetic problems (e.g., $2 + 8 = ?$, $12 - 4 = ?$) before recalling the second number of these problems (e.g., 8, 4). Sequences of equations ranged from 1 – 7 and participants had to process each arithmetic expression *and* recall the relevant digit correctly to score. This was presented on a computer screen. Instructions and two practice sequences were completed before the task began. Each expression was presented for 2 seconds, for which the participants had to provide an answer; a recall page appeared at the end of each sequence. This was the prompt for participants to recall the second number in each of the expression as quickly as possible. The next sequence appeared on the screen following a mouse click by the participant. This continued until all sequences were completed. A point for each correct response was given from a maximum total of 56 ($M = 23.6$, $SD = 8.8$).

Corsi block task. In this version of the Corsi Block task (Corsi, 1972) participants were shown ten sequences of shaded blocks in a 4 x 4 matrix on a computer screen. The number of blocks to be remembered in each sequence increased from 2 to 6 blocks in length. Participants scored one point for each correctly remembered block location, thus the maximum score was 40 ($M = 33.2$, $SD = 4.4$).

Maths Expertise. A questionnaire was developed to evaluate maths expertise. Three questions were related to maths grades in high school with a further three questions on university courses and current employment where applicable. The scoring comprised of three components. First, one point was allocated for each level of maths completed in high school; second, a score from 0 (no grade), to 4 (A

or A*¹) was given for grades attained; finally a score of 1 (no degree) to 4 (maths-related degree or job) was allocated for the maths weighting of a degree or employment. These scores were aggregated to provide a continuous numerical measure of maths expertise. The maximum score possible was 19 ($M = 11.3$; $SD = 5.0$).

Task Engagement Scale. The Task Engagement Scale (TES) was developed to gauge a participant's engagement and enjoyment during a task. The 9-item scale was based on three key components of task engagement identified by Shernoff et al. (2003): concentration, enjoyment, and interest. Participants were asked to rate such items as how bored, challenged and relaxed they felt while undertaking the task. Each item was scored on an 8-point Likert scale, labelled from 0 (definitely not) to 7 (definitely yes): The higher the score the more positive the attitude toward the task. Each participant completed the TES scale twice, once following each of the two sets of sums across interactivity conditions. The reliability of the nine-item scale was good (Cronbach's $\alpha = .80$).

Procedure

The length of the additions (11 or 17 digits) and level of interactivity (high or low) were repeated measures factors in a 2x2 design yielding four conditions. The presentation order of these conditions was counterbalanced across participants. The sets of sums for each interactivity condition were separated by at least one other task (either the MAS-UK, BAS, Computation-span or Corsi Block). The other tasks were presented at either the beginning or the end of the session with the order counterbalanced across participants. Each experimental condition was followed by the TES, and the experiment ended with the maths experience questionnaire. The

¹ The top grade achievable for mathematics in the English education system.

working memory tasks were presented on a computer with all other tasks presented on paper. The experimental session lasted approximately one hour.

Mental arithmetic performance was measured in terms of accuracy (proportion of sums correct), latency to solution, absolute deviation error (ADE) and efficiency. ADE was defined as the absolute error from the correct answer. Therefore, if the participant answered 52 and the correct answer was 50, the ADE would be two; if the participant answered 48 the ADE would also be 2. A participant's efficiency was his or her proportion of correct answers over the proportion of time used to calculate a set of sums. This proportion was derived by taking the participant's mean latency to calculate a set of five sums in a given condition divided by the mean latency of the slowest quartile of participants in that condition. Thus, if a participant's accuracy for a series of five sums was .6, and her average latency to complete these sums was 40% of the average of the slowest participants, then her efficiency ratio would be $.6/.4$ or 1.5. Ratios at or above 1 reflect efficient reasoning; ratios below 1 reflect inefficient reasoning.

Results

Overall Arithmetic Performance

Accuracy. The mean percent correct, illustrated in the top left panel of Figure 2, was greater in the high interactivity condition than the low interactivity condition for both sum lengths. A 2 (Interactivity: Low and high) x 2 (Sum length: 11-digit and 17-digit) repeated measures analysis of variance (ANOVA) revealed a significant main effect of interactivity, $F(1, 59) = 30.0, p < .001, \eta_p^2 = .34$ and sum length $F(1, 59) = 21.2, p < .001, \eta_p^2 = .265$; however, the interaction was not significant, $F < 1$.

