UNIVERSITY OF LEEDS

This is a repository copy of *Cognitive Offloading: Structuring the Environment to Improve Children's Working Memory Task Performance*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/146506/

Version: Accepted Version

Article:

Berry, EDJ, Allen, RJ orcid.org/0000-0002-1887-3016, Mon-Williams, M orcid.org/0000-0001-7595-8545 et al. (1 more author) (2019) Cognitive Offloading: Structuring the Environment to Improve Children's Working Memory Task Performance. Cognitive Science, 43 (8). e12770. ISSN 0364-0213

https://doi.org/10.1111/cogs.12770

© 2019 Cognitive Science Society, Inc. This is the peer reviewed version of the following article: Berry, E. D., Allen, R. J., Mon-Williams, M. and Waterman, A. H. (2019), Cognitive Offloading: Structuring the Environment to Improve Children's Working Memory Task Performance. Cogn Sci, 43: e12770., which has been published in final form at https://doi.org/10.1111/cogs.12770. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Cognitive offloading: structuring the environment to improve children's working memory task performance

Ed D. J. Berry, Richard J. Allen, Mark Mon-Williams, Amanda H. Waterman*

School of Psychology, University of Leeds, LS2 9JT, UK

*corresponding author

Email: a.h.waterman@leeds.ac.uk

Tel: +44 (113) 3435731

Acknowledgements

This work was completed as part of EJDB's PhD, jointly funded by the University of Leeds and the Bradford Institute of Health Research. Two of the authors of this paper (AHW and MMW) were supported by the NIHR CLAHRC Yorkshire and Humber. www.clahrc-yh.nihr.ac.uk. The views and opinions expressed are those of the authors, and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

Keywords: cognitive offloading; working memory; developmental; metacognition

Abstract

Research has shown that adults can engage in cognitive offloading, whereby internal processes are offloaded onto the environment to help task performance. Here we investigate an application of this approach with children, in particular children with poor working memory. Participants were required to remember and recall sequences of colors by placing colored blocks in the correct serial order. In one condition the blocks were arranged to facilitate cognitive offloading (i.e., grouped by color), whereas in the other condition they were arranged randomly. Across two experiments (total N = 166) the ordered condition improved task performance for children with low working memory ability. In addition, participants in Experiment 2 rated the difficulty of the two arrangements, and performed a further condition in which they were given an opportunity to freely arrange the blocks before completing the task. Despite performing better in the ordered condition, children with low working memory ability did not rate the ordered arrangement as easier, nor did they chose an ordered arrangement when given the opportunity to do so. This research shows that cognitive offloading can also be a useful process in populations other than typical adults, and the implications of this work for supporting children with poor working memory are discussed.

1. Introduction

Making use of our environment to support cognition is ubiquitous, from writing notes (Clark, 2008; Intons-Peterson & Fournier, 1986) to arranging a kitchen (Kirsh, 1995). The importance of spatial organization to performance has been demonstrated experimentally in adults with computerized tasks (Hess, Detweiler, & Ellis, 1999), pen and paper tasks (Zhu & Risko, 2016), and virtual reality tasks (Ragan, Bowman, & Huber, 2012). The organization of material is not only relevant to performance, but also affects metacognitive judgements of one's knowledge, such that individuals who organize their environment (e.g. office space) are more confident in their knowledge (Hamilton, McIntyre, & Hertel, 2016; Hertel, 1988).

Here we consider the role of task-relevant groupings of objects in the environment in the context of children completing a working memory task. Task-relevant arrangements of items in the environment can be seen as an 'intelligent use of space' (Kirsh, 1995), where space can be used to simplify choice, performance, or internal computation to assist with performing a given task. For example, Beach (1993) notes that bartenders make use of the fact that different glasses are used for different drinks. By arranging the glasses in the relevant spatial sequence as the drinks order is given, they off-load some of the memory requirements onto the environment. Each glass only affords a small number of possible drinks, and therefore acts as a cue to recall. Zhu & Risko (2016) also looked at how adults arranged items that were necessary to complete a particular task, within their immediate environment. They found that when the task required more effort, participants arranged the items in order to maximize performance. Both these examples show the way in which people use spatial configurations in their environment to support on-going performance (see also, Scribner, 1986).

The role of arranging objects in the environment has also been characterized as cognitive offloading (Dunn & Risko, 2015; Gilbert, 2015; Risko & Gilbert, 2016). A (sub)task is offloaded onto the environment when the environment is organized to substitute a functional role previously carried out internally. Clark (2008) gives the example of the dynamic flow of information between internal and external representations via a notepad when working on some problem. This offloading then frees-up cognitive resources to allow for more complex and sophisticated forms of processing or behavior.

The ability to reduce the cognitive demands of a task is important within the context of working memory. Working memory is a limited capacity system that allows us to store and process information over short periods of time (seconds) (Baddeley, 2012; Cowan, 2005). The fact that working memory is a limited system makes the concept of offloading particularly relevant within this paradigm. Transferring some of the processing requirements of a task onto the environment reduces the demands placed on working memory, releasing capacity to help with successful task completion (Risko & Gilbert, 2016).

Within the working memory literature, the storage of information is sometimes characterized as short-term memory, with the label 'working memory' applying to conditions also involving executive control. However, several models include both aspects within the broader conceptualization of working memory (Allen & Waterman, 2015; Baddeley, 2012; Cowan, 2005; Logie, 2011; Waterman, Atkinson, Aslam, Holmes, Jaroslawska, & Allen, 2017), with storage (both verbal and visuospatial) labelled as 'simple working memory' and storage plus executive control as 'complex working memory'. Both simple and complex working memory are key to the successful completion of a large range of tasks (Cowan, 1999; Gathercole, Durling, Evans, Jeffcock, & Stone, 2008; Waterman et al, 2017), although will make different contributions depending on the specific task elements under consideration. For example, the retention of orally presented information primarily engages simple verbal working memory, whereas having to remember and manipulate information (for example, adding two digits in your head) would increase the demands on complex working memory resources. Thus, when investigating how offloading may affect the performance on a working memory task, it is important to consider its relationship with the different components of working memory.

Working memory is an important predictor of key developmental outcomes, is fundamental to learning, and predicts academic attainment (Alloway & Alloway, 2010; Alloway, Gathercole, Kirkwood, & Elliott, 2009; Cragg & Gilmore, 2014; Gathercole, Pickering, Knight, & Stegmann, 2004; Monette, Bigras, & Guay, 2011). In order to complete classroom learning activities, children must encode, retain, and manipulate information (Baddeley, 2007; Gathercole & Baddeley, 1993; Gathercole et al, 2008; Gathercole, Woolgar, Kievit, Astle, Manly, & Holmes, 2016). Therefore, children with poor working memory perform less well on measures of numeracy and literacy (Alloway et al., 2009; Holmes, Hilton, Place, Alloway, Elliott, & Gathercole, 2014; Gathercole et al. 2016), and this link continues into secondary education (Gathercole & Alloway, 2008).

Given the pervasive difficulties experienced by children with poor working memory abilities, researchers have investigated ways in which to support such children. Several researchers have looked at the possibility of improving working memory directly through training (e.g. Holmes, Gathercole, & Dunning, 2009; St. Clair-Thompson, Stevens, Hunt, & Bolder, 2010). However, attempts to train working memory have hitherto been unsuccessful in producing robust far transfer to outcomes such as mathematics or IQ (Melby-Lervåg & Hulme, 2013; Sala & Gobet, 2017). An alternative to training is to provide support through task adaptation, or appropriate management of the environment (Elliott, Gathercole, Alloway, Holmes, & Kirkwood, 2010; St Clair-Thompson et al., 2010). Given the limited capacity of working memory, can we adapt the environment to enable cognitive offloading of aspects of the task, which in turn will support performance?

The current research therefore explores whether children's performance on a working memory task can be supported by making use of the environment. Specifically, we examined whether task-relevant groupings of objects in the environment can help support the ability to remember sequences of instructions. Manipulating the task-relevance of the materials serves as a type of cognitive offloading, whereby processes linked to successful task completion can be partially offloaded onto the environment, thereby reducing the cognitive load on a limited capacity system. Given the reduced capacity limits of children with poor working memory, we might expect to see increased benefits of offloading in such populations.

To the best of our knowledge, no studies to date have investigated the effect of cognitive offloading in children. Given that skills related to offloading, such as goal setting, information processing, and use of control strategies, develop over the early and middle-school years (Anderson, 2012; Chevalier, 2015; Davidson, Amso, Anderson, & Diamond, 2006), it is important to investigate offloading in child populations, rather than making assumptions based on the adult literature. Indeed, in their review paper, Risko and Gilbert (2016) state that understanding how cognitive offloading behavior emerges over childhood is one of the key, as yet unanswered, research questions in this field.

Understanding how cognitive offloading might operate in childhood is important given the substantial improvements in working memory that have been observed up to adulthood (e.g. Gathercole et al., 2004). Studies investigating how children's performance on comparable working memory tasks might vary with other forms of experimental manipulation have typically focused on children in the age range 7- to 11-years old (Atkinson, Waterman, & Allen, in press; Berry et al, 2018; Jaroslawska, Gathercole, Allen, & Holmes, 2016; Shimi, Noble, Astle, & Scerif, 2014; Waterman et al, 2017), as working memory undergoes substantial improvement across this age range. The current study therefore focused initially on 9- to 10-year olds. This ensured consistency with previous literature on working memory, but also maximized the chances that participants were able to engage appropriately with the task by selecting children towards the upper end of the age range. Based on children's performance in the first experiment, Experiment 2 extended the age range to include 8-year-olds.

