



Mental Arithmetic and Interactivity: the Effect of Manipulating External Number Representations on Older Children’s Mental Arithmetic Success

Wendy Ross¹ · Frédéric Vallée-Tourangeau¹ · Jo Van Herwegen²

Received: 2 October 2018 / Accepted: 3 April 2019 / Published online: 20 June 2019

© The Author(s) 2019

Abstract

Manipulative artefacts are considered useful scaffolds of arithmetic during early years education, but their use is considered less important as children get older. Yet adult arithmetic performance often recruits artefacts to improve accuracy and efficiency, and so the same benefits should accrue to children beyond early years. We propose that interacting with manipulatives supports working memory by extending the mental workspace. This is a hitherto poorly explored aspect of the pedagogical benefits of manipulatives. Forty-three children aged between 7 and 9 years old were invited to assess whether interacting with artefacts supported arithmetic when under a working memory load. Interactivity and load were manipulated in a 2×2 repeated measures design. Children calculated the total of seven digits either with hands down—low interactivity—or moving numbered tokens—high interactivity. Additionally, in half the conditions, the children performed a second task designed to tax working memory. The children were also assessed across a battery of measures examining individual differences. As expected, performance was considerably worse under load. More sums were answered correctly in high-interactivity conditions regardless of working memory load. In the high-interactivity condition with a load, there were significant correlations between performance and both numeracy and fluid intelligence suggesting that even for more numerate children manipulatives confer benefits. The expression of arithmetic skills depends on the nature of the external resources offered by the environment. This dynamic agent–environment transaction highlights the pedagogical importance of external resources when assessing children’s arithmetic abilities.

Keywords Expertise · Interactivity · Manipulatives · Mental arithmetic · Working memory

✉ Wendy Ross
w.ross@kingston.ac.uk

Jo Van Herwegen
f.vallee-tourangeau@kingston.ac.uk

Extended author information available on the last page of the article

Alongside literacy, mathematical fluency is a cornerstone of most education systems. Both formal mathematical qualifications and numeracy levels positively correlate with earnings and employment status (Lane & Conlon, 2016). Furthermore, as life is becoming more reliant on advances in the science, engineering and technology sectors, mathematical fluency across the population becomes increasingly important (Office for Standards in Education, 2012). The ability to quickly carry out calculations is an important part of this mathematical fluency, and in education, this fluency is tested by mental arithmetic tests that form part of national assessments. This assessment strategy is predicated on the theoretical position that calculations are generated by internal numerical representations and exclude resources external to the participants. In practice, however, complex mental arithmetic may be less mental based than classical understandings might suppose. Rather, external constraints and resources augment and influence internal cognitive abilities. A full understanding of the most appropriate way to support children's calculations requires us to examine how they are scaffolded by external resources.

Interactivity. Gestures and interaction with external resources pose a formal challenge to the cognitive sciences and these features of cognitive behaviour are typically excluded in models of cognitive processing (Núñez, Edwards, & Matos, 1999). Yet, it is becoming increasingly clear that ignoring these might lead to an unsatisfactory explanation of what drives and explains the so-called mental arithmetic as it unfolds in the classroom. We argue from the perspective of distributed cognition that thinking has to be cast as an emergent product of a cognitive ecosystem (Hutchins, 2010) rather than something that happens in isolation. From an education perspective, knowledge and learning is often dynamically socially constructed and also offloaded onto environmental artefacts such as diagrams or concrete blocks. Yet, the traditional focus on intelligence casts intelligence as something a child possesses rather than a situated activity (Pea, 1993).

Current research in decision making and problem solving in adults has focused on how problem solvers recruit external resources to aid their thinking and extend their mental workspace. The recruitment of external resources to augment thinking comes under the umbrella term of 'interactivity', a theoretical position that suggests that a thinker's performance is a product of a system configured in terms of both external and internal resources. Research in this area examines how thinking is scaffolded through interacting with an external representation of the problem, be it numbered or lettered tiles (Maglio, Matlock, Raphaely, Chernicky, & Kirsh, 1999) or physical models of problems to be solved (Vallée-Tourangeau, Sirota & Vallée-Tourangeau 2016). The research contrasts how thinking proceeds in task environments that afford different levels of interactivity. Low-interactivity task environments isolate mental resources: hands are kept still, and the environment remains static and unmovable, and thus, thinking can only proceed in the head as it were. In conditions of high interactivity, the thinker is able to recruit different resources to augment his or her thinking: hands are free to move and the environment offers opportunities to interact with a reification of internal concepts. Thinking in these conditions involves physical transformations of the external resources that promote the enactment of new ideas and skills.

Vallée-Tourangeau and Vallée-Tourangeau, (2017) have previously argued that the classic information processing model should be modified to include these

external resources. In this model (SysTM), external resources are not something to be ignored or tacked onto the end of cognitive architecture but are instead key parts in a cognitive act. Predicated on a transactional logic (to adapt Malafouris, 2018), this model rejects linear processing and rather includes processing loops which emphasise a reciprocal, non-linear nature. The inductive loop supports action that does not emanate from a mental plan but one that is guided by the environment. Crucially for a consideration of the role of external resources in arithmetic, the model predicts that as working memory is overloaded, thinkers will change pathways and recruit additional resources.