Absolute Deviation Error. The mean ADE (Figure 2, bottom left panel) was lower in the high interactivity condition than in the low interactivity condition

regardless of the sum length. A 2 x 2 repeated measures ANOVA indicated a significant main effect of interactivity, $F(1, 59) = 11.0, p = .002, \eta_p^2 = .16$ and sum length $F(1, 59) = 17.2, p < .001, \eta_p^2 = .23$. However, there was no significant interaction, $F < 1$.

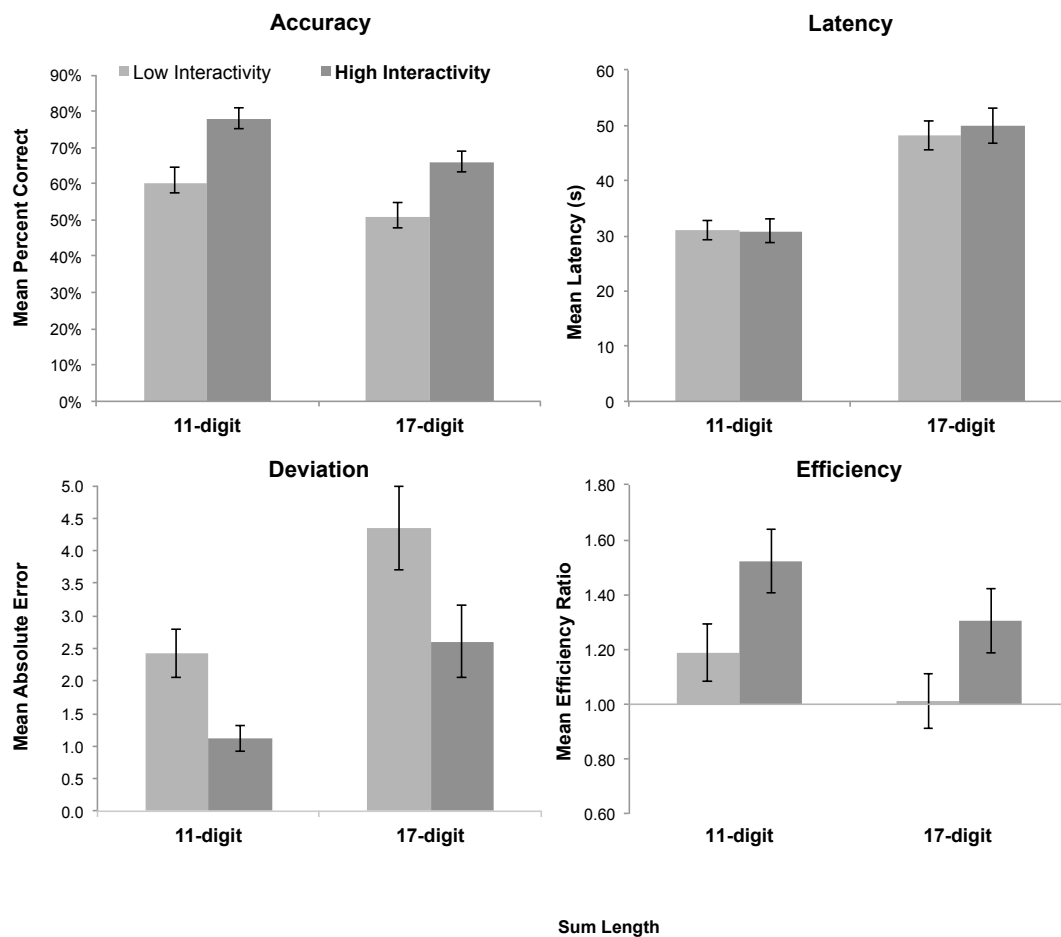


Figure 2. Mean percent correct (top left), mean latency (top right), mean absolute deviation error (bottom left), and mean calculation efficiency (bottom right) as a function of sum length (11-digit and 17-digit sums) in the low (light grey bars) and high (dark grey bars) interactivity condition. Error bars are standard errors of the mean.

