

THE COGNITION OF SEQUENCES

EDITED BY: Snehlata Jaswal

PUBLISHED IN: Frontiers in Psychology



frontiers

Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-398-6

DOI 10.3389/978-2-88945-398-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

THE COGNITION OF SEQUENCES

Topic Editor:

Snehlata Jaswal, L M Thapar School of Management, India



Sunflowers are brilliant natural exemplars of the Fibonacci sequence. Fibonacci numbers, and their relative ratios, appear in nature and in art, often underlying our cognition of beauty and proportion. The Fibonacci sequence also has important applications in computer science. Image: zcool.com.cn.

It is impossible to perceive the innumerable stimuli impinging on our senses, all at once. Out of the myriad stimuli, external and internal, a few are selected for further processing; and even among these, we try to put each in some sort of relation with the others, to be able to make some sense about them all. Time, of course, is an elementary dimension we use to organize our experiences. Thus, the perception of sequences is basic to human cognition. Nevertheless, research

addressing sequences is rather sparse. Partly, this is due to difficulty in designing experiments in this area due to huge individual differences. Then, there is the assumption that temporal order has more to do with memory than perception. Another problem is that sequences seem endemic to the auditory world. So much so that some researchers have suggested that sound provides the 'auditory scaffolding' for sequencing behavior. Little wonder that research studies addressing sequences in modalities other than audition are extremely rare.

This research topic aimed to gather a holistic picture of sequencing behaviour among humans by collecting snapshots of the current research on the topic of sequencing. We particularly sought contributions which addressed sequences beyond the auditory modality. The single unifying criteria for these diverse contributions was that they shed new light on previously unexplored empirical relationships and/or provoked new lines of research with incisive ideas regarding sequencing behavior. Seasoned researchers contributed their views on perception, memory, and production of sequences.

Citation: Jaswal, S., ed. (2018). *The Cognition of Sequences*. Lausanne: Frontiers Media.
doi: 10.3389/978-2-88945-398-6

Table of Contents

- 05 Editorial: What Next - The Cognition of Sequences**
Snehlata Jaswal
- 08 A Unified Theoretical Framework for Cognitive Sequencing**
Tejas Savalia, Anuj Shukla and Raju S. Bapi
- 18 No Apparent Influence of Reward upon Visual Statistical Learning**
Leeland L. Rogers, Kyle G. Friedman and Timothy J. Vickery
- 28 Episodic Short-Term Recognition Requires Encoding into Visual Working Memory: Evidence from Probe Recognition after Letter Report**
Christian H. Poth and Werner X. Schneider
- 42 Spatio-Temporal Structure, Path Characteristics, and Perceptual Grouping in Immediate Serial Spatial Recall**
Carlo De Lillo, Melissa Kirby and Daniel Poole
- 60 Working Memory for Sequences of Temporal Durations Reveals a Volatile Single-Item Store**
Sanjay G. Manohar and Masud Husain
- 71 Isochronous Sequential Presentation Helps Children Orient Their Attention in Time**
Katherine A. Johnson, Marita Bryan, Kira Polonowita, Delia Decroupet and Jennifer T. Coull
- 85 The Neurocognitive Performance of Visuospatial Attention in Children with Obesity**
Chia-Liang Tsai, Fu-Chen Chen, Chien-Yu Pan and Yu-Ting Tseng
- 95 The Nature of Verbal Short-Term Impairment in Dyslexia: The Importance of Serial Order**
Steve Majerus and Nelson Cowan
- 103 Spatial-Sequential Working Memory in Younger and Older Adults: Age Predicts Backward Recall Performance within Both Age Groups**
Louise A. Brown
- 114 Differences in Verbal and Visuospatial Forward and Backward Order Recall: A Review of the Literature**
Enrica Donolato, David Giofrè and Irene C. Mammarella
- 128 Commentary: Distinct neural mechanisms for remembering when an event occurred**
Sarah DuBrow and Lila Davachi



Editorial: What Next - The Cognition of Sequences

Snehlata Jaswal*

LM Thapar School of Management, Thapar University, Punjab, India

Keywords: working memory, sequences, cognition, sequential learning, sequential effects

Editorial on the Research Topic

What Next - The Cognition of Sequences

Sequences are ubiquitous in our lives, yet mysterious and difficult to research because of many confounds. The research topic “What next: The Cognition of Sequences” aimed to understand sequencing behavior by gathering theoretical and empirical articles showcasing current views. Gratifyingly, contributions not only addressed the theoretical debates but also focused on memory for sequences in special populations such as the autistic and the dyslexic, hinting at possible applications in this area. The recurring theme in the contributions is that sequencing is a process within Working Memory rather than merely a perceptual entity.

Proposing a unified theoretical framework for cognitive sequencing, Savalia et al. bring together two diverse debates in the sequencing literature—the implicit vs. explicit nature of sequencing, and the goal directed vs. habit-oriented response systems. They propose that the brain implicitly (automatically) extracts regularities from the myriad, ever changing stimuli, but uses attention to organize them in a hierarchical way. Attention is also needed to organize sequences of responses/actions to achieve a future goal, although when repeated often enough, these sequences acquire the force of habit with a concomitant release from the processes of attention. This theoretical framework serves for both humans and animals, and is perhaps best exemplified by skill acquisition.

In line with these thoughts, Rogers et al. provide empirical evidence that statistical learning of stimulus sequences is indeed implicit, being unaffected by reward contingencies. In their experiment, they found significant visual statistical learning effects, but no-, low-, or high-reward conditions did not cause any differences in the strength of learning. Thus, the amount of rewards did not affect statistical learning of sequences. They conclude that the system that detects links and regularities among stimuli, functions independently of the system that identifies reward contingencies.