2. Experiment 1

Participants were asked to complete a working memory task which involved encoding short verbal sequences and recalling them by acting on physical objects. These objects were either arranged randomly or were grouped by color. This latter ordered arrangement offloads aspects of the task (e.g., visual search) onto the environment. We therefore expected that participants would perform better when the blocks were grouped by color. We also took additional working memory measures to explore the influence of individual differences in working memory on performance on the primary task. If grouping objects in the environment allows participants to offload some processing demands, we might expect participants with low working memory particularly to benefit.

Given the primary task involves both storage and processing elements, measures of both simple working memory (storage) and complex working memory (processing) were taken. Further, in line with models suggesting modality-specific storage (Baddeley, 2012; Logie, 2011), one of the simple working memory tasks measured verbal storage, and one measured visuospatial storage. Indeed, the primary task requires retaining verbal information (the colors are presented orally) as well as being visuospatial in nature (selecting the correct block from a row of blocks). Therefore, using tasks that differentiate between different aspects of working memory enables investigation of the pattern of results across these different sub-components. Given this is the first study to examine offloading in a working memory task with children, we did not make any a priori predictions about potential differences in the relationships between specific working memory tasks and performance on the primary task.

2.1 Method

2.1.1. Participants

88 participants took part in Experiment 1. Four were excluded due to having special educational needs leaving 84 children (Mean age = 10.34, SD = 0.28, Range = 9.89 – 10.85). Participants were recruited from a predominately low SES Pakistani British area of Bradford, UK. Ethical approval was obtained from the School of Psychology Ethics Committee, University of Leeds, UK.

2.1.2. Materials

Primary task. For the primary task participants were presented with 12 colored blocks measuring 3 x 3 x 2cm. The blocks were placed in a line with a gap of 3cm between each block. The experimenter read out a sequence of colors at a rate of approximately one per second in a neutral voice. At the end of the sequence the participant was required to pick up a single colored block for each color in the sequence and place it in a response area in front of them. This was self-paced. At the end of a trial the blocks were returned to their original location. Four colors were used (red, blue, yellow, and green). A span procedure was used starting at sequences of length three and increasing to a maximum of eight. Colors could be repeated in a sequence. An item was scored as correct only if it was recalled in the correct location. Participants completed three trials at each span length, with span being increased by one if the participant got two or more trials correct. This procedure was continued until the participant got fewer than two trials correct at a given length. A participant’s span score was then the longest length at which they got two or more trials correct. An additional 0.33 was added on to the score if they got one trial correct at the next sequence length. For example, if a participant got 2/3 trials at span five and 1/3 trials at span six, then they would have a span score of 5.33. Participants completed two conditions for this task. In one condition the blocks were not grouped by color but were displayed in a pseudorandom arrangement (so that adjacent colors always differed, but there was no discernible or consistent order), while in the second condition the blocks were grouped according to their color. The order of the two arrangement conditions was counterbalanced. A schematic of the task is depicted in Fig. 1

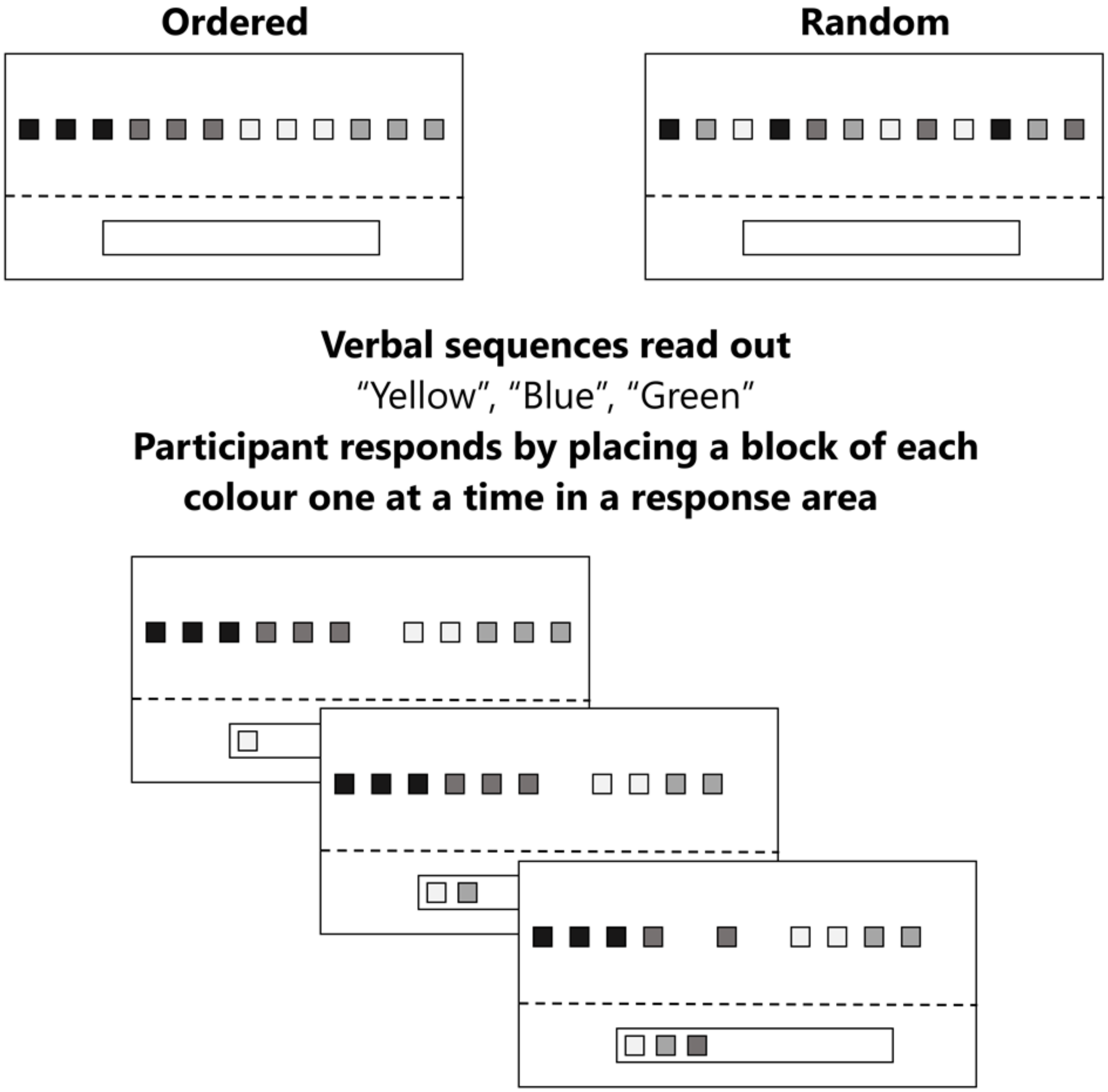

Figure 1. *A schematic of the procedure for the primary task. Shades of grey represent different colors (not to scale)*

Working Memory Measures

The following tasks have been used extensively in the literature as measures of simple and complex working memory (e.g. Berry et al., 2018; Gathercole, Pickering, Ambridge, & Wearing, 2004; Waterman et al., 2017)

Forward digit recall (FDR). This task measures simple working memory in the verbal domain. The experimenter read out digits at a rate of one per second and the participant was required to repeat them back in the same order. The same span procedure as the primary task was used with sequence lengths starting at three and increasing to a maximum of eight. Performance was scored in the same way as the primary task.

Backward digit recall (BDR). This task measures complex working memory. It was identical to forward digit recall except that participants were required to repeat the digits in backwards order. The task began with sequences of length two increasing to a maximum of six.

Corsi blocks. This task measures simple working memory in the visuospatial domain. A board with nine blocks (arranged in a pseudorandom configuration, consistent across participants) was placed in front of the participant. The experimenter touched a sequence of blocks at a rate of approximately one per second. The participant was then required to touch the blocks in the same order. This task used the same span procedure as previous tasks beginning at sequences of three up to a maximum of eight.

2.2. Results

2.2.1. Primary task

The data were analyzed using a Bayesian linear mixed-effects model in R (R Core Team, 2016) using rstanarm (Stan Development Team, 2016)[1]. The Supplementary Materials provide more information on the model. Descriptive statistics for the different variables are provided in Table 1

---

[1] The data and analysis scripts for both experiments are available at https://osf.io/j5qt8/?view_only=9153bc42208143ae81d7058786a8654f.