Latency. While latency to completion was influenced by sum length, interactivity level resulted in very little difference in latency (see Figure 2, top right panel). In a 2 x 2 repeated measures ANOVA the main effect of interactivity was not significant, $F(1, 59) = 1.42, p = .239, \eta_p^2 = .02$; however, there was a significant

main effect of sum length $F(1, 59) = 201, p < .001, \eta_p^2 = .78$ and a significant interaction between sum length and condition $F(1, 59) = 6.68, p = .012, \eta_p^2 = .10$. Post hoc tests revealed that the interaction was driven by the fact that the mean latencies did not differ significantly between interactivity conditions for the shorter sums ($M_{low11} = 31.0, SD_{low11} = 13.8, M_{high11} = 30.9, SD_{high11} = 16.2, t < 1$) but the difference was marginally significant for the longer sums ($M_{low17} = 48.1, SD_{low17} = 20.3, M_{high17} = 51.0, SD_{high17} = 25.7, t(59) = -1.811, p = .075$).

Efficiency. Participants were less efficient when calculating the sums in the low interactivity condition than when using tokens across both sets of sums (see Figure 2, bottom right panel). The efficiency ratio decreased for longer sums, although it was still larger in the high interactivity condition. A 2 x 2 repeated measures ANOVA indicated a significant main effect of interactivity, $F(1, 59) = 22.0, p < .001, \eta_p^2 = .27$ and sum length $F(1, 59) = 17.1, p < .001, \eta_p^2 = .225$; the interaction, however, was not significant, $F < 1$.

Individual Differences

We profiled participants in terms of their working memory capacity (measured with a computation span and a Corsi block test), numeracy, maths expertise and maths anxiety. The measures of working memory were correlated, $r = .40, p = .002$, indicating that both tests measured overlapping aspects of working memory capacity ($df = 58$ for all correlation coefficients). Our measure of expertise correlated highly with both measures of working memory (computation span, $r = .59, p < .001$; Corsi, $r = .36, p = .005$). As expected, maths expertise was negatively correlated with maths anxiety, $r = -.68, p < .001$, and positively correlated with numeracy, $r = .65, p < .001$ (Ericsson & Charness, 1994; Sternberg, 1999; Ramirez et al., 2013). In line with previous research (e.g., Ashcraft, 2002), maths anxiety was correlated with the

computation span, $r = -.47$, $p < .001$, although the correlation with the Corsi test, $r = -.25$, $p = .055$, was only marginally significant, which might be explained by the fact that this working memory test did not involve the storage and processing of numbers.

However, of greater interest was the degree to which these individual differences correlated with mean arithmetic performance in the low and high interactivity conditions. Of fundamental interest is the impact of interactivity on the participants' degree of calculation accuracy. Hence, for the correlational analyses reported below, the absolute deviation error was selected as the preferred measure of performance as it offered a finer grained capture of participants' calculation accuracy.

We determined the degree of correlation between each of these measures of individual differences and absolute deviation error for the short and long sums in the low and high interactivity conditions. We anticipated that these individual differences would better predict calculation error in the low than in the high interactivity condition. These correlations, and a z test of their difference, are reported in Table 1. Examining Table 1, we first note that of the 10 correlations with calculation error in the low interactivity condition across both sum lengths, 9 were significant; in contrast, 3 of the 10 correlations were significant in the high interactivity conditions. In addition, for each of the five measures of individual differences, the correlations were always larger in the low than in the high interactivity condition, and this for both the short and long sums. The largest difference between correlations with calculation error involved maths expertise ($z = 2.23$, $p = .010$) and the smallest involved Corsi scores ($z = 0.21$, $p = .417$), both for the longer 17-digit sums. Using an exact sign test, this pattern of correlations—that is 10 out of 10 coefficients being larger in the low than in the high interactivity condition—was significant, $z = 3.16$, $p = .002$.