Poth and Schneider explore how we remember objects from previous episodes. This could be because we remember visual features of objects or we remember the objects stored in VWM. Using a new paradigm combining letter report and probe recognition, they evaluated the dependence of episodic short term recognition on VWM. The first experiment showed that participants recognized probes more often if they had reported them earlier in a whole report. The second experiment required partial report of one letter, and probes were either for this letter, or those near it, or those far from it. Probe recognition was better for near than for far letters, indicating that episodic short term recognition is only possible for a limited number of simultaneously presented objects due to the encoding limitation of VWM. presentations.

De Lillo et al. provide evidence for VWM factors being more crucial than perceptual grouping in the retention of spatial sequences. They used variants of the Corsi task on touch screen monitors and in virtual reality to establish that serial spatial recall is least affected by path length. It is the structure or organization imposed on the stimuli which is the most important factor in the performance of the participants. Their experiments show that visual perceptual grouping factors

OPEN ACCESS

Edited and reviewed by:

Serge Thill,
Plymouth University, United Kingdom

*Correspondence:

Snehlata Jaswal
sneh.jaswal@gmail.com

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 22 October 2017

Accepted: 28 November 2017

Published: 13 December 2017

Citation:

Jaswal S (2017) Editorial: What Next -
The Cognition of Sequences.
Front. Psychol. 8:2160.
doi: 10.3389/fpsyg.2017.02160

are not necessary for the benefit of structure, and thus they conclude that encoding of structure happens at a post perceptual stage in VWM.

Manohar and Husain explored the retention of sequences of brief time intervals in the auditory modality, which were to be reproduced by the participants by holding down a key. Analogous to verbal and visuo-spatial working memory, they found an effect of set size as well as serial position effects. Attention and expectation were also significant factors. Performance was significantly worse when only one item (the lowest set size) was to be remembered, indicating the vulnerability of memory for single items. They conclude the mechanisms used to remember auditory time durations in Working Memory are similar to those used to remember verbal or visual stimuli.

Johnson et al. compared children and adults in their orientation/attention to temporal vs. spatial cues. In one experiment, location of a target was predicted by an arrow, while target onset was predicted by a short or long tone. Adults' showed a greater response latency in invalid trials as compared with valid trials in both spatial and temporal domains. However, children's (Mean age 11.4 years) responses were slowed only in the spatial domain, and were not affected in the temporal domain. In the second study, a series of sounds were presented in a rhythmic series. In this experiment, children were slowed significantly by invalid cues. Thus, sequential rather than single presentation, helped children orient attention in time.

Sequences, are however, not only important for perception of time. Deficits in sequencing are associated with other problems as well. Tsai et al. provide evidence of deficits in processing of simple cue-target sequences being associated with childhood obesity. They compared the performance of children with obesity and healthy weight controls using behavioral as well as ERP measures on a visuospatial attention task. The task used the Posner paradigm in which the children had to respond to a cue-target sequence correctly as well as quickly. Simultaneously, ERP activity was recorded. Children with obesity showed poorer behavioral performances (slower reaction times as well as deficits in attentional inhibition) and aberrant neural activity (e.g., smaller P3 amplitudes) when performing the task.

Majerus and Cowan review the evidence regarding verbal STM shortcomings in people with dyslexia. Their contention is that this STM impairment in dyslexics is primarily an inability to process serial order in working memory. This impairment is found for verbal as well as non-verbal (visuo-spatial) material. However, it is not reported by every individual who has dyslexia, nor is it specific to only dyslexics. Thus, it is not a cognitive marker of dyslexia. Nor is it clear whether and how far this impairment contributes causally to dyslexia. Finally, research is also needed to disentangle the mechanisms and effects of deficits in serial order processing and phonological processing.

Age differences in forward and backward recall are documented in an original research article by Brown. She compared young adults (18–40 years) and older adults

(64–85 years) on a modified version of the Spatial Span subtest of the Wechsler Memory Scale. Spatial interference had the maximum effect as compared to visual interference and a control condition, indicating that the task was indeed assessing spatial memory. Further, using regression analyses within each age group, she provides evidence regarding age being a significant factor in backward spatial span performance, supporting other studies which show declines in a variety of working memory tasks with increasing age. Her study also shows backward span is more sensitive to aging than forward span, presumably because it relies more heavily on processing than forward span task.

Donolato et al. review the literature regarding differences in forward and backward order recall in the verbal and visuo-spatial domains of Working Memory. They begin with the proposition that order of presentation is crucial in verbal but not visuo-spatial memory. This is particularly evident in differences in forward and backward recall, with performance being worse in backward than forward recall in verbal tasks, but not always so in visuo-spatial tasks. Nevertheless, their mini review shows that in individuals with weak visuospatial abilities, performance is worse for backward recall than for forward recall. This indicates the importance of individual differences in cognitive tasks in general and Working Memory tasks in particular.

The commentary by Dubrow and Davachi on an original fMRI research article by Jenkins and Ranganath (2016) regarding neural mechanisms underlying the memory for events is included in the research topic because it provides an insightful comparison of various mechanisms that extant literature suggests support memory for temporal order. The authors begin the review with the intuitive notion that the item memory strength gives a cue regarding which of two stimuli is the most recent one. They move on to context differentiation account of order memory. Then they discuss the theories of temporal representation based on absolute time and position at which an event occurred, and those which are based on relative time and position (associative chaining). Their paper also addresses how fMRI data can be used to test these competing viewpoints, suggesting future avenues of research.