Table 1. *Descriptive statistics for the different variables for Experiment 1.*

Variable	Mean (SD)	Range
Backward digit recall (BDR)	3.36 (0.85)	1.33 – 6
Forward digit recall (FDR)	5.19 (0.88)	3.33 – 8
Corsi	4.55 (0.89)	3 – 6.33
Ordered condition	4.85 (0.54)	4 – 6.33
Random condition	4.66 (0.77)	3.33 – 6.33
Ordered advantage	0.19 (0.58)	-1.33 – 1

The model included condition (ordered and random), FDR, BDR, and Corsi as predictors as well as interactions for each of FDR, BDR, and Corsi with condition. Random intercepts for subjects were also included in the model. Posterior odds were used to compare the different effects in the model. The posterior odds describe the ratio between posterior samples where an estimate is positive versus negative. For example, posterior odds of 50:1 in favor of a positive coefficient would indicate that, for the 20,000 samples taken from the posterior, 50 estimates were positive for every 1 that was negative for the effect of interest. The interpretation of posterior odds is different from more traditional frequentist tests as there is no agreed cut-off whereby odds above or below a certain level are seen as ‘important’ or ‘significant’, although odds ratios in the thousands clearly represent stronger effects than those in the tens (Morey, Morey, van der Reijen, & Holweg, 2013). However, to aid interpretation, equivalent frequentist analyses were run for each relevant section. For the data presented in this paper, results relating to odds ratios of less than 20:1 were consistently non-significant, results relating to odds ratios higher than 30:1 were consistently and clearly significant. For odds ratios between 20:1 and 30:1 there was some support in the frequentist analyses for these results being significant. A table of correlations between the measures is also provided in the Appendix.

For the effect of condition the odds were 1052:1, with performance being superior in the ordered condition. The odds were 20000:1 in favor of a positive coefficient for FDR predicting overall performance, and were 23:1 in favor of a positive coefficient for Corsi predicting overall performance. However, for BDR the odds were only 6:1, and therefore did not support BDR predicting performance. The interaction between FDR and condition was favored 487:1, as was the interaction between Corsi and condition (odds of 35:1). The interaction between BDR and condition was not supported (odds of 1:1). Table A in the Supplementary Materials provides the parameter estimates and diagnostic for this model.

As there was support for FDR and Corsi predicting overall performance, this was evaluated at different values of a combined working memory score. Fig. 2 shows the distribution of posterior estimates from the model where the combined score was set at the sample minimum, mean, or maximum (see Table 1). This allows us to show how the predicted performance in the two conditions interacts with performance on these working memory measures. In other words, Fig. 2 shows the predicted difference in performance between the ordered and random conditions for an individual performing at the sample minimum, mean, and maximum of a combined score of FDR and Corsi.

At the sample mean of FDR and Corsi the posterior estimate for the difference between the ordered (posterior median = 4.85, 95% credible interval = 4.73 – 4.97) and random (posterior median = 4.66, 95% credible interval = 4.54 – 4.78) conditions was 0.19 (95% credible interval = 0.07 – 0.31). This suggests a slight boost in performance in the ordered condition for the average participant. However, as Fig. 2 shows, the estimated boost in performance in the ordered condition is clearly more pronounced for participants with poor working memory scores (i.e., at the minimum).

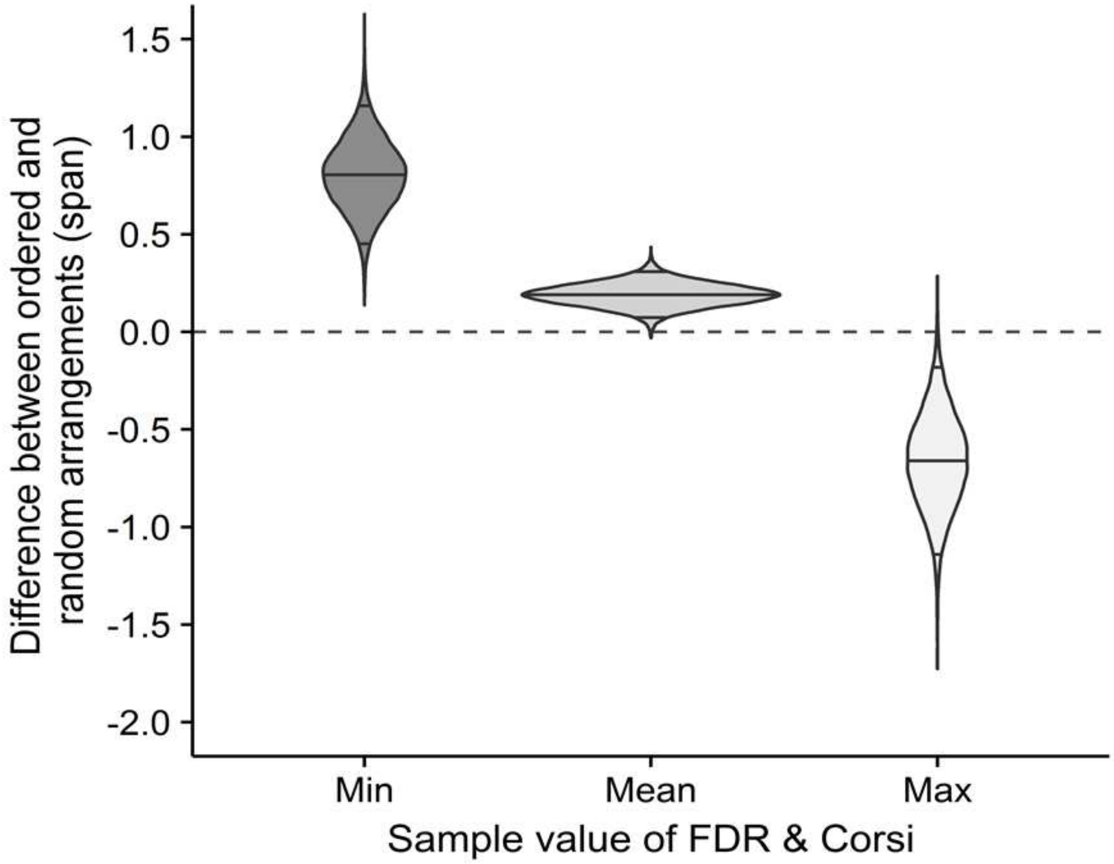

Figure 2. *The distribution of predicted differences between the ordered and random conditions at the sample minimum, mean and maximum for forward digit recall and Corsi. Violin plots show the probability density for different values of a variable such that there is more density over values where the shape is wider. Here the variables are posterior samples from our model. The central line shows the median value, with the outer lines showing the 2.5% and 97.5% quantiles.*

2.2.2. Extreme groups

The interaction between condition and working memory was further explored using an extreme-groups approach. Low (N = 24) and High WM (N = 22) groups were selected by taking participants at or below the 15th percentile (Low WM) and at or above the 85th percentile (High WM) on the forward digit recall task (e.g. Gathercole, et al., 2016). Although there was evidence for an effect of both FDR and Corsi on task performance, the posterior odds ratio for FDR was substantially larger (20000:1) than for Corsi (23:1). Therefore, for the current analysis, only FDR was used to select the high and low WM groups. However, an analysis was run using a composite FDR and Corsi score to create the high and low working memory groups: all results were the same using this alternative metric. As with the primary task analyses a linear mixed-effect model was fitted to these subgroups, with random intercepts for subjects. The model included condition (ordered; random), WM group (high; low), and a condition by WM group interaction. Fig. 3 shows the estimated difference between the ordered and random condition for the two WM groups.

Posterior odds were again used to gauge support for different effects in the model. The effect of condition was favored 1999:1. The effect of WM group was favored by a factor of 20000:1. Finally, the interaction between condition and WM group was favored 6666:1. The posterior odds did not support a difference between conditions in the high WM group (odds of 2:1). In contrast for the low WM group the odds were 20000:1 in favor of better performance in the ordered condition. Table B in the Supplementary Materials provides the parameter estimates and diagnostics for this model.

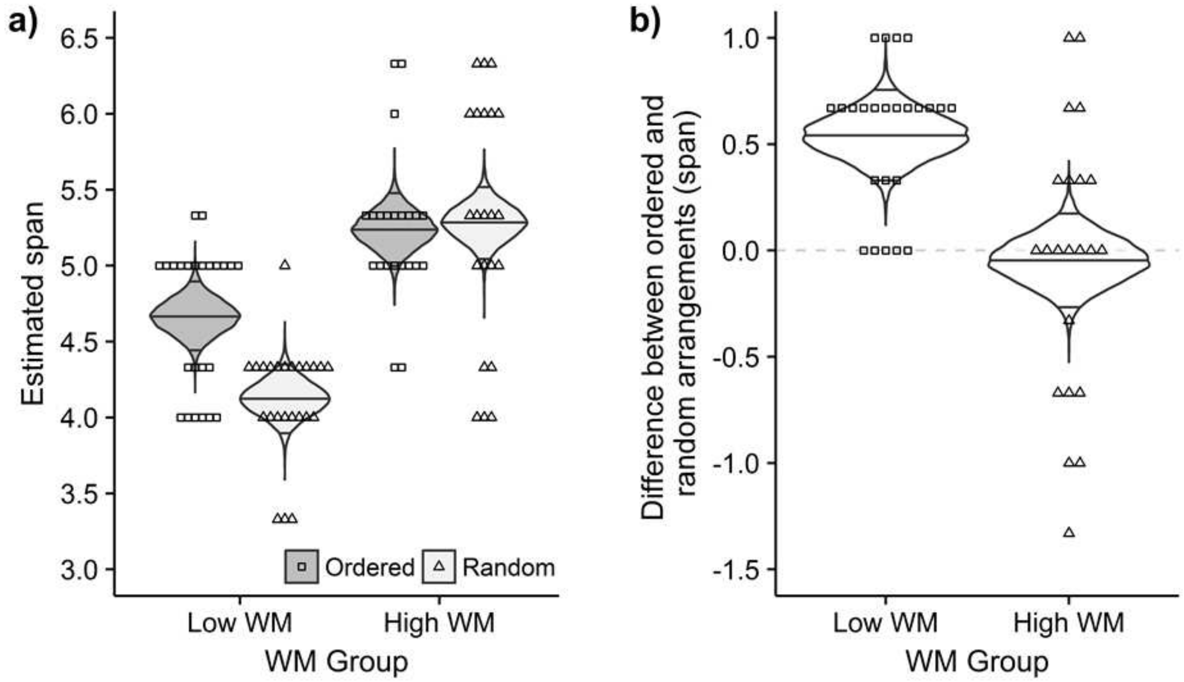

Figure 3. *a) Estimated performance in the ordered and random conditions for the low and high WM groups. (b) The estimated difference between the ordered and random conditions for the low and high WM groups. Unfilled shapes show the raw data. Violin plots show the probability density for different values of a variable such that there is more density over values where the shape is wider. Here the variables are posterior samples from our model.*

2.3. Discussion

The results showed that structuring the environment, such that materials were arranged in a task-relevant manner and thus afforded the opportunity for cognitive offloading, improved performance. This effect was particularly pronounced for participants with lower scores on the working memory task.