Table 1

Summary of correlations for absolute deviation errors and individual difference measures in both interactivity conditions ($df = 58$) with z-scores (Fisher's r to z transformation) and associated p -values

| Predictor | 11-digit Sums | | | | | |
|-----------|---------------|--------------|-------|--------------|------------|--------------|
| | Low | | High | | Difference | |
| | r | 95% CI | r | 95% CI | z | p (1 tail) |
| MAS | .49 ** | [.27, .66] | .28 * | [.03, .50] | 1.33 | .092 |
| OBJ-N | -.43 ** | [-.61, -.20] | -.12 | [-.36, .14] | 1.81 | .035 |
| C-Span | -.49 ** | [-.66, -.27] | -.17 | [-.40, .09] | 1.95 | .026 |
| Corsi | -.31 * | [-.52, -.06] | -.24 | [-.46, -.01] | 0.40 | .345 |
| Expertise | -.46 ** | [-.64, -.23] | -.13 | [-.37, .13] | 1.96 | .025 |

| Predictor | 17-digit Sums | | | | | |
|-----------|---------------|--------------|--------|--------------|------------|--------------|
| | Low | | High | | Difference | |
| | r | 95% CI | r | 95% CI | z | p (1 tail) |
| MAS | .44 ** | [.21, .62] | .16 | [-.10, .39] | 1.66 | .049 |
| OBJ-N | -.43 ** | [-.61, -.20] | -.28 * | [-.50, -.03] | 0.92 | .178 |
| C-Span | -.41 ** | [-.60, -.17] | -.28 * | [-.50, -.03] | 0.79 | .215 |
| Corsi | -.06 | [-.31, .20] | -.02 | [-.27, .23] | 0.21 | .417 |
| Expertise | -.46 ** | [-.64, -.23] | -.06 | [-.31, .20] | 2.33 | .010 |

Note: **MAS** = Maths anxiety; **OBJ-N** = Objective numeracy (basic arithmetic skill); **C-Span** = Computation-span task assessing working memory capacity; **Corsi** = adapted Corsi block task assessing visuospatial working memory; **Expertise** = Maths expertise (continuous measure); **Low** = Low interactivity; **High** = High interactivity. CI = confidence interval. * $p < .05$, ** $p < .01$.

A reviewer of a previous version of this manuscript queried whether the absence of correlation between calculation error and maths expertise with interactivity was because participants with low expertise got better or because participants with higher maths expertise somehow were adversely affected by a higher degree of interactivity and their performance was closer to the poorer performance of the participants with lower maths expertise. To address this query, we computed the change in calculation error with and without interactivity and correlated the resulting value with level of maths expertise, for short and long sums. As participants generally made more substantial calculation errors without interactivity, the change values were generally positive (e.g., it was generally the

case that the low interactivity error was greater than the high interactivity error). We plotted these change values (the more positive the change, the more participants improved in the high interactivity condition) against maths expertise scores (see Figure 3). The resulting correlations were significantly negative, $r = -.418, p = .001$, and $r = -.426, p = .001$ for the 11-digit and 17-digit sums respectively. What these correlation coefficients indicate is that it was the participants with lower levels of expertise who improved the most with interactivity, and this for both sum lengths. Thus the absence of correlation between maths expertise and calculation error in the high interactivity condition is attributable to the improvement in accuracy among participants with lower levels of maths expertise.