High interactivity has a consistent beneficial effect on arithmetic performance even in the presence of a drain on working memory resources. Participants in high-interactivity conditions get more calculations correct and have a lower absolute calculation error, whether adults (Cary & Carlson, 1999; Guthrie & Vallée-Tourangeau, 2018) or children (Allen & Vallée-Tourangeau, 2016). Indeed, interactivity appears to be most effective as an aid to arithmetic when there is a working memory load, whether that load is caused by overloading a particular aspect of working memory artificially (through articulatory suppression; Vallée-Tourangeau, Sirota & Vallée-Tourangeau, 2016) or as a function of the length of the sum (e.g. adding 11 single-digit numbers or 17 single-digit numbers; the latter creates a heavier burden on working memory, Guthrie & Vallée-Tourangeau, 2018).

It is commonly accepted that mental arithmetic relies on both working and long-term memory (De Stefano & Lefevre, 2004). Interim sums are held in memory and subvocalized to keep a running count, place markers must also be rehearsed (to identify which numbers have been included in the sum, and which have not), and executive function skills deploy attentional resources and retrieve long-term memory knowledge of simple sums (Vallée-Tourangeau, 2013). Recent meta-analyses have focused on different areas of the relationship between working memory and mathematics and have demonstrated that working memory correlates with mathematics performance across all age ranges (Raghubar, Barnes, & Hecht, 2010) and mathematical areas (Friso-van Den Bos, Van der Ven, Kroesbergen, & Van Luit, 2013). In children under 10, there appears to be more reliance on counting and effortful processing than in adults (Thevenot, Barrouillet, Castel, & Uittenhove, 2016). A counting procedure such as commonly seen in children of this age is more reliant on working memory and short-term storage because the total is constantly updated. The child must keep track of both the addends and the number counting (Witt, 2010). Therefore, the role of working memory in addition of children under 10 is likely to be greater and thus, high-interactivity conditions that allow the child to recruit external resources should be particularly beneficial.

When children's arithmetic ability is studied in a cognitive scientist's lab, it is often done in a controlled manner to eliminate extraneous variables. However, such variables may be integral to an understanding of the phenomenon under consideration. Ecological plausibility is especially important in considering the application of cognitive models in education settings. For example, it is becoming increasingly clear that the use of fingers during supposedly wholly mental arithmetic aids addition in children (Dupont-Boime & Thevenot, 2018; Moeller, Martignon, Wessolowski, Engel, & Nuerk, 2011). Furthermore, the use of external resources appears to alter the way that children approach a problem and encourage the use of conceptually more developed strategies (Manches & O'Malley, 2016).

Manipulatives in Education

Research on the use of manipulatives from developmental and educational psychology has led to mixed conclusions. The most recent meta-analysis found a small to medium effect size of the use of manipulatives in education (Carboneau, Marley, & Selig, 2013) moderated by age, perceptual richness of the manipulatives, and instructional guidance. The main theoretical reasons for the success of manipulatives as outlined in the meta-analysis were (a) supporting the development of abstract reasoning, (b) stimulating real-world knowledge, (c) encoding via motoric channels, and (d) allowing the learners to explore mathematical concepts for themselves. However, in this review, there was no consideration of the pivotal role of working memory in arithmetic and the scaffolding effect of manipulatives.

This lack of focus on the potential benefits manipulatives offer to working memory characterises the otherwise rich literature on manipulatives in mathematical education. There has been very little focus on the support that external artefacts offer to children's working memory by extending the mental workspace beyond the thinker. As outlined above, interactivity supports working memory in adults and Allen and Vallée-Tourangeau (2016) showed that being able to manipulate tokens with number representations mitigated the deleterious effects of maths anxiety in 10- and 11-year olds, suggesting that interactivity in the form of manipulatives can have beneficial effects on children's arithmetic and support working memory. This theoretical aspect of the benefits of manipulatives is worth examining to clarify some of the debate in this area. Manipulatives can take various forms. They are often used to represent relationships, such as Numicon, but they can also make abstract numbers concrete.

The state-funded English education system is a phase-based education system. Primary schools are divided into three main phases: Early Years (aged 4–5), Key Stage 1 (aged 5–7) and Key Stage 2 (7–11). Mathematics education moves from a concrete play based direction in the Early Years curriculum to be increasingly abstract in the final years of primary education. In Year 2 (aged 6–7), the National Curriculum teachers' guidance explicitly encourages the use of 'materials and a range of representations' (Department for Education, 2013, p. 11). One year later in year 3 (aged 7–8), however, there is an abrupt change and the emphasis is on mental computation, especially in addition and subtraction (see Department for Education, 2013, p. 18). The current study assesses the performance of children who are learning in this later phase (ages 7–9) and should not have regular access to manipulatives in their everyday school experience according to the National Curriculum. For those schools that do use manipulatives, it is more likely to be used with lower attainers and in an haphazard way. Focus interviews (Griffiths, Back, & Gifford, 2017) state that teachers' use of manipulatives was 'almost accidental' (p. 5). It further revealed that they were mainly used with younger children. In addition, 92 % of teachers reported using manipulatives for 'lower attainers' in the 9–11 age range, whereas only 45% used them with the 'higher attainers'. Given that the use of external resources has been shown to support arithmetic even in adults, the assumption that manipulatives are only useful to younger children and lower attainers should be examined.