There are different possible mechanisms via which cognitive offloading could occur. One possibility is that the task-relevant arrangement reduces the visual search demands, which are lower in the ordered condition, and therefore enable more effective utilization of spatial location as a cue to support task performance. The importance of spatial location for the successful recall of visual and auditory stimuli presented at those locations has been shown even in infants (Benitez & Smith, 2012; Morey, Mareva, Lelonkiewicz, & Chevalier, 2017; Richardson & Kirkham, 2004; Samuelson, Smith, Perry, & Spencer, 2011). This supports the idea that facilitating the ease with which location information can be used to complete a task might lead to improved performance. Alternatively, another task element that may be offloaded onto the environment in the ordered condition is action planning. Indeed, previous work has shown that children with poor working memory skills are less accurate on tasks related to visual search (Holmes et al, 2014), and action planning (St Clair-Thompson, 2011).

When considering the different working memory tasks, the strongest predictor of performance on the primary task was the measure of simple verbal working memory. There was also some support for simple visuospatial working memory predicting performance, although this was less strong. The primary task requires the retention of verbal information; the participant has to store a series of orally presented color words in order to complete the task successfully. The relationship with the simple visuospatial working memory task might also be expected given that the participant has to engage with a physical set of blocks in a particular spatial order. When considering the complex working memory measure, there was no evidence of a relationship between that and the primary task, over and above the effect of simple verbal working memory. This suggests that the storage aspects of working memory were more important to task performance than executive control. Whether or not this pattern was replicated was explored in Experiment 2.

In summary, this experiment showed that task-relevant grouping of objects in the environment improved performance, and was particularly beneficial for children with low working memory ability. One interesting question is whether children are aware of this difference between ordered and random arrangements, and, further, whether children will spontaneously make use of the environment to structure a task to optimize performance. These issues relate to the concept of metacognition; a range of abilities that involve knowledge about cognition, and that can impact on task performance (Kuhn, 2000; Schneider, 2008). Research has shown that metacognition develops over childhood (Flavell, 1979; 1999; Siegler, 1996), with children's metastrategic knowledge (awareness of task difficulty, task performance, and appropriate strategies to employ) improving with age (Kuhn, 2000). In addition, children with poor working memory show poorer metacognition, for example, they are less effective at selecting higher-level strategies, and are less able to transfer taught strategies across domains or tasks (Alloway et al, 2009).

Experiment 2 therefore sought to replicate the findings from Experiment 1, but also to assess children's awareness of the potential benefits of structuring the environment. Research on cognitive offloading in adults has explored this relationship with metacognition (Dunn & Risko, 2015; Risko & Gilbert, 2016). Risko and Gilbert (2016) found that participants' metacognitive beliefs about expected task performance were an important factor in whether or not they engaged in cognitive offloading. Exploring children's perception of task difficulty, and how that relates to spontaneous use of the environment, is an interesting extension of the adult literature. Further, understanding children's awareness of how task structure may impact on performance will be equally important when considering possible practical applications. Finally, one limitation of the current experiment was that span scoring tends to result in a small number of possible scores, thus limiting the ability to discriminate between participants. Therefore, in Experiment 2 we used a method where children are given a fixed number of sequences at lengths three to six (see Waterman et al., 2017), and performance was scored at the item-level giving a greater range of possible values (for a review of scoring methods see Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). In addition, the working memory tasks were also measured using this fixed length scoring, and an extra task was introduced that specifically measures complex visuospatial working memory, to enable a more detailed exploration of the relationship with the primary task and different components of working memory.

3. Experiment 2

Experiment 2 attempted to replicate and expand on Experiment 1. Firstly, we expected to observe both the main effect of condition, and the interaction with working memory. Secondly, given the 9- to 10-year old children were successfully able to complete the task in Experiment 1, we expanded the age range to include 8-year-olds. Finally, we investigated whether children were aware that an ordered arrangement can boost performance, and whether they would spontaneously choose such an arrangement. To address this, participants rated the difficulty of the ordered and random conditions and were asked to judge which condition was easier. In addition, a 'free arrangement' condition was added where participants were asked to arrange the blocks for themselves before completing a set of trials.

If a participant chooses an ordered arrangement in the 'free arrangement' condition, they would need to have recognized that the ordered arrangement made the task easier, and then translate this knowledge in selecting an ordered arrangement (Kuhn, 2000; Schneider, 2008). Recent work using a working memory task has shown that 7- to 12-year-old children are sensitive to the cognitive effort required to complete a task, and will avoid a more difficult task because it requires greater cognitive effort to complete (Chevalier, 2017). Therefore, children in this age range can show awareness of task difficulty (see also Efklides, Kourkoulou, Mitsiou, and Ziliaskopoulou, 2006). As such, we expected that children in the current experiment would recognize that an ordered arrangement facilitates better performance. However, the ability to anticipate task difficulty, and then plan an effective strategy to improve performance (proactive control), is less well developed in younger children, and is likely influenced by concurrent developments in working memory (Chevalier 2015). Therefore, it could be that children may not choose an ordered arrangement in the free arrangement condition, even if they can retrospectively judge the ordered condition as easier. In addition, children with poor working memory may be particularly poor at using effective strategies due to difficulties holding and manipulating possible strategies in mind.

3.1. Method

3.1.1. Participants

100 participants took part in Experiment 2. Fifteen were excluded due to having special educational needs and a further two participants were excluded due to be distracted on the primary task. Finally, one participant was excluded due to missing data on one of the additional measures. This left 82 children (Mean age = 9.62, SD = 0.62, Range = 8.51 – 10.49). Participants were recruited from a predominately low SES Pakistani British area of Bradford, UK. None of these children had taken part in Experiment 1. Ethical approval was obtained from the School of Psychology Ethics Committee, University of Leeds, UK.

3.1.2. Materials

For Experiment 2 all participants completed the same set of trials, rather than using a span procedure as per Experiment 1. For each condition participants completed a block of four trials at sequence lengths three, four, five and six. Performance was then scored as the total number of colors correctly recalled in the correct position. An additional color (orange) was added to reduce the chance of participants getting items correct by guessing. This meant there were 15 objects, three of each color. The order of the ordered and random conditions was counterbalanced. Participants then completed an additional block of trials where they were instructed to arrange the blocks in whatever order they wanted, so long as they remained equally spaced in a row, as in the ordered and random conditions. After all three blocks had been completed participants were asked to rate whether the task was easier when the blocks were ‘mixed up’ or ‘grouped by color’. They then rated the difficulty of each condition on a 7-point Likert scale.

Additional measures

For Experiment 2 computerized versions of the additional measures from Experiment 1 were used as well as an additional complex visuospatial working memory measure. Computerized tasks remove potential differences in task presentation that come from different researchers presenting stimuli. All the tasks were written using PsychoPy 1.83.01 (Peirce, 2007) and presented on a touchscreen tablet (see Berry et al, 2018). Unlike the tasks in Experiment 1, all participants completed a fixed set of trials, with overall proportion correct at the item level used as the outcome measure.

Forward digit recall (simple verbal WM). At the start of a trial, the word 'Listen' was presented on screen for 1000ms followed by a 1000ms blank screen. Sequences of digits were presented over headphones with utterances lasting 350-550ms followed by an ISI of 1000ms. A 1250ms retention interval followed each sequence of digits. They had to be recalled in the same order at a participant's own pace by pressing response boxes on screen. Recall was followed by a 1000ms inter-trial interval. The response boxes were approximately 1.7cm by 1.9cm and were evenly spaced in a line from one to nine. Three practice trials at sequence length three were completed followed by 16 test trials. The test trials were divided into four blocks of four trials. Three numbers were presented for each trial on the first block, increasing by one to six digits per trial on the fourth block. After the last trial of each block a message was presented stating that the span length would increase. The full 16 trials were completed by all participants. A digit was not presented twice within a trial.

Backward digit recall (complex verbal WM). The procedure for this task was identical to the forward digit recall task except that the numbers were recalled in the reverse order to how they were presented. Additionally, the trials began at a sequence length of 2 building up to a length of 5.