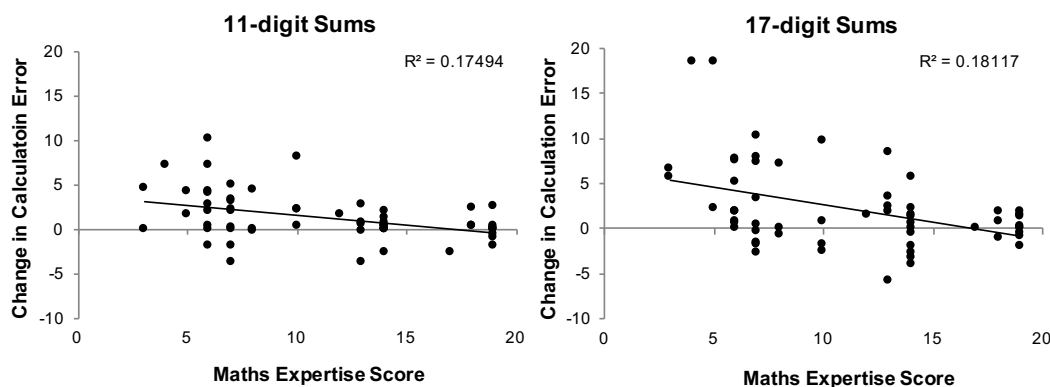


Figure 3. Change in absolute calculation error derived from the difference in the average absolute deviation from the correct answer for the sums in the low interactivity condition and the high interaction condition (low interactivity performance minus high interactivity performance)—the larger the change, the more the participant’s performance improved in the high interactivity condition—as a function of maths expertise score for the 11-digit sums (left panel) and 17-digit sums (right panel).

Task Engagement Scale

Participants were more engaged in the high interactivity condition ($M = 44.1$, $SD = 9.2$) than in the low interactivity condition ($M = 37.8$, $SD = 8.8$). This difference was significant, $t(59) = -6.16, p < .001$. There were no significant correlations

between the task engagement scale (TES) and expertise (low, $r = .147$, $p = .264$; high, $r = .023$, $p = .863$), or the TES and the average deviation for the 11-digit sums or the 17-digit sums (low, $r = -.231$, $p = .076$; high, $r = .175$, $p = .182$).

Qualitative Observations on Actions in the High Interactivity Condition

Participants' instructions in the high interactivity condition read: "In completing the additions you can move the tokens in any way you want to help you calculate the sum of numbers". If the participant asked whether he or she must move the tokens the researcher always replied that it was the participant's choice. Although participants' movement of the tokens in the high interactivity condition were not systematically analysed, researchers noted that all participants moved the tokens rather than simply touching or pointing at them during the mental arithmetic task. These informal observations are confirmed on the basis of a small subset of participants ($n = 11$) whose session was video recorded, as a means to check procedural consistency across participants: All participants in these videos moved the tokens in calculating the sums in the high interactivity condition. The recordings showed some participants moving the numbered tokens into groups creating congenial sums, such as grouping 5, 5, 2, and 3 to produce a total of 15. As the tokens were moved about on the board, the rearrangement appeared to provide the participants with different pathways for determining the answer. This was consistent with recorded observations for a similar task in Guthrie and Vallée-Tourangeau (2015) as well as in Vallée-Tourangeau (2013). Other participants were observed moving tokens to one side, once added to the running total, with no apparent strategic grouping or pattern. A few of the participants created lines of tokens to facilitate scanning.

Discussion

The main objective of this experiment was to explore the possible change to mental arithmetic performance for individuals with varying maths abilities and levels of anxiety when maths problems were presented in differing reasoning contexts. Participants completed two sets of addition problems: one set was completed with restricted hand movement reducing interactivity; the other using round numbered wooden tokens increasing the opportunity to reconfigure the problem presentation as the sum was calculated. Notably, all participants chose to move the tokens when calculating the sums in the high interactivity condition. Generally, participants answered more sums accurately, made smaller errors, and performed the calculations more efficiently in the high than in the low interactivity condition. Latency however, remained constant across the two levels of interactivity for the short and long additions, suggesting improvements in other measures were related to the mode of problem solving, rather than the time required to complete the problems. This improvement in performance could not be attributed to extraneous between-subject factors because of the repeated measures design employed in this experiment: all participants completed the sums in both interactivity conditions. The results support the claim that a high degree of interactivity improves the performance of those with less maths expertise for these simple arithmetic problems.

The strong correlation between objective numeracy and expertise indicated that our measure of expertise was an acceptable measure of the arithmetic proficiency of an individual. With a static problem presentation and hands down on the table, participants' performance reflected their arithmetic skills and working memory capacity. A high degree of interactivity improved the performance of participants with lower maths expertise. The absence of correlation between

arithmetic performance and expertise in the high interactivity condition implied that the manipulation of number tokens augmented the arithmetic skills of participants with less maths expertise.