To conclude, it seems sequences are constructed within Working Memory from the raw material provided by perception, sometimes implicitly, but more often explicitly. Further this process is important in intact cognitive processing, and aberrations are associated with behavioral/clinical problems which we are just beginning to explore. Future research is envisioned in two related avenues—determining the factors and processes in Working Memory which contribute to sequencing, and studying how to ameliorate problems of sequencing behavior through education and training in general, and treatment and rehabilitation efforts in clinical populations.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

Jenkins, L. J., and Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. *Hippocampus* 26, 554–559. doi: 10.1002/hipo.22571

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Jaswal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Unified Theoretical Framework for Cognitive Sequencing

Tejas Savalia¹, Anuj Shukla¹ and Raju S. Bapi^{1,2*}

¹ Cognitive Science Lab, International Institute of Information Technology, Hyderabad, India, ² School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology Jodhpur,
India

Reviewed by:

Rainer Schwarting,
University of Marburg, Germany
Christopher Conway,
Georgia State University, USA

*Correspondence:

Raju S. Bapi
raju.bapi@iiit.ac.in

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 14 August 2016

Accepted: 03 November 2016

Published: 18 November 2016

Citation:

Savalia T, Shukla A and Bapi RS
(2016) A Unified Theoretical
Framework for Cognitive Sequencing.
Front. Psychol. 7:1821.
doi: 10.3389/fpsyg.2016.01821

The capacity to sequence information is central to human performance. Sequencing ability forms the foundation stone for higher order cognition related to language and goal-directed planning. Information related to the order of items, their timing, chunking and hierarchical organization are important aspects in sequencing. Past research on sequencing has emphasized two distinct and independent dichotomies: implicit vs. explicit and goal-directed vs. habits. We propose a theoretical framework unifying these two streams. Our proposal relies on brain's ability to implicitly extract statistical regularities from the stream of stimuli and with attentional engagement organizing sequences explicitly and hierarchically. Similarly, sequences that need to be assembled purposively to accomplish a goal require engagement of attentional processes. With repetition, these goal-directed plans become habits with concomitant disengagement of attention. Thus, attention and awareness play a crucial role in the implicit-to-explicit transition as well as in how goal-directed plans become automatic habits. Cortico-subcortical loops basal ganglia-frontal cortex and hippocampus-frontal cortex loops mediate the transition process. We show how the computational principles of model-free and model-based learning paradigms, along with a pivotal role for attention and awareness, offer a unifying framework for these two dichotomies. Based on this framework, we make testable predictions related to the potential influence of response-to-stimulus interval (RSI) on developing awareness in implicit learning tasks.

Keywords: implicit sequence learning, explicit sequence knowledge, habitual and goal-directed behavior, model-free vs. model-based learning, hierarchical reinforcement learning

1. INTRODUCTION

Cognitive Sequencing can be viewed as the ability to perceive, represent and execute a set of actions that follow a particular order. This ability underlies vast areas of human activity including, statistical learning, artificial grammar learning, skill learning, planning, problem solving, speech and language. Many human behaviors ranging from walking to complex decision making in chess involve sequence processing (Clegg et al., 1998; Bapi et al., 2005). Such sequencing ability often involves processing repeating patterns—learning while perceiving the recurrent stimuli or actions and executing accordingly. Sequencing behavior has been studied in two contrasting paradigms: *goal-directed* and *habitual* or under the popular rubric of *response-outcome* (R-O) and *stimulus-response* (S-R) behavior. A similar dichotomy exists on the computational side under the alias of *model-based* vs. *model-free* mechanisms. The model-based vs. model-free computational paradigm has proved vital in designing algorithms for planning and learning in various intelligent

system architectures—leading to the proposal of their involvement in human behavior as well. In this article, we use another dichotomy on the learning side: *implicit* vs. *explicit* along with a pivotal role for attention and awareness to connect these dichotomies and suggest a unified theoretical framework targeted toward sequence acquisition and execution. In the following, the three dichotomies will be described along with a summary of the known neural bases of these.

1.1. Habitual vs. Goal-Directed Behavior

Existence of a combination of habitual and goal-directed behaviors is shown in empirical studies on rats and humans. In the experiments to study these behaviors two phenomena have been used to differentiate: *outcome devaluation*—sensitivity to devaluation of the goal and *contingency degradation*—sensitivity to an omission schedule. Outcome devaluation is achieved by satiating the rats on the rewarding goal; making the reward less appealing whereas contingency degradation is achieved by omitting a stimulus within a sequence of stimuli leading to the goal. Results demonstrate that overtrained rats and humans seem to be insensitive to both the phenomena. That is, even though the outcome of following a path is devalued or a stimulus in the sequence is omitted, habits lead rats to follow the same path, thus relating overtraining to habitual or stimulus-response (S-R) kind of control (Adams and Dickinson, 1981; Killcross and Coutureau, 2003). On the other hand, moderately trained rats have had little or no difficulty adapting to the new schedule relating this behavior to a goal-directed or response-outcome (R-O) kind of control (Dickinson, 1985; Balleine and Dickinson, 1998; Tricomi et al., 2004; Balleine and O'Doherty, 2010; Dolan and Dayan, 2013). Based on this proposal of two contrasting mechanisms, quite a few notable neuroimaging studies have attempted establishing the neural substrate related to the two modes of control. fMRI studies related to outcome devaluation point to two sub areas of the ventromedial prefrontal cortex (vmPFC)—medial orbito-frontal cortex (OFC) and medial prefrontal cortex (PFC) as well as one of the target areas of the vmPFC structures in the human striatum, namely, the anterior caudate nucleus, to be involved in goal-directed actions (Valentin et al., 2007; Balleine and O'Doherty, 2010). Studies aimed at finding the neural substrate for habitual behavior suggest an involvement of the subcortical structure, dorsolateral striatum (Hikosaka et al., 1999; Yin and Knowlton, 2006; Tricomi et al., 2009).