Corsi block task (simple visuospatial WM). Nine pseudo-randomly arranged boxes (approximately 2.5cm^2) were displayed on screen. A sequence of boxes lit up, with participants responding by pressing the boxes in the same order that they lit up. Like the forward digit recall task, participants completed 16 trials from sequence lengths three to six, split into four blocks of four trials. Each trial began with a 1000ms fixation cross followed by a 1000ms blank screen. The boxes then lit up in yellow for 500ms with a 1000ms ISI. Self-paced recall was preceded by a 1250ms retention period. A 1000ms inter-trial interval followed recall.

Odd-one-out (complex visuospatial WM). This task was an adaption of the existing odd-one-out task (Alloway, Gathercole, & Pickering, 2006; Russell, Jarrold, & Henry, 1996). Each trial began with a 1000ms fixation cross followed by a 1000ms blank screen. Three shapes were then presented simultaneously, with participants having to press the shape that was different from the other two ('the odd-one-out'), as well as remembering its location. After a series of such processing decisions there was a 1250ms retention interval. Five boxes were then displayed on-screen and the location of the "odd-ones-out" had to be recalled, in order, by pressing the appropriate boxes. A 1000ms inter-trial interval followed recall. Eight different shapes were used, each measuring approximately 2.5cm^2. There were eight possible combinations of the shapes with each shape serving as the odd-one-out and as the distractor once. Three shapes were always presented, with one shape serving as the odd-one-out and one serving as the two 'distractor' shapes. A 1000ms ISI followed participants pressing one of the shapes. The same nine locations as the Corsi task were used for presenting the sets of three shapes. A subset of five locations was used within a trial, with a location being used no more than once for the odd-one-out. Participants completed three practice trials followed by 16 trials in four blocks from span lengths two to five.

3.2. Results

3.2.1. Primary task

For Experiment 2 the outcome measure was accuracy (rather than span, as in Experiment 1). To respect the fact accuracy in not normally distributed the data were analyzed using a binomial logistic regression (see Supplementary Materials). This approach prevents the spurious effects that can sometimes emerge from averaging over accuracy data and using analyses such as ANOVAs (Dixon, 2008). Descriptive statistics are provided in Table 4.

Table 4. *Proportion correct for the variables in Experiment 2.*

Variable	Mean (SD)	Range
Backward digit recall	0.7 (0.18)	0.14 – 1
Forward digit recall	0.78 (0.12)	0.44 – 1
Corsi	0.74 (0.15)	0.36 – 0.97
Odd-one-out	0.59 (0.21)	0.16 – 1
Ordered	0.75 (0.1)	0.53 – 0.94
Random	0.75 (0.11)	0.46 – 0.93
Ordered advantage	0.01 (0.08)	-0.17 – 0.24

As with Experiment 1 the models included random intercepts for subjects. Additionally, the effect of condition was allowed to vary by subject. The model included condition (ordered; random), FDR, BDR, Corsi, and Odd-one-out as predictors[2]. Interactions between each predictor

---

[2] The model was run with age group as a predictor. While there was a small main effect of age group (whereby older children showed a small improvement in overall performance) it did not interact with other variables. The effect of condition, and working memory ability, resulted in the same pattern of results across age groups. Thus, the simplified model is reported to aid comparison with Experiment 1.

and condition were also included. See the Supplementary Materials (Table C) for coefficients and diagnostics. See Appendix for correlations between measures, as per Experiment 1.

The effect of condition was not supported by the model (odds of 4:1 in favor of superior performance in the ordered condition). The odds were 20000:1 in favor of FDR positively predicting overall performance. However, for BDR the odds were only 13:1 in favor, and for Corsi and Odd-one-out the odds were only 2:1 in favor of positive relationships with overall performance. There was some support for the interaction between FDR and condition (odds of 29:1). On the other hand, the interactions between BDR and condition, and Odd-one-out and condition were not supported (both with odds of 2:1); neither was the interaction between Corsi and condition (odds of 4:1). The model was used to make the estimates shown in Fig. 4 by varying the value of FDR to make predictions from the model. Given that the effect of Corsi on the primary task was not supported in this Experiment, only FDR was used to create the plots.

To summarize, the main effect of condition was not replicated, but the interaction between condition and working memory was supported. As with Experiment 1, this was driven by the performance on the FDR task.

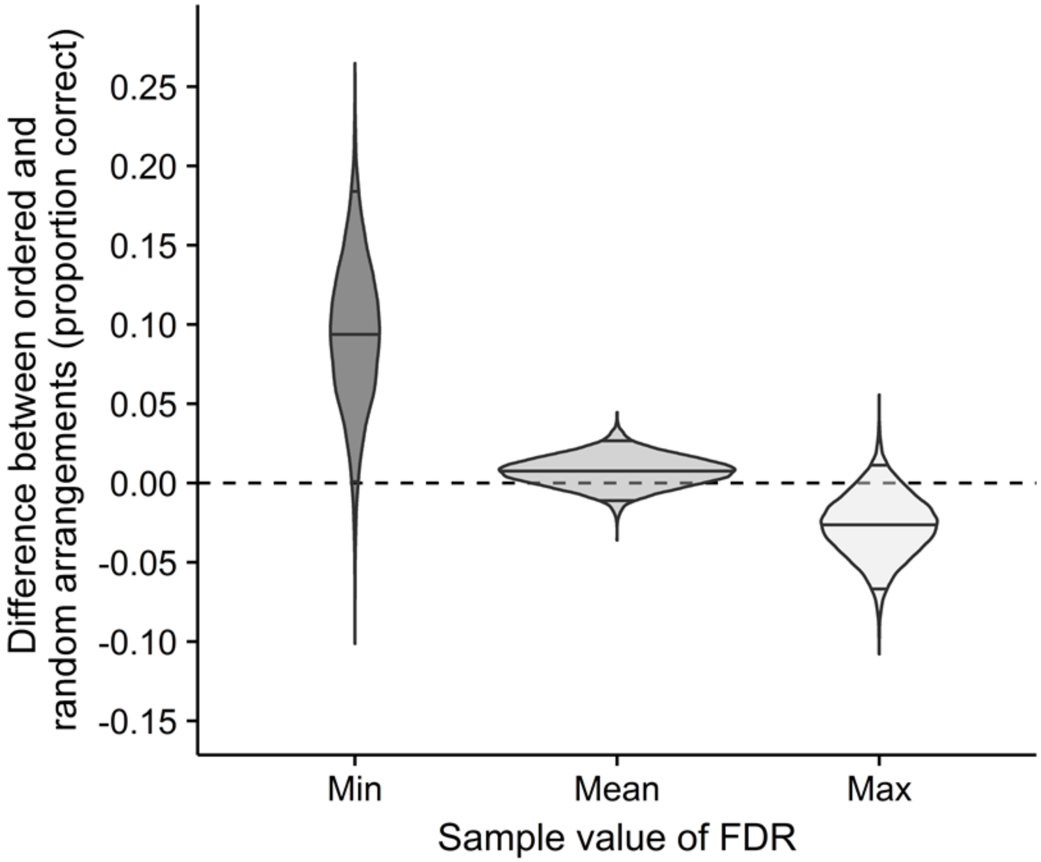

Figure 4. *The distribution of predicted differences between the ordered and random conditions at the sample minimum, mean and maximum for forward digit recall. Violin plots show the probability density for different values of a variable such that there is more density over values where the shape is wider. Here the variables are posterior samples.*

3.2.2. Extreme groups

As with Experiment 1 an extreme group analysis was also carried out, in order to explore further the interaction between condition and working memory. Participants were again selected by being at or above the 85th (high WM; N = 20), or at or below the 15th (low WM; N = 13) percentiles for FDR. The model included condition and working memory group as predictors, as well as their interaction. The effect of condition was not supported (odds of 3:1 in favor of better performance in the ordered condition). The effect of working memory group was strongly favored by a factor of 20000:1. The interaction between condition and working memory group was supported with odds of 99:1. When looking at the working memory groups separately, the effect of condition was not strongly supported for the high WM group (odds of 8:1). In contrast, for the low WM group the odds were 43:1 in favor of superior performance in the ordered condition. The estimates from this model are shown in Fig. 5.