The influence of maths anxiety on performance was different as a function of interactivity, especially with the longer sums. When interactivity with the problem was low, maths anxiety had a significant impact on performance for both short and long sums. Maths anxiety was correlated significantly with performance for the short sums in the high interactivity condition, but for the more demanding long sums, maths anxiety was no longer a significant predictor of performance in the high interactivity condition. These findings lend some support to the thesis advanced here, namely that a dynamic presentation that offers a greater level of interactivity may assist in reducing or controlling the impact of maths anxiety. A higher degree of interactivity exploits the dynamic changes in the problem presentation to scaffold storage capacity; the spatial re-arrangements of the tokens may also facilitate the expression of arithmetic knowledge by helping participants more efficiently recognize groupings that facilitate computation (Vallée-Tourangeau, 2013). Interpreted from the perspective of processing efficiency (Ashcraft, 2002), a participant's anxious rumination exacts a proportionally smaller toll on resources when those resources can be augmented through interactivity.

The two measures of working memory, computation-span and the Corsi block task, were moderately correlated. Computation-span correlated highly with numeracy and expertise supporting claims that working memory is a contributing factor to mental arithmetic skill (see Butterworth, 2006). Our computation-span test was designed to reflect a conventional complex span task requiring some numerical skills; unsurprisingly, this correlated with mental arithmetic performance in the low

interactivity condition, more interestingly it only weakly correlated with performance in the high 17-digit interactivity condition. The Corsi task, as a measure of visuospatial working memory was deliberately selected to reduce the reliance on numeracy. With the exception of a marginally significant correlation for short sums in the low interactivity condition, Corsi scores did not predict mental arithmetic performance. Span tasks, such as the computation-span assess an individual's working memory in both processing and storage, whereas the Corsi test as designed here gauges storage capacity of visuospatial information only. These patterns of correlations indicate that high interactivity does not simply function as a means for off-loading working memory storage. It also helps participants deploy improved executive function skills. This is because the allocation of attentional resources is transformed by the physical changes to the problem presentation. These changes help participants look at the problem differently and may thus improve their ability to perceive groupings that cue long term arithmetic knowledge. In addition, participants can exploit the physical space to segregate tokens that have been processed from those that have not, which may make it easier for them to identify and combine congenial groupings and update the running interim total. Thus enhanced storage and dynamic perceptual feedback may scaffold executive function skills synergistically (see also McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010).

Individuals were more engaged in the task when given the opportunity to use the tokens than when they had to maintain their hands on the table. However, the level of engagement did not change as a function of maths expertise as indicated by the lack of correlations between the TES and expertise scores. Notably performance, as measured by absolute deviation error, was also not influenced by how engaged participants were in the task. Participants might have felt more engaged when

completing the task with tokens, but the level of engagement did not in itself explain the improvement in arithmetic performance in the high interactivity condition.

Expertise in the domain of mathematics may be attributable to factors including practice, intrinsic reward and components of working memory. Here we have shown that a systemic perspective on mental arithmetic reveals how resources internal and external to participants can be configured dynamically to better reflect abilities in solving simple mathematical problems. These findings also encourage us to reflect critically on the importance of the physical context of reasoning in mapping the psychometric predictors of performance in mental arithmetic.

References

- Alibali, M.W., & DiRusso, A.A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development, 14*, 37-56. doi: 10.1007/s11031-011-9252-7
- Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science, 11*, 181-185. doi: 10.1111/1467-8721.00196
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*, 224-237. doi: 10.1037/0096-3445.130.2.224
- Ashcraft, M. H., & Ridley, K. S. (2005) Math anxiety and its cognitive consequences: A tutorial review. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 315-327). New York: Psychology Press.
- Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and “choking under pressure” in math. *Psychological science, 16*, 101-105. doi: 10.1111/j.0956-7976.2005.00789.x
- Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology & Psychiatry, 46*, 3-18. doi: 10.1111/j.1469-7610.2004.00374
- Butterworth, B. (2006). Mathematical Expertise. In K.A. Ericsson, N. Charness, P.J. Feltovich, & R.R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 553-568). New York, NY: Cambridge University Press.
- Carlson, R. A., Avraamides, M. N., Cary, M., & Strasberg, S. (2007). What do the hands externalize in simple arithmetic? *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 33, 747–756. doi: 10.1037/0278-7393.33.4.747

Csikszentmihalyi, M. (1978). Intrinsic rewards and emergent motivation. In M.R. Lepper, & D. Greene (Eds.). *The hidden costs of reward: New perspectives on the psychology of human motivation* (pp. 205-216). Hillsdale, NJ: Erlbaum Associates.