1.2. Model-Free vs. Model-Based Paradigm

In order to understand the learning process in goal-directed and habitual paradigms, two contrasting computational theories have been proposed. Goal-directed learning and control processes have been related to a *model-based* system (Doya et al., 2002; Khamassi and Humphries, 2012) whereas the habitual paradigm to a *model-free* system (Dolan and Dayan, 2013). Typically, a goal-directed system uses its view of the environment to evaluate its current state and possible future states, selecting an action that yields the highest reward. A model-based mechanism conceives this as building a search tree leading toward goal states. Such a system can be viewed as using the past experiences to understand

the environment and using this view of the environment to predict the future.

In contrast, a habitual behavior can be viewed as a repetition of action sequences based on past experience. Acquisition of a habit can be viewed as learning a skill—correcting the residual error in the action sequence leading to the goal. Typically skills are acquired over time, are sensitive to the effectors used for acquisition (Bapi et al., 2000) and there seems to be a network of brain areas involved at various stages of skill learning (Hikosaka et al., 1999). A model-free system conceives the residual error as a sort of prediction error called the *temporal difference* (TD) error (Sutton, 1988). Dopamine has been noted to be a key neurotransmitter involved in encoding the prediction error and thus leading to the view that dopamine plays a crucial role in habit learning (Schultz et al., 1997). The influence of dopamine is in no way limited to habitual behaviors. Role of dopamine in functions mediated by the prefrontal cortex such as working memory, along with the observation that there are wide projections of dopamine to both the caudate and the putamen and studies manipulating dopamine levels in the prefrontal cortex affecting goal-directed behavior indicate its involvement in the goal directed mechanism as well (see Dolan and Dayan, 2013 for a review).

Further work has been directed at establishing a connection between the dichotomies—the behavioral dichotomy of goal-directed vs. habitual and the computational dichotomy of model-based vs. model-free. For example, Daw et al. (2005) suggested an uncertainty based competition between the two behaviors—the computationally simpler process at any point acting as a driver process. Another interesting aspect of such a combination comes from the *hierarchical reinforcement learning* (HRL) framework (Sutton and Barto, 1998; Sutton et al., 1999; Botvinick et al., 2009; Botvinick, 2012). An important aspect of acquisition of sequences is the formation of chunks among the sequence of stimuli. The striatum has been emphasized to be involved in the chunking process, the chunks then selected and scaled by the output circuits of the basal ganglia (Graybiel, 1998).

1.3. Implicit and Explicit learning

For the past three decades, there has been significant interest in sequence learning. A large body of experimental evidence suggests that there is an important distinction between *implicit* and *explicit learning* (see for example Howard and Howard, 1992; Shea et al., 2001 studies). Howard and Howard (1992) used a typical serial reaction time (SRT) task (Nissen and Bullemer, 1987), where the stimuli appeared in one of the four locations on the screen, and a key was associated with each location. The participants' job was to press the key as soon as possible. The stimuli were presented in a sequential order. With practice, participants exhibited response time benefits by responding faster for the sequence. However, their performance dropped to chance-level when they were asked to predict the next possible location of the stimulus suggesting that the participants might have learned the sequence of responses in an implicit manner and thus can not predict the next move explicitly. Another example is the study by Shea et al. (2001), where participants were given a standing platform task and were asked to mimic

the movements of the line presented on a screen. The order of the stimuli was designed in such a way that the middle segment was always fixed whereas the first and the last segments varied but participants were not told about the stimulus order. It was found that the performance of the middle segment improves over time. During the recognition phase participants fail to recognize the repeated segment, pointing to the possibility that they may have acquired this via implicit learning. A recent study done with a variant of SRT task called oculomotor serial response time (SORT) task also suggests that motor sequence can be learned implicitly in the saccadic system (Kinder et al., 2008) as well and does not pose attentional demands when the SORT task is performed under dual-task condition (Shukla, 2012). Another way of differentiating implicit vs. explicit learning would be to see whether an explicit instruction was given about the presence of a sequence prior to the task. An instruction specifying the presence of a sequence in the task would, in turn, drive attentional learning. Without such explicit prior knowledge, however, it may take more number of trials for the subjects to become aware of the presence of a sequence—requiring them to engage their attention toward the sequence, turning the concomitant learning and execution explicit. Apart from these studies, there is a large body of clinical literature which confirms the distinction of implicit and explicit learning. Most of the clinical evidence comes from the *artificial grammar learning* (AGL) paradigm, where patients learned to decide whether the string of letters followed grammatical rules or not. Healthy participants were found to learn to categorize grammatical and ungrammatical strings without being able to verbalize the grammatical rules. Evidence from amnesic patients points toward implicit learning being intact in patients even though their explicit learning was severely impaired (Knowlton et al., 1992; Knowlton and Squire, 1996; Gooding et al., 2000; Van Tilborg et al., 2011). Willingham et al. (2002) suggested that activation in the left prefrontal cortex was a prerequisite for such awareness along with activation of the anterior striatum (Jueptner et al., 1997a,b). Results of the positron emission tomography (PET) study of Grafton et al. (1995) when participants performed a motor sequence learning task under implicit or explicit learning task conditions suggest that the motor cortex and supplementary motor areas were activated for implicit learning whereas the right premotor cortex, the dorsolateral cingulate, anterior cingulate, parietal cortex and also the lateral temporal cortex were associated with explicit procedural memories (Gazzaniga, 2004; Destrebecqz et al., 2005).