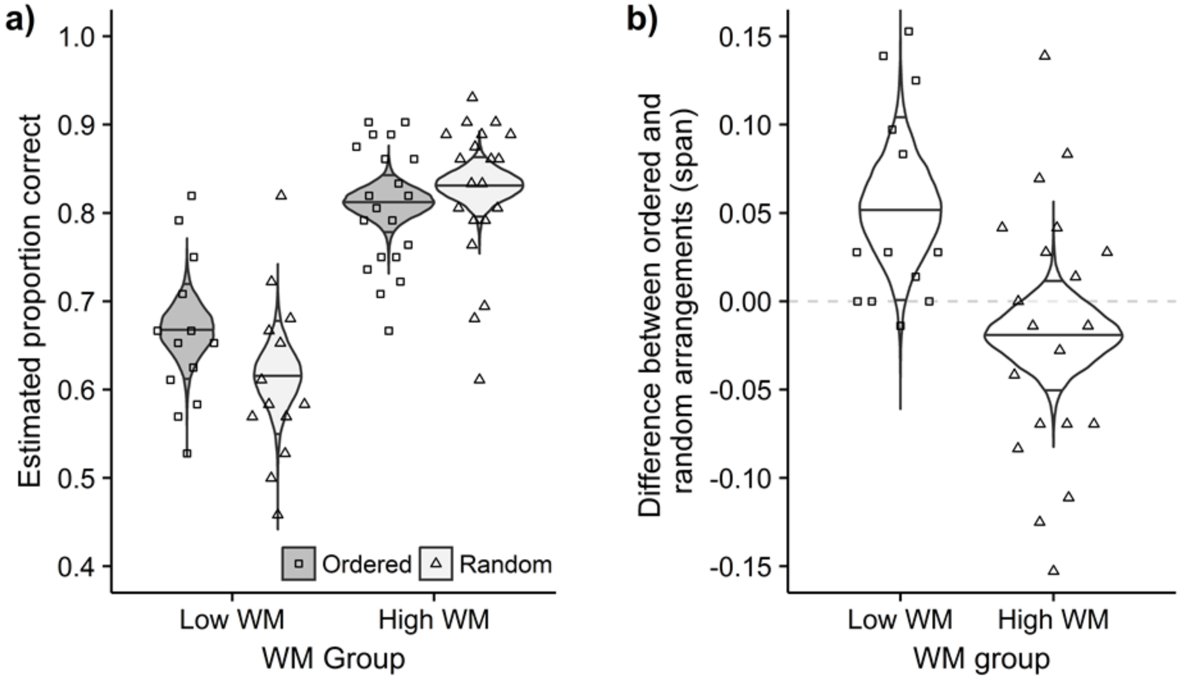

Figure 5. a) Estimated performance in the ordered and random conditions for the low and high WM groups. (b) The estimated difference between the ordered and random conditions for the low and high WM groups. The unfilled shapes show the raw data. Violin plots show the probability density for different values of a variable such that there is more density over values where the shape is wider. Here the variables are posterior samples from our best-fitting model.

3.3.3. Free arrangement condition and metacognitive measures

Overall, the majority (78%) of participants rated the ordered condition as easier than the random condition. The modal response on a 7-point Likert scale for the ordered condition was 7 (i.e. 'very easy'). For the random condition the modal response was 4 (i.e. 'neither easy nor difficult'). Fig. 6 shows the distribution of difficulty ratings for the two conditions. However, when looking at the extreme groups, only 57% of the children in the low WM group rated the ordered condition as easier, compared to 85% for the high WM group, despite performance being better in the ordered condition for the low WM group.

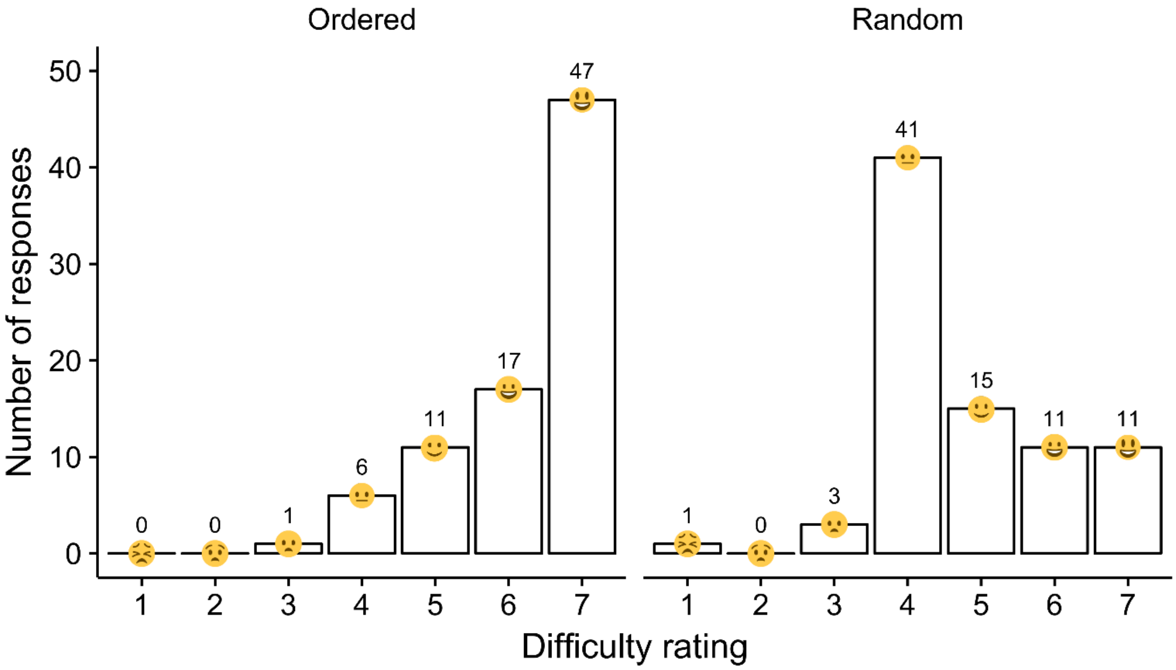

Figure 6. *The distribution of difficulty ratings for the two conditions. 1 = very difficult, 4 = neither difficult nor easier, 7 = very easy. The emoji atop the bars were placed on a scale for participants to make the difficulty ratings. (Emoji, Copyright 2016 Twitter, Inc and other contributors, CC-BY 4.0; graph created with help from emoGG, https://github.com/dill/emoGG)*

To evaluate the randomness of each participant's chosen arrangement a 'randomness metric' was created. The 15 locations in which the blocks could be placed were numbered 1 to

15. The absolute distance between each pair of blocks of the same color was then calculated and squared. The sum of these squared distances was then adopted as a metric of the randomness of the chosen arrangement. The distribution of this randomness metric can be seen in Fig. 7, where lower values indicate a less random arrangement (the ordered arrangement used for the primary task would have a score of 30 on this metric). As Fig. 7 shows, the majority of participants chose a relatively random arrangement. In addition, the mean randomness for the low (Mean = 504.43, SD = 278.58) and high (Mean = 568, SD = 194.78) WM groups was very similar, so this comparison was not analyzed further.

With regards to overall performance in the free condition (*M* = 0.76, SD = 0.11), this was very similar to performance in the ordered (*M* = 0.75, SD = 0.10) and random (*M* = 0.75, SD = 0.11) conditions (this is also the case when looking at the high and low WM groups separately). When considering the impact of participants' chosen arrangement on task performance, the randomness metric did not relate to performance in the free arrangement condition (odds of 10:1). However, there was support for a negative relationship between randomness metric and overall performance in the ordered and random conditions, averaging across condition (odds of 54:1), such that participants who performed better overall in the ordered and random conditions chose a less random arrangement in the free arrangement task.

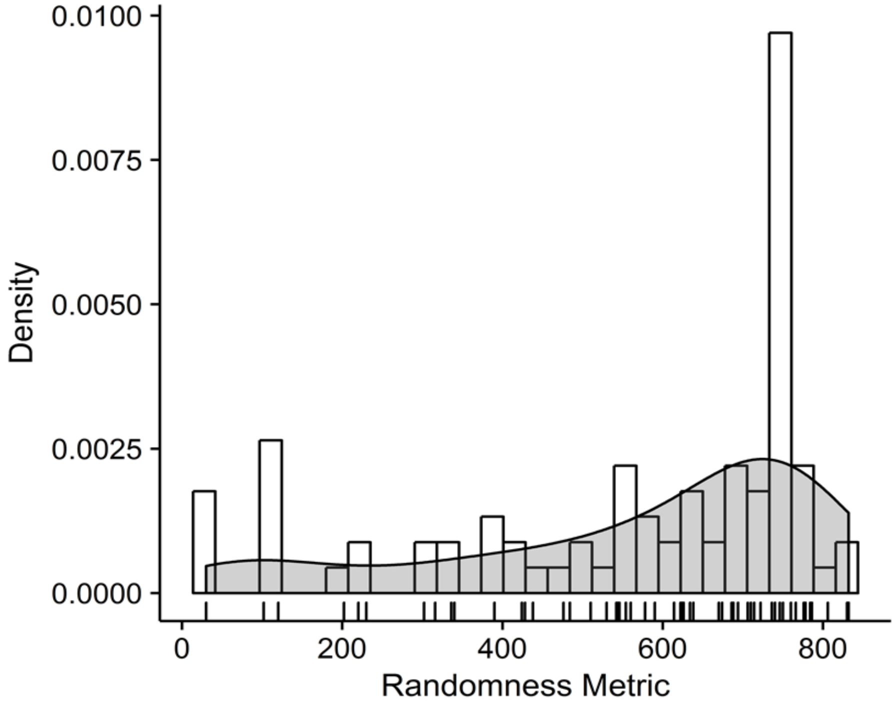

Figure 7. The distribution of scores on the randomness metric with an overlaid density estimate.

3.3. Discussion

We replicated the interaction with working memory observed in Experiment 1 whereby children with low working memory scores performed better in the ordered condition compared to the random condition. This supports the idea that children with poor working memory benefit from the ability to offload aspects of task performance onto the environment. As in Experiment 1, children with high working memory performed as well in the random condition as they did in the ordered condition. One possible reason for this might relate to task difficulty. It is possible that, even in the random condition, the task is relatively easy for children with high working memory. If the task does not overly tax working memory capacity, there would be no benefit from freeing up working memory resources via offloading. Future research could explore this by increasing task difficulty and observing how this changes the performance across conditions for children with high working memory.