The Current Experiment

The experiment reported here explored the arithmetic performance of 7- to 9-year-old children when tasked with completing sums adding seven single digits. Children performed these sums in contexts that varied in the degree of interactivity, that is, whether they were able to interact with and change the external representation of the sum. In line with adult studies on interactivity, in the low-interactivity condition, children had a static visual representation of the sum created from fixed tokens. They could not use their hands to complement their thinking and could not change the presentation or use finger-counting strategies. By contrast, in the high-interactivity condition, the children were allowed to move the same tokens at will. The low-interactivity condition assessed performance when reliant on purely mental strategies, whereas the high-interactivity condition assessed performance when children could recruit features from their environment. In the low-interactivity condition, the artefacts are iconic representations of the numbers and cannot be moved, whereas they become dynamic manipulatives in the high-interactivity condition affording a range of actions that can dynamically reconfigure the physical presentation of the problem, guiding attention and cueing further action.

To determine how working memory resources are implicated across the different levels of interactivity, the sums were completed in the presence and absence of a secondary task, namely reciting the alphabet. This task was expected to monopolise part of children's working memory, namely the phonological loop, and reduce their ability to rehearse interim sum totals. We predicted that loading the phonological loop during a dual task should result in substantial decrements in performance in both the high- and low-interactivity conditions. However, a higher degree of interactivity should enhance performance, even in the presence of a secondary task. Children were also profiled along a number of dimensions, including IQ, numeracy and maths anxiety. These measures informed exploratory analyses to determine the degree to which performance was moderated by these measures in different interactivity contexts.

Method

Participants

Parental permission was received from 48 pupils from a mainstream primary school in Surrey, England, to take part in the experiment. One child was excluded because this child had additional special educational needs, one child became distressed (before the study started), and one child refused to take part. The experimental data for a further two children were disregarded because the testing conditions were too noisy for them to fully concentrate on the task at hand. The final sample, therefore, consisted of 43 participants (23 females) aged between 93 and 117 months ($M = 102.7$, $SD = 6.97$).

The repeated measures design employed here is similar to the design employed in previous research with adult and children participants, that is, that participants took part in multiple conditions, and so the experimental power could be reliably calculated from previous research. The effect size of interactivity on arithmetic performance in similar experiments ranged from 0.138 (Vallée-Tourangeau, 2013, $N = 42$, adult participants) to

0.379 (Allen & Vallée-Tourangeau, 2016 $N = 59$, year 5 and 6 pupils). In an experiment that looked at articulatory suppression on arithmetic performance specifically, the effect size of interactivity was 0.21 (Vallée-Tourangeau, Sirota & Vallée-Tourangeau, 2016, $N = 52$, adult participants). On the basis of this previous work, we expected to observe an effect size of interactivity between 0.2 and 0.25 on proportion correct performance. On the basis of G*Power assuming $\alpha = 0.05$, $1 - \beta = 0.80$, a sample size of 34 would have sufficient power to detect an effect size of 0.25 and a sample size of 52 to detect one of 0.2.

Procedure

A 2 (interactivity—low, high) \times 2 (phonological load—without, with) repeated measures design was employed. In each condition, participants calculated five sums, each of which involved adding seven single digits, with totals varying from 23 to 44 (e.g. $2 + 5 + 7 + 5 + 4 + 1 + 3$). Each condition had the same five sums with the same digits in a different, random order. This allowed us to control for the difficulty of the sum. For all the conditions, numbered pebbles (approximately 4 cm in diameter) were used to present the sums; each sum was presented as a random cloud arrangement (see Fig. 1).

Participants completed these sums twice within each level of interactivity: once with a load on the phonological loop created by the continuous repetition of the alphabet and once without. In the high-interactivity condition, participants were invited to move the tokens as they wished and to use their hands freely. In the low-interactivity condition, the tokens remained fixed and participants were not allowed to use their hands to aid addition. This way we hoped to isolate the performance when reliant on purely internal, mental processes and contrast it with performance when participants were allowed to recruit resources external to the thinker. This dual-task procedure has been used in previous studies looking at the role of working memory in arithmetic in both adults (Lemaire, Abdi, & Fayol, 1996) and children (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).

Prior to the start of the experiment, the children were invited to play with the tokens as well as practise reciting the alphabet. This practice continued until the children were comfortable with the experimental procedure. The experiment was administered in a quiet room in their school during a school day. To minimise fatigue, the experiment was conducted over two sessions with a break of a fortnight between sessions. The order of the four experimental conditions was randomised for each participant with the following constraint: each testing session included one of the two high-interactivity



Fig. 1 The numbered tokens and their configuration for one of the sums employed in this experiment

conditions. The order of the sums was also counter balanced in each condition within each session.