It has been established that the brain areas involved in working memory and attentional processing are more active during explicit learning as compared to implicit learning. Further, the findings of functional magnetic resonance imaging (fMRI) studies suggest that the prefrontal and anterior cingulate cortex and early visual areas are involved in both implicit and explicit learning (Aizenstein et al., 2004). However, there is a greater prefrontal activation in case of explicit processing than implicit which is consistent with the findings from attention literature suggesting that prefrontal activation is associated with controlled and effortful processing (Aizenstein et al., 2004). However, the neural bases of implicit and explicit learning are still inconclusive. For example, Schendan et al. (2003) used fMRI

to differentiate brain activation involved in implicit and explicit processing. Their finding suggests that the same brain areas are activated in both types of processing. More specifically, the medial temporal lobe (MTL) is involved in both implicit and explicit learning when a higher order sequence was given to the participants. Furthermore, Pammi et al. (2012) observed a shift in fronto-parietal activation from anterior to posterior areas during complex sequence learning, indicating a shift in control of sequence reproduction with help of a chunking mechanism.

In this section we discussed the three dichotomies that have stayed mostly distinct in the literature. While there have been many significant attempts at combining goal-directed behavior with model-based mechanism and habitual behavior with model-free mechanism, we attempt to add the third implicit vs. explicit dichotomy to devise a unifying framework explaining both learning and execution.

2. COMPUTATIONAL EQUIVALENTS

In this section we present how explicit learning and goal directed behavior can be related to a model-based mechanism whereas implicit learning and habitual behavior can be related to a model-free system. Indeed, there have been previous such attempts at bringing together the contrasting paradigms (Doya et al., 2002; Daw et al., 2005; Dezfouli and Balleine, 2012; Dolan and Dayan, 2013; Dezfouli et al., 2014; Cushman and Morris, 2015).

2.1. Goal Directed Behavior As Model-Based Mechanism

A goal directed behavior can be viewed as keeping the end-point (goal) in mind and selecting the ensuing actions accordingly. This kind of learning and control can be explained by a simple markov decision process (MDP) framework. Typically, an agent estimates its environment and calculates the value of its current state and possible future states. This estimation can be described by the Bellman equations:

$$V(s_n) = \sum T(s_n, a, s_{n+1})[R(s_n, a, s_{n+1}) + \gamma V(s_{n+1})] \quad (1)$$

Here, $T(s_n, a, s_{n+1})$ is the transition probability: the probability of the agent landing in state s_{n+1} if it takes an action a from state s_n . $R(s_n, a, s_{n+1})$ denotes the reward the agent gets on taking an action a from state s_n and lands in state s_{n+1} , γ the discount factor enabling the agent to select higher-reward-giving actions first. Such a system can be viewed as building a search tree and looking forward into the future, estimating values of the future states and selecting the maximum valued one (denoted as $V^*(s_n)$).

$$V^*(s_n) = \max_a \sum T(s_n, a, s_{n+1})[R(s_n, a, s_{n+1}) + \gamma V(s_{n+1})] \quad (2)$$

This kind of behavior requires the agent to know the transition probabilities along with the reward function. While these values are not directly available in the environment, the agent builds these gradually over time while interacting with the environment—learning the transition probabilities iteratively. These values, in effect, form a model of the environment, hence

deriving its name. Rewards in all the future states are estimated by propagating back the rewards from the final goal state. The transition probabilities can be initially equal for all actions and subsequently getting refined with experience.

2.2. Habitual Behaviors As Model-Free System

Habitual behavior can be viewed as a typical S-R behavior, where the end-goal does not influence the current action selection directly. Instead, previous experiences of being in a particular state are cached (Daw et al., 2005). This can be conveyed by a model-free system through the well established temporal difference (TD) learning. TD learning follows the following update rules.

$$p_k = R(s_n, a, s_{n+1}) + \gamma V(s_{n+1}); \quad (3)$$

$$V(s_n) = (1 - \alpha)V(s_n) + (\alpha)p_k$$

$$V(s_n) = V(s_n) + \alpha(V(s_n) - p_k) \quad (4)$$

Here α is the learning rate. p_k encodes a sample evaluation of state s_n when the agent enters state s_n for the k th time in the form of a sum of two terms—the first term indicating reward the agent would receive from the current state and the second term computing discounted value of the next state s_{n+1} that it would enter, the value being returned is the agent's version of the value function $V(\cdot)$. The last term of Equation (4) refers to the prediction error signal. The definition of terms such as $R(\cdot)$ and $V(\cdot)$ are as defined earlier in Bellman equations. This system can be viewed as looking into the past—making a small adjustment to optimize performance and taking the next action. There is no explicit model of the system, the agent learns on-line—learning while performing.

Our proposed architecture attempts a combination of the two contrasting paradigms. We suggest that implicit learning and control can be viewed in a similar way as habitual behavior and in turn both can be modeled using a model-free computational system. Similarly, explicit learning and control seem to have similar requirements as goal-directed behavior and in turn both can be understood as using a model-based computational system. We aim to exploit the hierarchical reinforcement learning architecture and chunking phenomenon to propose how these contrasting dichotomies can be combined into a unified framework in the next section.