Corsi, P. M. (1972). *Memory and the medial temporal region of the brain* (Unpublished doctoral dissertation). McGill University, Montréal, QC, Canada.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, 16, 353-386. doi: 10.1080/09541440244000328

Ericsson, K.A. & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, 49, 725-747. doi: 10.1037/0003-066X.49.8.725

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, 12, 516-522. doi: 10.1111/1467-9280.00395

Guthrie, L.G., & Vallée-Tourangeau, F. (2015). Interactivity and mental arithmetic: Coupling mind and world transforms and enhances performance. *Studies in Logic, Grammar and Rhetoric*, 41, 41-59. doi: 10.1515/slgr-2015-0019

Hembree, R. (1990). The nature effects and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21, 33-46. doi: 10.2307/749455

Hoffman, B. (2010). “I think I can, but I’m afraid to try”: The role of self-efficacy beliefs and mathematics anxiety in mathematics problem-solving efficiency.

Learning and Individual Differences, 20, 276-283. doi:

10.1016/j.lindif.2010.02.001

Hunt, T. E., Clark-Carter, D., & Sheffield, D. (2011) The development and part validation of a UK scale for mathematics anxiety. *Journal of Psychoeducational Assessment*, 29, 455-466. doi: 10.1177/0734282910392892

Kirsh, D. (1995). Complementary strategies: Why we use our hands when we think. In J. M. Moore, & J. L. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 212-217). Mahwah, NJ: Lawrence Erlbaum Associates.

Kirsh, D. (2017). Thinking with external representation. In S. J. Cowley, & F. Vallée-Tourangeau (Eds.), *Cognition beyond the brain: Computation, interactivity and human artifice* (Second edition, pp. 61-84). London: Springer-Verlag.

Lee, K., Ng, S-F., Ng, E-L. & Lim, Z-Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology*, 89, 140-158. doi: 10.1016/j.jecp.2004.07.001

Lyons, I.M. & Beilock, S.L. (2011). Mathematics anxiety: Separating the math from the anxiety. *Cerebral Cortex*, 22, 2102-2110. doi: 10.1093/cercor/bhr289

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, 30, 520-540. doi:10.2307/749772

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, 24, 222. doi: 10.1037/a0017619

- Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. F. M. Newmann (Ed.) *Student engagement and achievement in American secondary schools* (pp. 11-39). New York: Teachers College Press Columbia University.
- Núñez-Peña, M. I., Suárez-Pellicioni, M., & Bono, R. (2013). Effects of math anxiety on student success in higher education. *International Journal of Educational Research*, 58, 36-43. doi: 10.1016/j.ijer.2012.12.004
- Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, 14, 187-202. doi: 10.1080/15248372.2012.664593
- Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18, 158-176. doi: 10.1521/scpq.18.2.158.21860
- Sternberg, R.J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, 24, 359–375. doi: 10.1006/ceps.1998.0998
- Storbeck, J., & Clore, G.L. (2007). On the interdependence of cognition and emotion. *Cognition and Emotion*, 21, 1212-1237. doi:10.1080/02699930701438020
- Vallée-Tourangeau, F. (2013). Interactivity, efficiency, and individual differences in mental arithmetic. *Experimental Psychology*, 60, 302-311. doi: 10.1027/1618-3169/a000200
- Vallée-Tourangeau, F., Sirota, M., & Villejoubert, G. (2013). Reducing the impact of math anxiety on mental arithmetic: The importance of distributed cognition. In M. Kauff, M.Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the*

Thirty-Fifth Annual Conference of the Cognitive Science Society (pp. 3615-3620). Austin, TX: Cognitive Science Society.

Vallée-Tourangeau, F., Sirota, M., & Vallée-Tourangeau, G. (2016). Interactivity mitigates the impact of working memory depletion on mental arithmetic performance. *Cognitive Research: Principles and Implications*, 1, 26. doi: 10.1186/s41235-016-0027-2