Participants were instructed to calculate the sum and announce the answer. The tokens were initially hidden behind a screen and overall latencies in seconds were calculated from when this was removed. In load conditions, participants started the alphabet before the screen was removed. The sessions were also voice recorded enabling subsequent analysis of the reciting rate of letters. The children were not given feedback on the outcome of the addition; rather, they were praised regardless of their performance. If a participant appeared to be struggling, the experimenter confirmed with the child that he or she was happy to continue. If a child did not want to continue with a sum, the child was praised for the effort but the performance was recorded as a 'fail'. Allowing the child to give up in this way minimised potential distress in the conditions with a load.

Measures of Individual Differences

In addition to completing the four sets of sums, the participants were profiled in terms of their overall fluid intelligence using Raven's coloured progressive matrices (RCPM; Raven, Court, & Raven, 1990). Their numeracy was assessed in two ways: with the Numerical Operations sub-test from the Wechsler's Individual Achievement Test (UK 2nd Edition; WIAT-II, Weschler, 2005) and by answering as many single digit additions as possible in 45 seconds in a test created specifically for this study (this latter test was employed in previous research with adult participants, correlating strongly with arithmetic performance and working memory capacity, Guthrie & Vallée-Tourangeau 2018). They also completed a forward digit span task and a reverse digit span from the British Ability Scales-II in which participants had to repeat a string of single digits back to the researcher, either in the exact same order (forward digit span) or in reversed order (reverse digit span) (Elliott, Smith, & McCulloch, 1996). Finally, the children's level of mathematics anxiety was assessed with the Maths Anxiety Scale for Young Children Revision (MASYC-R; Ganley & McGraw, 2016) with slight adjustments made for British English. These measures were administered by the researcher in separate sessions to the experiment with the exception of the MASYC-R which was administered in class by the classroom teacher.

Results

Analysis of the audio recording revealed that the children maintained the load well across both conditions with a mean rate of 1.40 ($SD = 0.75$) letters per second in the low-interactivity condition and 1.42 ($SD = 0.61$) letters per second in the high-interactivity condition; this difference was not significant, $t(36) = 0.270$, $p = .788$.

Before reporting the calculation performance and latency data, we report the rates of failed attempts for each of the five sums in each experimental condition. A failed calculation is different to one where the participant gave the wrong answer; a calculation was classified as failed when a child did not calculate the sum and did not give an answer. The inclusion of the option to give up and not return was important to ensure that the children were not put under undue pressure. Unsurprisingly, participants failed

to complete the sums more often in the two working memory load conditions when they concurrently repeated the alphabet (see Table 1).

With a concurrent load, mean failure was lower in the high-interactivity condition ($M=0.40$, $SD=0.89$) than in the low-interactivity condition ($M=0.86$, $SD=1.36$). In the absence of load, mean failure was at floor (one failure across all participants) in the high-interactivity condition ($M=0.02$, $SD=0.15$), and marginally higher in the low-interactivity condition ($M=0.12$, $SD=0.32$). A 2×2 repeated measures analysis of variance (ANOVA) revealed a significant main effect of interactivity, $F(1, 42)=8.45$, $p=.006$, $\eta^2=.167$ and a significant main effect of the concurrent task, $F(1, 42)=14.96$, $p<.001$, $\eta^2=.263$. The interaction was also significant, $F(1, 42)=4.30$, $p=.04$, $\eta^2=.093$. Post hoc tests revealed that interactivity made a significant difference in both load conditions: the number of failed attempts was always lower with high interactivity, in the absence of load, $t(42)=2.08$, $p=.044$, $d=0.32$, and in the presence of load, $t(42)=2.58$, $p=.013$, $d=0.39$. The interaction is best explained in terms of the difference in the number of fails in the absence and in the presence of load in the low ($M=0.74$, $SD=1.31$) and in the high ($M=0.37$, $SD=0.87$); these mean differences were significant, $t(42)=2.08$, $p=.044$, $d=0.32$.

The mean latencies in seconds per sum of all four conditions are reported in the middle of Table 1. While participants were substantially slowed down in the dual-task conditions, latencies were not influenced by the level of interactivity. In a 2×2 repeated measures ANOVA, the main effect of load was significant, $F(1, 39)=19.40$, $p<.001$, $\eta^2=0.332$, but the main effect of interactivity, $F<1$, and the interaction, $F<1$, were not. The mean percentage of correct answers for all four conditions is reported in the bottom portion of Table 1. Performance was always better in the high-interactivity conditions, but was substantially affected by the concurrent task. Thus, in the absence of a concurrent load, mean percent correct was higher in the high ($M=72.0$, $SD=25.9$) than in the low ($M=62.3$, $SD=29.0$) interactivity condition. In addition, performance was generally lower with the concurrent load, but participants performed better in the high ($M=43.7$, $SD=32.1$) than in the low ($M=32.7$, $SD=32.1$) interactivity condition. A 2×2 repeated measures ANOVA confirmed these impressions: the main effect of load was significant, $F(1, 40)=48.9$, $p<.001$, $\eta^2=.550$, as was the main effect of

Table 1 Mean number of failures, latency to solution (in seconds), and percent correct solutions (with standard deviations) in the low- and high-interactivity conditions when participants did not engage in a secondary task (no load) or when they concurrently recited the alphabet (load)

No load			Load		
Interactivity	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>
Low	0.12	0.32	Failures	0.86	1.36
High	0.02	0.15		0.40	0.89
Low	21.5	9.1	Latency (s)	31.3	23.5
High	21.5	8.2		33.6	17.9
Low	62.3	29.0	% Correct	32.7	32.1
High	72.0	25.9		43.7	32.1

interactivity, $F(1, 40) = 10.10$, $p = .003$, $\eta^2 = .202$, but the interaction was not significant, $F < 1$.