3. UNIFIED THEORETICAL FRAMEWORK

In a model-based mechanism—searching in a tree of possible states—as one looks further ahead into the future, the search tree starts expanding exponentially, making such a search computationally infeasible. Whereas in case of a model-free mechanism, the system has to be in the exact same state as it was before to enable an update in its policy. To make such an update account for something substantial, there have to be enough samples of a particular state which might take a larger number of trials exploring the entire state space. The respective inefficiency of the individual systems (Keramati et al., 2011) and

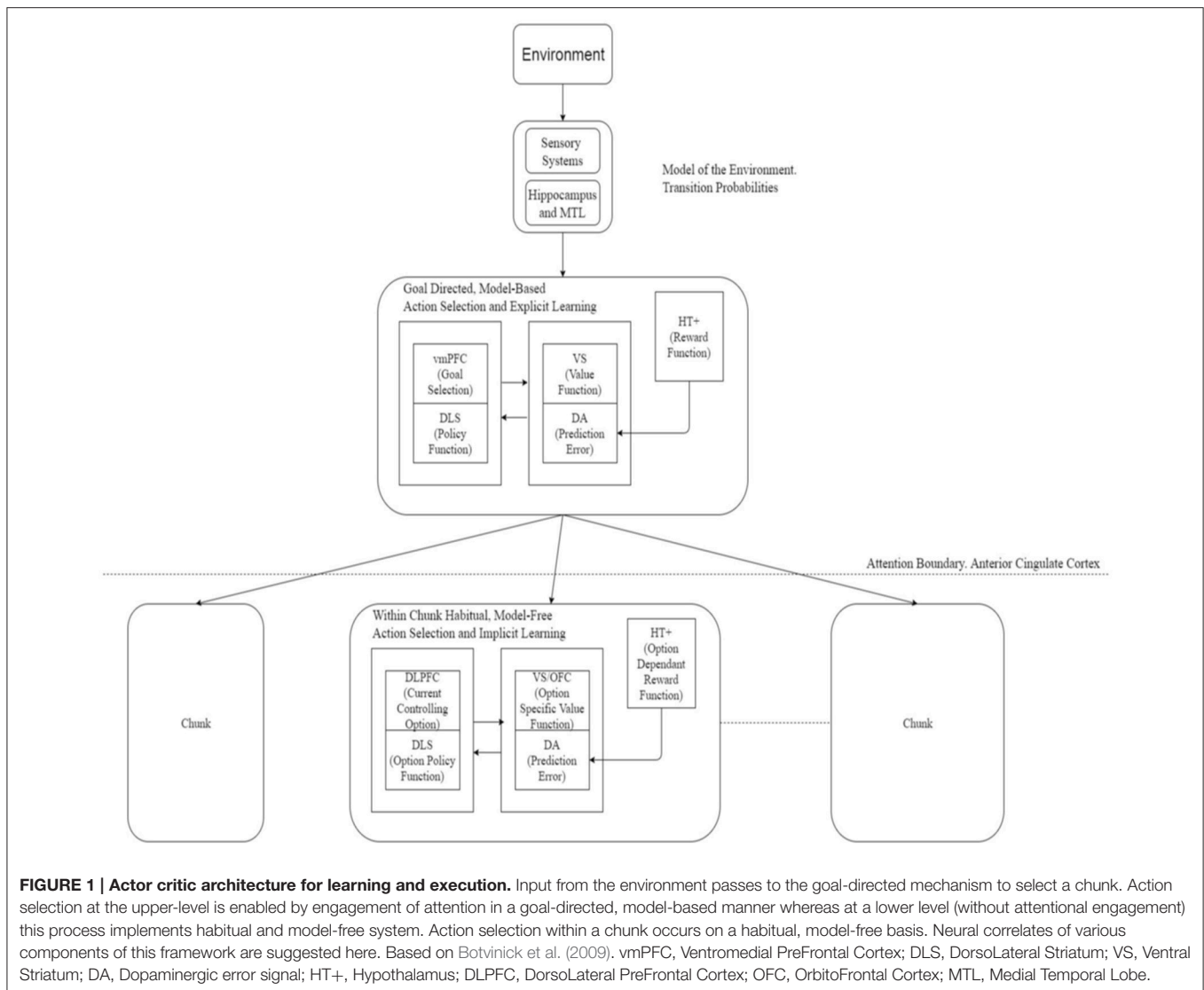
the evidence of existence of both as part of a continuum allow us to formulate a hybrid scheme combining both the computational mechanisms to explain sequence acquisition and execution in the brain.

In an attempt to formulate a unifying computational theory, we add the learning factor—implicit learning conceived as model-free learning whereas explicit learning conceived as model-based. One such idea in computational theories that suits our needs is the hierarchical reinforcement learning (HRL) framework (Sutton and Barto, 1998; Sutton et al., 1999). The HRL framework gives an additional power to the agent: the ability to select an “option” along with a primitive action for the next step. An option is a set of sequential actions (a motor program) leading to a subgoal. The agent is allowed to have separate policies for different options—on selection of an option, the agent follows that option's policy; irrespective of what the “external” policy for the primitive actions is.