With regards to the metacognitive data, there were differences between children in the high and low working memory groups. When asked which condition was easier, children in the low working memory group did not discriminate between the ordered and random conditions. In addition, they did not tend to choose an ordered arrangement in the free condition, despite benefitting from an ordered arrangement. These data suggest that children with poor working memory are less likely to show metacognitive awareness regarding how difficult a task is, and how strategies or offloading might help their task performance (Alloway et al, 2009). In contrast, children in the high working memory group rated the ordered condition as easier than the random condition, although they still did not choose an ordered arrangement in the free condition. Children in this group may therefore be showing an awareness of task difficulty (Efklides et al, 2006), but lack the ability to use appropriate strategies to support performance (Chevalier, 2015, 2017).

However, it is important to note that children in the high working memory group did not actually perform better in the ordered condition. As previously discussed, the task might not be sufficiently difficult for offloading to provide substantial benefit. The fact that these children *perceived* the ordered condition to be easier may be explained by factors relating to speed/accuracy trade-off (Schmidt, 1994; Smits-Englesman, Sugden, & Duysens, 2006). Children in the high working memory group may have been able to complete the task in the ordered condition more quickly, even though this did not affect accuracy. This might lead to a perception of reduced effort that is not captured by an accuracy performance metric. The current experiments used physical blocks, which made the collection of reaction times difficult. Future studies could investigate these possible speed-accuracy trade-offs by creating an equivalent automatized task that would enable collection of response latency as well as response accuracy.

To summarize the key findings from Experiment 2, children with poor working memory benefitted from the opportunity to engage in cognitive offloading but showed less developed metacognition, with a reduced ability to judge task difficulty, identify facilitative strategies, and actively adjust task set-up accordingly.

4. General Discussion

This study investigated whether aspects of a working memory task could be offloaded onto the environment in order to boost children's performance. Across two experiments, the results showed that task-relevant arrangements improved performance for children with low working memory ability. In addition, Experiment 2 found that, despite performing better in the ordered condition, children with poor working memory did not consistently rate this condition as easier, nor did they choose an ordered arrangement when allowed to arrange the blocks themselves before completing the task.

The consistent advantage of an ordered arrangement for children with poor working memory is consistent with embodied accounts of cognition that allow for the offloading of cognitive processes onto the environment. Grouping the blocks in a task-relevant manner may have allowed children to offload processing functions that would otherwise need to be performed internally (Dunn & Risko, 2015; Gilbert, 2015; Risko & Gilbert, 2016). This would then particularly benefit children with reduced availability of cognitive resources, i.e., poor working memory. This cognitive offloading may occur via mechanisms such as reduced visual search time, or through facilitating action planning (Holmes et al, 2014; St Clair Thompson, 2011). Future research could explore the mechanisms via which children offload cognitive processes onto the environment in this type of task. For example, eye-tracking could be used to explore visual search times (e.g. Deubel & Schneider, 1996) and the extent to which children fixate on the appropriate target colors during encoding and maintenance (e.g. Morey, Mareva, Lelonkiewicz, & Chevalier, 2018). Indeed, further research investigating whether children with high versus low working ability engage differently in visual search strategies might also help us to understand why children with better working memory ability did not benefit from the ordered arrangement. This links to research which has found some evidence of a relationship between visual search tasks and working memory in older children (Leonard, Weismer, Miller, Francis, Tomblin, & Kail, 2007) and adults (Poole & Kane, 2009).

Whilst children in the low working memory group showed improved performance in the ordered condition, their performance in that condition still did not reach the levels of the children in the high working memory group. Therefore, structuring the task materials in this way did not completely compensate for the reduced capacity of children with poor working memory. This suggests that the amount of offloading available via task-relevant grouping was not enough to offset the more limited cognitive resources in low working memory group. This makes sense, given that even in the structured condition children were required to temporarily retain an ordered sequence of color names. Indeed, it is possible that while structuring the environment may improve performance in children with poor working memory, it can never entirely compensate for the reduction in capacity. Future research could utilize differing levels of support, and via different modalities (visuospatial vs. verbal), to investigate the extent to which functional performance can be fully ameliorated in the face of working memory deficits.

When considering the individual working memory tests, the measure that consistently predicted task performance was Forward Digit Recall, which taps into simple verbal working memory (sometimes referred to as verbal short-term memory). This reflects the fact that the primary task necessitated the storage of verbal codes (i.e., the color words), as a central component for successful completion. There was some weak support for simple visuospatial working memory being linked to task performance in Experiment 1, but this was not replicated in Experiment 2. Clearly the task involved visuospatial information, so we might have expected a stronger relationship here. However, some working memory models eschew the idea of domain specific storage, instead proposing a single ‘focus of attention’ (e.g. Cowan, 1999, 2005), or at least argue against the need to include a specialized visual storage capacity (Gray et al., 2017;

Morey, 2018; but see Hanley & Young, in press; Baddeley, Hitch, & Allen, in press). Indeed, even models that contain modality-specific sub-components acknowledge that working memory is an inter-linked, holistic construct (Baddeley, 2012; Logie, 2011). Therefore, if children's primary focus was retaining color words in order to complete the task, simple verbal storage would then serve as the most effective measure of the involvement of the working memory system.

The research presented here also explored children's metacognitive judgments. Risko and Gilbert (2016) discuss the importance of metacognition for successful offloading in adults. Appropriate strategy choice involves awareness of the internal resources available, and how external aids can be successfully employed to improve performance. Experiment 2 found that children showed some awareness of task difficulty, but this was related to working memory ability. Children with poor working memory were less able to differentiate between conditions, even though their performance levels were influenced by condition difficulty. Thus, the children who were most likely to benefit from using the environment to structure task performance were the least likely to understand the efficacy of task-relevant arrangements. This links to research demonstrating that children with poor working memory ability are also less likely to select appropriate strategies, and less likely to transfer learnt strategies across to other domains or tasks (Alloway et al, 2009; Swanson, 2016; Swanson, Lussier, & Orosco, 2013).

With regards to more practical applications, structuring the classroom environment may provide additional ways to improve learning opportunities for children with low working memory. However, as the results of our metacognitive questions demonstrated, it would not necessarily be sufficient to ensure that the task environment had been adapted appropriately. Children would also need to be aware of such changes, and of the benefits they might afford (Schneider, 2008; Chevalier, 2015). However, before firm conclusions are drawn relating to the application to classroom contexts, it would be useful to investigate whether this pattern of results replicates in a more naturalistic setting using meaningful objects, and in tasks that are more closely analogous to learning activities typically encountered in the classroom.

To summarize, the results of the current studies extend work on cognitive offloading in adults (Dunn & Risko, 2015; Risko & Gilbert, 2016; Zhu & Risko, 2016) to a novel population, and show that the environment can also be utilized to support children’s performance when cognitive resources are limited.

References

Allen, R. J., & Waterman, A. H. (2015). How does enactment affect the ability to follow instructions in working memory? *Memory & Cognition*, *43*(3), 555–561. http://doi.org/10.3758/s13421-014-0481-3

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. http://doi.org/10.1016/j.jecp.2009.11.003

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The Cognitive and Behavioral Characteristics of Children with Low Working Memory, *Child Development, 80*(2), 606–621. http://doi.org/10.1111/j.1467-8624.2009.01282.x

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development, 77*(6), 1698–1716. http://doi.org/10.1111/j.1467-8624.2006.00968.x

Anderson, P. (2002). Assessment and development of executive function during childhood. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *8*(2), 71–82.

Baddeley, A.D. (2007). *Working Memory, Thought, and Action*. Oxford: OUP.

Baddeley, A.D. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology, 63*, 1-29. http://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A.D., Hitch, G.J., & Allen, R.J. (in press). From short-term store to multicomponent working memory: The role of the modal model. *Memory & Cognition.*

Beach, K. (1993). Becoming a bar-tender: the role of external memory cues in a work-directed educational activity. *Applied Cognitive Psychology, 7*, 191-204.

Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition, 125*(3), 339-352.

Berry, E. D., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). The limits of visual working memory in children: Exploring prioritization and recency effects with sequential presentation. *Developmental Psychology*, *54*(2), 240-253. http://dx.doi.org/10.1037/dev0000427

Chevalier, N. (2015). Executive Function Development: Making Sense of the Environment to Behave Adaptively. *Current Directions in Psychological Science*, *24*(5), 363–368. http://doi.org/10.1177/0963721415593724

Chevalier, N. (2017). Willing to Think Hard? The Subjective Value of Cognitive Effort in Children. *Child Development*, *1*, 1–13. http://doi.org/10.1111/cdev.12805

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, *3*(9), 345–351. http://doi.org/10.1016/S1364-6613(99)01361-3

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. http://doi.org/10.1037/0096-3445.133.2.189

Cowan, N. (1999). An embedded-processes model of working memory. *Models of working memory: Mechanisms of active maintenance and executive control*, *20*, 506.

Cowan, N. (2005). *Working memory capacity.* Hove, East Sussex, UK: Psychology Press.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*, 63–68. http://doi.org/10.1016/j.tine.2013.12.001

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision research*, *36*(12), 1827-1837.