Individual Differences and Predictors of Performance

The mean scores in each of the individual difference measures are reported in Table 2, alongside the correlations between these scores and the percent correct performance in each of the four conditions. There were significant correlations between children's performance on the numeracy test and their WIAT scores, $r(39) = 0.628$, $p < .001$, as well as their working memory as assessed by the reverse digit span, $r(39) = .387$, $p = .010$. There were significant correlations between performance in the high-interactivity condition while under a load with numeracy, $r(41) = 0.451$, $p = .002$, RCPM, $r(41) = .304$, $p = .047$, WIAT, $r(41) = .312$, $p = .041$, and reverse digit span scores, $r(41) = .362$, $p = .017$. The correlation between working memory capacity as assessed with the reverse digit span and performance in the four experimental conditions is further illustrated in Fig. 2.

Discussion

The current study explored the impact of interactivity both with and without a load on working memory capacity during mental arithmetic in primary school children. In line with previous studies (Lemaire et al., 1996), the results indicated that children rely on working memory when engaging in mental arithmetic. They took longer and made more errors when their working memory resources were burdened with a secondary task. Similar to performance in both child (Allen & Vallée-Tourangeau, 2016) and adult studies (Guthrie & Vallée-Tourangeau 2018; Vallée-Tourangeau, Sirota, Vallée-Tourangeau, 2016), participants were also significantly more accurate when allowed to move their hands and tokens, either with or without a concurrent load. The effect size of interactivity on performance was commensurate with what was observed in similar experiments exploring the impact of interactivity on mental arithmetic performance (in fact, replicating closely the effect size reported in an experiment on articulatory suppression with adult participants, Vallée-Tourangeau et al. 2016). Moreover, children were able to complete more sums in the interactive conditions (fewer failures); indeed, in the high-interactivity condition without a load, only one sum was not attempted. It is important to bear in mind that each child worked with the same five sums across all four conditions, so the significant differences in performance cannot be attributed to differences in the inherent difficulty of the sum or individual capabilities but rather reflect differences in the experimental conditions. In other words, given the same calculation but a varied environment, children's performance was significantly different. This supports arguments that intelligence should be considered as distributed (Pea, 1993) and that external artefacts constitute an important part of the whole cognitive ecosystem.

Significant correlations between performance and numeracy, standard mathematical abilities, IQ and working memory capacity were only observed in the high-interactivity condition with a concurrent load. These results indicate that those with better internal resources benefitted most from a high degree of interactivity. Current educational perspectives view manipulatives as scaffolding the learning of those who are less able,

Table 2 Descriptive statistics for the measures of individual differences along with their correlations with the percentage of correct sums in all four conditions

Variables	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10
1. MAS	38.4	9.0	–									
2. Numeracy	22.6	7.9	.296	–								
3. RCPM	28.4	3.7	.188	.120	–							
4. WIAT	20.5	3.8	.208	.628**	.305	–						
5. F-DS	23.9	3.3	.186	–.005	.155	.181	–					
6. R-DS	12.9	3.6	–.012	.387*	.181	.181	.255	–				
7. Low/No L	62.3	29.0	.162	.182	.207	.207	.130	.169	–			
8. Low/Load	32.7	32.1	.209	.304	.214	.246	.106	.305	.456**	–		
9. High/No L	72.1	25.9	–.009	.250	.109	.144	.151	.305*	.507**	.348*	–	
10. High/Load	43.7	32.1	.052	.451**	.304*	.312*	.043	.362*	.361**	.611**	.472**	–

MAS maths anxiety score, RCPM Raven's coloured progressive matrices, WIAT/WIAT II, F-DS forward digit span, R-DS reverse digit span, Low low-interactivity condition, High high-interactivity condition, No L no load on the phonological loop, Load load on the phonological loop