We propose that learning within an option—the policy of the primitive actions within an option occurs in a model-free way. The most granular set of actions a human performs are learned by a habitual mechanism and implicitly. As one moves to learning of a less granular set of actions the roles start to change—a habitual model-free learning gradually transforms to a goal-directed one. At some point, one becomes aware of the recurring patterns being experienced and the attentional processes thereafter enable a shift from implicit state to explicit learning. Indeed, in the serial reaction time studies, it has been observed that as the subjects became aware of the recurring pattern or sequence, their learning might have moved from implicit to explicit state. We attribute this conversion to explicit learning to the formation of *explicit motor programs* or *chunks*. Chunks are formed when the subject becomes aware of the sequential pattern and implicit, model-free learning then turns into an explicit and model-based learning process. One interesting theory that can be used to explain the chunking process is the *average reward Reinforcement Learning* model (Dezfouli and Balleine, 2012).

As depicted in **Figure 1**, a similar analogy can be applied in control or performance of sequences with a change in direction in the process described above. The most abstract, top level goals are executed explicitly in a goal-directed way using model-based mechanism, the goal directed mechanisms gradually relinquishing control as the type of actions proceed downward in the hierarchy. At the finest chunk-level, subject loses awareness of the most primitive actions executed and those are then executed entirely in a habitual, model-free, implicit manner.

In neural terms, the ventromedial prefrontal cortex (vmPFC) along with the caudate nuclei may be involved in the goal-directed part, the dorsolateral striatum and dorsolateral prefrontal cortex (DLPFC) may be engaged in the habitual part—dopamine providing the prediction error signal while the anterior regions of the striatum and left prefrontal and medial frontal cortex playing a role in attentional processes. Gershman and Niv (2010) suggested a role for the hippocampus in task structure estimation which could be extended to estimating the world model and hence the transition probabilities required for the model-based system. Neural correlates for the options framework are detailed in **Figure 1**.



In this section we presented our unifying framework for combining the three dichotomies. In the subsequent section we attempt to specify the roles of response-to-stimulus interval (RSI) [or more generally, inter-stimulus interval (ISI)] and prior information along with the pivotal role of attention in switching between the two contrasting mechanisms in the explicit vs. implicit and goal-directed vs. habitual dichotomies.

4. ROLE OF RESPONSE-TO-STIMULUS-INTERVAL (RSI) AND PRIOR INFORMATION IN THE UNIFIED FRAMEWORK

A model-based search leading to explicit learning is typically slower—subject is required to deliberate over possible choices leading to the goal. In contrast, subject does not need to think

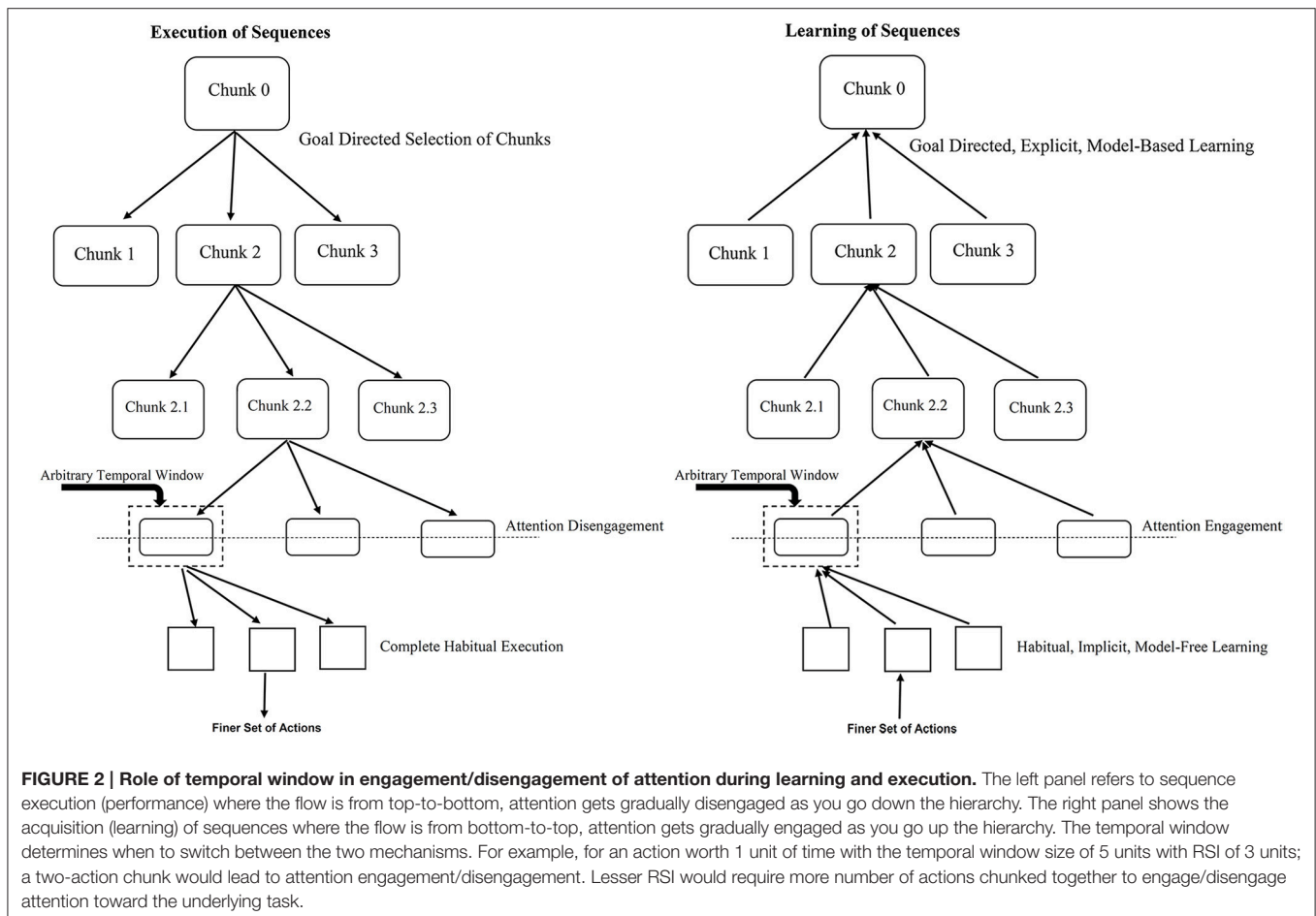
while performing an action habitually or learning implicitly—a model-free mechanism does not deliberate, it performs an action based on an already available “cache” of previous experiences and updates the cache as it proceeds further. Based on this, we propose that *response-to-stimulus interval* (RSI) [or more generally, inter-stimulus interval (ISI)] plays a key role in serial reaction time (SRT) experiments. Larger RSIs allow the subject enough time to form a model of the system, deliberate over the actions and hence this kind of learning and control corresponds to a model-based (explicit) system. On the other hand smaller RSIs do not allow the subject to form an explicit model and as is well known from the literature of serial reaction time experiments, subjects do remain sensitive to (implicitly acquire) the underlying sequential regularities (Robertson, 2007). This sort of implicit learning can be explained with temporal difference (TD) learning, where the error signal leads to an adjustment in action selection keeping the general habitual control the same.