Dixon, P. (2008) Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59(*4), 447-456*.* http://dx.doi.org/10.1016/j.jml.2007.11.004

Dunn, T. L., & Risko, E. F. (2015). Toward a Metacognitive Account of Cognitive Offloading. *Cognitive Science*, *40*. 1080-1127 http://doi.org/10.1111/cogs.12273

Efklides, A., Kourkoulou, A., Mitsiou, F., & Ziliaskopoulou, D. (2006). Metacognitive knowledge of effort, personality factors, and mood state: Their relationships with effort-related metacognitive experiences. *Metacognition and Learning*, *1*(1), 33–49. http://doi.org/10.1007/s11409-006-6581-0

Elliott, J., Gathercole, S., Alloway, T., Holmes, J., & Kirkwood, H. (2010). An evaluation of a classroom-based intervention to help overcome working memory difficulties and improve long-term academic achievement. *Journal of Cognitive Education and Psychology, 9*(3), 227–250.

Flavell, J. H., (1977). *Cognitive Development.* Englewood Cliffs, NJ: Prentice Hall.

Flavell, J. H., (1999). Cognitive development: children's knowledge about the mind. *Annual Review of Psychology, 50,* 21-45.

Gathercole, S. E., & Baddeley, A. D. (1993). Working Memory and Language. Lawrence Erlbaum.

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, *22*(8), 1019–1037. http://doi.org/10.1002/acp.1407

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*(1), 1–16. http://doi.org/10.1002/acp.934

Gathercole, S., & Alloway, T. P. (2008). *Working Memory and Learning: A Practical Guide for Teachers*. SAGE Publications.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177-190. http://dx.doi.org/10.1037/0012-1649.40.2.177

Gathercole, S. E., Woolgar, F., Kievit, R. A., Astle, D., Manly, T., & Holmes, J. (2016). How common are WM deficits in children with difficulties in reading and mathematics?

Journal of Applied Research in Memory and Cognition, *5*(4), 384–394. http://doi.org/10.1016/j.jarmac.2016.07.013

Gilbert, S. J. (2015). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology*, *68*(5), 971–92. http://doi.org/10.1080/17470218.2014.972963

Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language*, *92*, 183-201.

Hamilton, K. A., McIntyre, K. P., & Hertel, P. T. (2016). Judging Knowledge in the Digital Age: The Role of External-Memory Organization. *Applied Cognitive Psychology, 30*, 1080–1087

Hanley, J. R., & Young, A. W. (in press). ELD revisited: A second look at a neuropsychological impairment of working memory affecting retention of visuo-spatial material. *Cortex.*

Hertel, P.T. (1988). Monitoring external memory. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds), *Practical aspects of memory: Current research and issues* (pp. 221-226). New York: Wiley.

Hess, S. M., Detweiler, M. C., & Ellis, R. D. (1999). The utility of display space in keeping track of rapidly changing information. *Human Factors, 41*(2), 257-281.

Holmes, J., Gathercole, S. E., & Dunning, D. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12*(4), F9-F15. DOI: 10.1111/j.1467-7687.2009.00848.x.

Holmes, J., Hilton, K. A., Place, M., Alloway, T. P., Elliott, J. G., & Gathercole, S. E. (2014). Children with low working memory and children with ADHD: same or different? *Frontiers in Human Neuroscience*, *8*(July 2015), 976. http://doi.org/10.3389/fnhum.2014.00976

Intons-Peterson, M. J., & Fournier, J. (1986). External and internal memory aids: when and how often do we use them? *Journal of Experimental Psychology: General, 115*, 267-280.

Jaroslawska, A. J., Gathercole, S. E., Allen, R. J., & Holmes, J. (2016). Following instructions from working memory: Why does action at encoding and recall help? *The Quarterly Journal of Experimental Psychology*, *8*, 1183-1191. http://doi.org/10.3758/s13421-016-0636-5

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, *73*(1–2), 31–68. http://doi.org/http://dx.doi.org/10.1016/0004-3702(94)00017-U

Kuhn, D. (2000). Metacognitive Development. *Current Directions in Psychological Science*, *9*(5), 178–181.

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research, 50*, 408-428.

Logie, R. H. (2011). The Functional Organization and Capacity Limits of Working Memory. *Current Directions in Psychological Science*, *20*(4), 240–245. http://doi.org/10.1177/0963721411415340

Melby-Lervåg, M., & Hulme, C. (2013). Is Working Memory Training Effective? A Meta-Analytic Review. *Developmental Psychology*, *49*(2), 270–291. http://doi.org/10.1037/a0028228

Monette, S., Bigras, M., & Guay, M. C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, *109*(2), 158–173. http://doi.org/10.1016/j.jecp.2011.01.008

Morey, C. (2018). The case against specialized visual-spatial short-term memory. *Psychological Bulletin*, 144, 849-883.

Morey, C. C., Mareva, S., Lelonkiewicz, J. R., & Chevalier, N. (2018). Gaze-based rehearsal in children under 7: a developmental investigation of eye movements during a serial spatial memory task. *Developmental Science*, *21*(3), e12559.

Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: implications for models of working memory. *Journal of Memory and Language, 69*(3), 324-348.

Peirce, J. W. (2007) PsychoPy - Psychophysics software in Python. *J Neurosci Methods, 162*(1-2), 8-13.

Poole, B. J. & Kane, M. J. (2009). Working-memory capacity predicts the executive control of visual search among distractors: the influences of sustained and selective attention. *The Quarterly Journal of Experimental Psychology, 62*(7), 1430-1454.

R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Ragan, E. D., Bowman, D. A.,, & Huber, K. (2012). Supporting cognitive processing with spatial information presentations in virtual environments. *Virtual Reality*, *16,* 301-314. DOI:10.1007/s10055-012-0211-8

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General, 133*,46–62.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive Offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. http://doi.org/10.1016/j.tics.2016.07.002

Russell, J., Jarrold, C., & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *37*(6), 673–686. http://doi.org/10.1111/1469-7610.ep9702145190

Sala, G., & Gobet, F. (2017). Working Memory Training in Typically Developing Children: A Meta-Analysis of the Available Evidence. *Developmental Psychology*. Advance online publication. http://dx.doi.org/10.1037/dev0000265

Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding word learning in space. *PLoS One, 6*(12), 1–13.

Schmidt, R. A. (1994). *Motor learning and performance*. Champaign; Human Kinematics.

Schneider, W. (2008). The Development of Metacognitive Knowledge in Children and Adolescents: Major Trends and Implications for Education. *Mind, Brain, and Education. 2*(3), 114–121. http://doi.org/10.1111/j.1751-228X.2008.00041.x

Scribner, S. (1986). Thinking in action: some characteristics of practical thought. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical Intelligence: Nature and origins of competence in the everyday world.* Cambridge University Press.

Shapiro, L. (2010). *Embodied Cognition*. Taylor & Francis. Retrieved from http://books.google.co.uk/books?id=Msi5OZek9vwC

Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting Attention Within Visual Short-Term Memory: Development and Mechanisms. *Child Development*, *85*(2), 578–592. http://doi.org/10.1111/cdev.12150

Siegler, R. S. (1996). A grand theory of development. *Monographs of the Society for Research in Child Development, 61*, 266-275.

Smits-Englemsan, B C M., Sugden, D, & Duysens, J (2006). Developmental trends in speed accuracy trade-off in 6- to 10-year-old children performing rapid reciprocal and discrete aiming movements. *Human Movement Science, 25*, 37-49

St Clair-Thompson, H. L. (2011). Executive functions and working memory behaviours in children with a poor working memory. *Learning and Individual Differences*, *21*(4), 409–414. http://doi.org/10.1016/j.lindif.2011.02.008

St Clair-Thompson, H., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology*, *30*(2), 203–219. http://doi.org/10.1080/01443410903509259

Swanson, H. L. (2016). Word Problem Solving, Working Memory and Serious Math Difficulties: Do Cognitive Strategies Really Make a Difference? *Journal of Research in Memory and Applied Cognition, 5*, 368-383.

Swanson, H. L., Lussier, C. & Orosco, M., (2013). Effects of Cognitive Strategy Interventions and Cognitive Moderators on Word Problem Solving in Children at Risk for Problem Solving Difficulties. *Learning Disabilities Research & Practice, 28*(4), 170-183.

Waterman, A. H., Atkinson, A. L., Aslam, S. S., Holmes, J., Jaroslawska, A. and Allen, R. J. (2017). Do actions speak louder than words? Examining children's ability to follow instructions, *Memory & Cognition*, *45*(6), 877–890. doi: 10.3758/s13421-017-0702-7.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. http://doi.org/10.3758/BF03196322

Zhu, M. J. H., & Risko, E. F. (2016). Spatial habit competes with effort to determine human spatial organization. The Quaterly Journal of Experimental Psychology, 69(7), 1255-1264. http://dx.doi.org/10/1080/17470218.2016.1158301

Appendix

Correlation tables for Experiment 1 and Experiment 2

	FDR	BDR	Corsi	Ordered Condition	Random Condition
FDR	--	.378**	.246*	.429***	.581***
BDR	--	--	.365**	.285**	.342**
Corsi	--	--	--	.202^{+}	.361**
Ordered Condition	--	--	--	--	.652***

$^{+}$p = .066; *p <.05; **p<.01; ***p< .001

Table 1: Correlations for Experiment 1

	FDR	BDR	Corsi	Odd one out	Ordered Condition	Random Condition
FDR	--	.673***	.359**	.494***	.541***	.719***
BDR	--	--	.338**	.456***	.451***	.563***
Corsi	--	--	--	.452***	.293**	.314**
Odd one out	--	--	--	--	.349***	.372***
Ordered Condition	--	--	--	--	--	.665***

**p<.01; ***p< .001

Table 2: Correlations for Experiment 2