p* < .05; *p* < .01

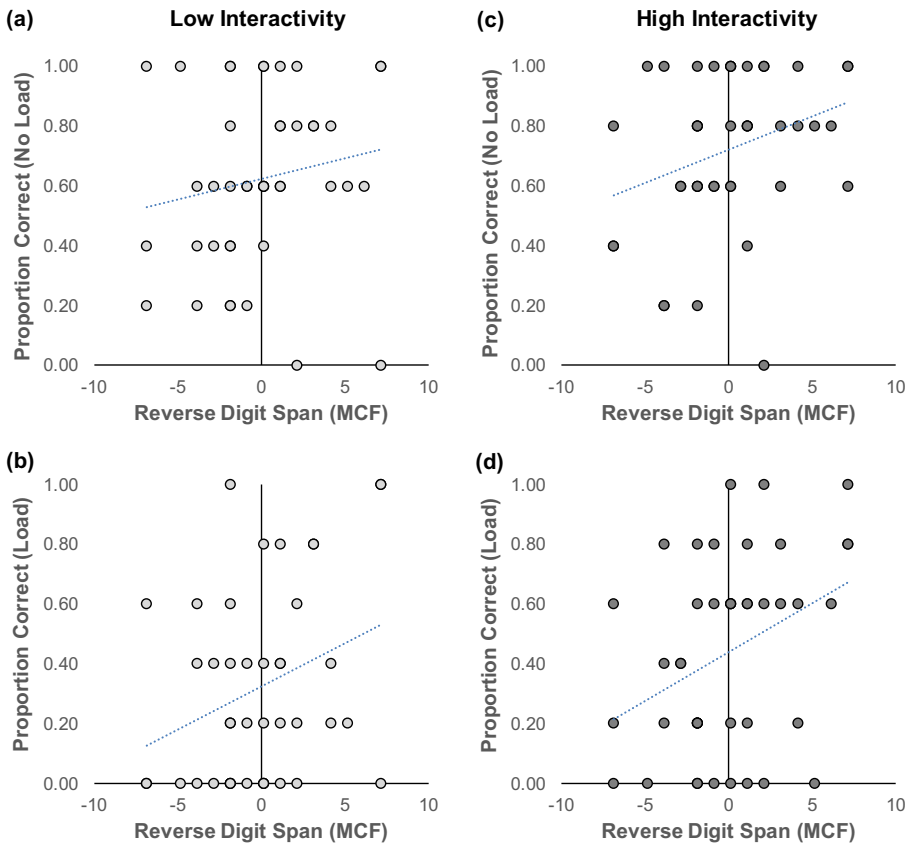


Fig. 2 Proportion correct answers as a function of reverse digit span score (mean centred form) in the low-interactivity condition with no load (a) and with load (b) and in the high-interactivity condition with no load (c) and with load (d)

rather than being integral to everyday thinking for all children (Griffiths et al., 2017). Yet, the results of this experiment support the notion that external resources should be considered across all abilities.

Working Memory

Working memory is a limited capacity system. Anything which reduces the need for working memory will necessarily mean cognitive acts become more efficient. Reifying internal representations—in this instance with numbered artefacts—affords manipulation without drawing on internal working memory resources. However, if the only benefit of interactivity were extending the mental workspace, then benefits should accrue most to those with limited initial resources. That is not what the current results show. While the support interactivity offers to working memory is an important part of its benefits, it also boosted performance in conditions without a concurrent load and aided those who had high working memory resources. In light of these findings, we suggest that an explanation solely based on working memory offloading, as initially hypothesised, would be unsatisfactory. Rather, performance on this task is best

described as an emergent property of the cognitive ecosystem (Hutchins, 2010) that is configured through the dynamic coupling of an agent's internal resources with physical resources from his or her environment. The expression of arithmetic skills can be better understood from a systemic perspective that takes into account the internal abilities of the reasoner and the action affordances of the environment within which the reasoner tackles a task. From this perspective, skills are enacted, and it becomes important to better understand how different types of physical environment, including the classroom, can facilitate the expression of these skills.

Maintaining the low-interactivity condition may itself be a working memory load as it required participants to inhibit the tendency to move their hands. However, the effects of the load were broadly similar across high- and low-interactivity conditions, which suggest that low interactivity did not create an additional working memory load. This aligns with suggestions from Goldin-Meadow et al. (2001) and Vallee-Tourangeau et al. (2016) that an environment low in interactivity is not itself a working memory load.

Pedagogical Implications

There is plenty of evidence from educational psychology and teacher education (Griffiths et al., 2017) that children benefit from external representations of numbers in numerical processing. Yet, the role of artefacts and the external environment does not feature prominently in models of number processing in mainstream cognitive science. In educational psychology, mathematical manipulatives are primarily seen as providing concrete representations of symbols, allowing children to visualise quantities and relationships and form internal representations of numbers (Carbonneau et al., 2013; Griffiths et al., 2017).

While this use of manipulatives is worth consideration, it may be ignoring some of the additional benefits of manipulative artefacts. If manipulatives are only seen as grounding conceptual knowledge, their relationship with working memory may be disregarded. It is unlikely that the benefits reported in this study came from linking the concrete to the symbolic because of the nature of the manipulatives employed. The tokens relied upon pre-existing symbolic knowledge. Rearranging them would not in itself yield an answer to the problem and nor would it facilitate counting procedures (unlike simple blocks where each block represents a number) or visuospatial processes (unlike manipulatives such as Cuisenaire rods). In other words, it is unlikely that success in the high-interactivity conditions was driven by any of the commonly proposed benefits of manipulatives (helping to develop an abstract code, activating real-life representations, invoking motoric learning or encouraging learner exploration [Carbonneau et al., 2013]). Rather, their most likely benefit came from enabling the externalisation of already formed internal representations and extending the mental workspace.

Furthermore, children with the highest internal cognitive resources exploited the external resources most effectively. These findings would be unlikely if the only benefit of tokens were to scaffold new knowledge because, if this were the case, the benefits would be more likely to accrue to the least able children, rather than those already secure in their knowledge. Evidence from other studies suggests that, by increasing working memory capacity through cognitive offloading, children were free to make decisions that maximised their underlying conceptual knowledge (Cary & Carlson, 1999; Manches & O'Malley, 2016), potentially explaining why those who had the most robust mathematical skills demonstrated the greatest benefit. Equally, recently

published findings relating to finger counting suggest that those high in working memory resources are more likely to fully exploit the benefits of using their fingers (Dupont-Boime & Thevenot, 2018). It seems likely that the higher ability children were able to use the artefacts to enact conceptually more sophisticated strategies. However, this inference cannot be explored fully with the data collected. Further research is required to establish the relationship between strategy use and ability. The initial findings here that there was an advantage to using artefacts for higher ability children counters the way educational manipulatives are used in schools (Griffiths et al., 2017). A more detailed exploration of the relationship between individual differences and manipulative use will enable both better design and better pedagogical applications.

Moreover, the fact that the children were much less likely to give up on the task in the conditions with interactivity has further important pedagogical implications. Guthrie & Vallée-Tourangeau's (2015) study of adult arithmetic linked interactivity with positive affect and flow. They found that when participants rated their attitude towards the tasks, those tasks with interactivity elicited more favourable ratings. In addition, the problem-solving efficiency increased as levels of interactivity increased. The low rate of failure in the conditions with interactivity found in the current study indicates that the children were more comfortable in those conditions. This may explain the teachers' view that manipulatives are merely fun (Griffiths et al., 2017). Yet, the current study demonstrates a benefit beyond merely positive affect.

Future Directions

The current experiment contrasted a highly interactive environment, reflective of real-world thinking, against one which was tightly controlled and, beyond anchoring the numbers, relied solely on the child's internal resources and did not offer opportunities to manipulate reified representations. In the low-interactivity condition, the tokens functioned merely as iconic representations, once they could be moved and played with they became useful artefacts (Carbonneau et al., 2013). However, it is unclear from the current results which aspect of the external environment had the most influence on the improved performance. Indeed, it seems likely that different aspects encouraged and supported different strategies. Informal observations suggested that some children grouped the tokens into congenial totals (such as 7 and 3), indicating the use of a strategy predominantly based on the retrieval from long term store, while others used a combination of place marking and finger counting indicating a counting strategy. However, the data gathered here do not let themselves to a systematic classification of strategies. Further studies should disentangle what aspects of manipulating the external representations yield the greatest benefits.

Mathematical knowledge requires both procedural fluency and conceptual knowledge (Rittle-Johnson, Schneider, & Star, 2015). Procedural fluency is often tied to specific problem types and refers to the steps needed to solve a problem, whereas conceptual knowledge is broader and refers to the knowledge of general principles underlying mathematical competency (Rittle-Johnson, 2017). There is some evidence that the use of external resources allows people to draw on their conceptual knowledge, thus strengthening their overall mathematical fluency. For example, when adults were provided with pen and paper, to calculate wages

based on an hourly rate added to different commission rates, participants answered more questions correctly and used strategies which corresponded to the conceptual nature of the task compared to the performance of those who performed the task without a memory aid (Cary & Carlson, 1999). A comparison of the use of physical blocks with paper-based representation also showed a significant difference in strategies used when the children were invited to solve the problems with concrete blocks (Manches & O'Malley, 2016). This suggests that performance on these tasks are best described as an emergent property of the cognitive ecosystem (Hutchins, 2010) that is configured through the dynamic coupling of an agent's internal resources with physical resources from his or her environment. It may be that, as with calculator use in the classroom, the reason that the children with the highest cognitive resources were most able to exploit them was because their conceptual understanding was higher (Ruthven & Chaplin, 1997). By increasing working memory capacity through interacting with tokens, the children were free to make decisions that maximised conceptual knowledge rather than conserving working memory (Cary & Carlson, 1999). More fluent and creative strategies could be employed which may lead to greater conceptual knowledge.

Conclusion

The difference in performance across the conditions underscores the importance of both external resources and working memory in mental calculation and models of numerical cognition. Considering concrete manipulatives as scaffolds to learning, rather than a dynamic part of everyday problem solving, undersells their usefulness to an education environment. A more accurate description is that there is a dynamic system configured through an agent's action, reflecting the internal resources of an agent, and the external resources offered by the physical environment. We argue that it is more productive to profile the cognitive resources and abilities of the system rather than characterising an agent's internal resources and the physical resources offered by the environment independently.

Clearly, much remains to be explored, especially in terms of the strategies that might be enacted in a high-interactivity environment. A high degree of interactivity offers a dynamic landscape of action affordance and possibilities. As a result, different strategies might be enacted. In turn, these strategies may be contingent on the specific token configuration—both as it is presented initially, but perhaps more important as it is shaped dynamically throughout the problem. Hence, we might observe a participant employ different strategies as a function of how he or she interacted with the tokens in a given problem. Certainly, as our informal qualitative observations suggest, the same environment encourages different strategies from different problem solvers, and elicits different types of mathematical knowledge. This will add more clarity to the inconclusive evidence around the use of manipulatives in education. Future study in this area should improve the granularity of the analysis of an agent's actions, and the cognitive system that is configured as a result. The challenge is that the granularity of such analysis must be fine enough to isolate important variables in behaviour yet broad enough to draw general conclusions.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Allen, M. & Vallée-Tourangeau, F. (2016). Interactivity defuses the impact of mathematics anxiety in primary school children. *International Journal of Science and Mathematics Education*, *14*, 1553–1566
- Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, *105*, 380–400.
- Cary, M., & Carlson, R. A. (1999). External support and the development of problem-solving routines. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1053–1070.
- De Stefano, D., & Lefevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, *16*, 353–386.
- Department for Education and Skills. (2013). *Mathematics programmes of study: Key stages 1 and 2. National curriculum in England*. London: Her Majesty's Stationary Office.
- Dupont-Boime, J., & Thevenot, C. (2018). High working memory capacity favours the use of finger counting in six-year-old children. *Journal of Cognitive Psychology*, *30*, 35–42.
- Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales second edition (BAS II). Administration and scoring manual*. London, England: Nelson.
- Friso-Van Den Bos, I., Van Der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44.
- Ganley, C. M., & McGraw, A. L. (2016). The development and validation of a revised version of the Math Anxiety Scale for young children. *Frontiers in Psychology*, *7*, 1181–1188.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*, 516–522.
- Griffiths, R., Back, J., & Gifford, S. (2017). *Using manipulatives in the foundations of arithmetic: Main report*. Leicester: England: University of Leicester Retrieved from [http://www.nuffieldfoundation.org/sites/default/files/files/Nuffield%20Main%20Report%20Mar%202017web\(1\).pdf](http://www.nuffieldfoundation.org/sites/default/files/files/Nuffield%20Main%20Report%20Mar%202017web(1).pdf).
- Guthrie, L. G., & Vallée-Tourangeau, F. (2018). Numbers in action: individual differences and interactivity in mental arithmetic. *Cognitive processing*, *19*, 317–326.
- Guthrie, L. & 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.
- Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science*, *2*, 705–715.
- Lane, M., & Conlon, G. (2016). The impact of literacy, numeracy and computer skills on earnings and employment outcomes. (OECD Education Working Papers No. 129). Paris, France: OECD Publishing.
- Lemaire, P., Abdi, H., & Fayol, M. (1996). The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology*, *8*, 73–104.
- Maglio, P. P., Matlock, T., Raphaely, D., Chemicky, B., & Kirsh, D. (1999). Interactive skill in Scrabble. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society* (pp. 326–330). Mahwah, NJ: Lawrence Erlbaum.
- Malafouris, L. (2018). Mind and material engagement. *Phenomenology and the Cognitive Sciences (online first)*, *18*(1), 1–17.
- Manches, A., & O'Malley, C. (2016). The effects of physical manipulatives on children's numerical strategies. *Cognition and Instruction*, *34*, 27–50.
- Moeller, K., Martignon, L., Wessolowski, S., Engel, J., & Nuerk, H. C. (2011). Effects of finger counting on numerical development: The opposing views of neurocognition and mathematics education. *Frontiers in Psychology*, *2*, 328.
- Núñez, R. D., Edwards, L., & Matos, J. (1999). Embodied cognition as a grounding for situatedness and context in mathematical education. *Educational Studies in Mathematics*, *39*, 45–65.
- Office for Standards in Education. (2012). *Mathematics: Made to measure*. London, England: Her Majesty's Stationary Office.

- Pea, R. (1993). Practices of distributed intelligence and designs for learning. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations*. Cambridge, England: Cambridge University Press.
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110–122.
- Raven, J. C., Court, J. H., & Raven, J. C. (1990). *Manual for Raven's progressive matrices and vocabulary scales – section 2: Coloured progressive matrices*. Oxford, England: Oxford Psychologists Press.
- Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives, 11*, 184–190.
- Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review, 27*, 587–597.
- Ruthven, K., & Chaplin, D. (1997). The calculator as a cognitive tool: Upper-primary pupils tackling a realistic number problem. *International Journal of Computers for Mathematical Learning, 2*(2), 93–124.
- Thevenot, C., Barrouillet, C., Castel, C., & Uittenhove, K. (2016). Ten-year-old children's strategies in mental addition: A counting model account. *Cognition, 146*, 48–57.
- Vallée-Tourangeau, F. (2013). Interactivity, efficiency and individual differences in mental arithmetic. *Experimental Psychology, 60*, 302–311.
- 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.
- Vallée-Tourangeau, G., & Vallée-Tourangeau, F. (2017). Cognition beyond the classical information processing model: Cognitive interactivity and the systemic thinking model (SysTM). In S.J. Cowley & F. Vallée-Tourangeau, F. (Eds) *Cognition Beyond the Brain* (pp. 133–154). London: Springer
- Wechsler, D. (2005). *Wechsler individual achievement test – second UK edition (WIAT-II UK)*. London, England: Pearson Assessment.
- Witt, M. (2010). Cognition in children's mathematical processing: Bringing psychology to the classroom. *Electronic Journal of Research in Educational Psychology, 8*, 945–970.

Affiliations

Wendy Ross¹ · Frédéric Vallée-Tourangeau¹ · Jo Van Herwegen²

¹ Department of Psychology, Kingston University, Kingston upon Thames, Surrey KT1 2EE, UK

² UCL, Institute of Education, 25 Woburn Square, London WC1H 0AA, UK