Further, knowledge (prior information) about the existence of sequential regularities in the SRT task leads to the learning and control being explicit and model-based. This can be said to engage attentional processes in our proposal. With attentional engagement, habitual control ceases to exercise control over behavior. While we propose attention-mediated arbitration between model-based and model-free systems, Lee et al. (2014) suggest that such mediation is driven by estimates of reliability of prediction of the two models. Emerging awareness of the presence of a sequence plays a similar role in mediating learning as explicit attentional processes. The complete architecture is depicted in **Figure 2**.

Our proposal relies heavily on the *hierarchical chunking* mechanism and engagement or disengagement of attention to the underlying repeating pattern or sequence. While learning which begins implicitly and in a model-free manner, eventually as the formation of chunks proceeds up the hierarchy, at some point, the size of chunk—defined in terms of the time it takes to execute the set of actions within the chunk, crosses a threshold thus engaging the attentional resources of the subject. At this point explicit model-based learning starts taking control. Similarly, during control (or execution) of a sequence, the top most selection of chunks happens via a goal-directed, model-based

mechanism, on proceeding down the chunk hierarchy after the point of crossing some chunk size threshold, the subject no longer pays attention to the execution—it goes on in a habitual, model-free manner. Learning or execution of a set of actions within a chunk proceeds in a habitual, model-free fashion, – which at “attentive” level in the hierarchy can be explained by a habitual control of goal selection as suggested by Cushman and Morris (2015).

Attention engagement or disengagement occurs when the chunk size is equivalent to a certain *temporal window*. Such a temporal window includes the RSI for a typical SRT task. For instance, larger RSIs need fewer physical actions to reach the threshold size of the temporal window during bottom-up learning and hence cause attentional engagement toward the underlying sequential pattern sooner than in case of a trial with smaller RSIs. Based on this proposal, it will be interesting to empirically investigate the impact of varying the size of temporal window and studying resultant influence on the awareness levels of the presence of an underlying sequence in the standard SRT task. According to our proposal, implicit (associative) learning in the lower-level of the hierarchy proceeds without engagement of attention. Further we propose that as the response-stimulus interval (RSI) increases the width of the



temporal window available for integration of information related to the previous response and the subsequent stimulus increases. Thus, increasing the temporal window allows deliberative and reflective (analytical) processes to kick in, enabling a transition to explicit (awareness-driven) top-down mechanisms. This prediction can be verified experimentally and seems to be supported by preliminary evidence from the work of Cleermans et al. (see Cleermans 2014).

Such a hierarchical chunking mechanism for behavior generation has been suggested by Albus (1991), albeit from intelligent control architecture perspective. According to Albus (1991), a hierarchy producing intelligent behaviors comprises seven levels covering at the lowest level, the finest reflex actions, and spanning all the way up to long term planning of goals. At each higher level, the goals expand in scope and planning horizons expand in space and time. In neural terms, the motor neuron, spinal motor centers and cerebellum, the red nucleus, substantia nigra and the primary motor cortex, basal ganglia and prefrontal cortex and finally the temporal, limbic and frontal cortical areas get involved in increasing order of the hierarchy.

5. COMPARISON WITH OTHER DUAL SYSTEM THEORIES

Many dual system theories related to goal-directed vs. habitual behavior or implicit vs. explicit learning have been proposed in the recent past. For example, Keele et al. (2003) suggest a dual system where implicit learning is typically limited to a single dimensional or a unimodal system whereas explicit learning involves inputs from other dimensions as well. Our model incorporates this duality in a different sense and does not distinguish the dichotomy between different modalities. Inputs from multiple modalities are treated as actions in an abstract sense and when a bunch of such actions crosses the threshold (acquisition or execution time), this would lead to attentional modulation (engagement in the case of acquisition or disengagement in the case of performance). A similar idea has been discussed by Cleermans (2006) who suggested that a representation obtained from exposure to a sequence may become explicit when strength of activation reaches a critical level. Formation of the chunks is, however, assumed to be driven by bottom-up, unconscious processes. These chunks become available later for conscious processing (Perruchet and Pacton, 2006). We concur with the suggestions of Keele et al. (2003) on the neural correlates of implicit and explicit learning; learning in the dorsal system being implicit whereas that in the ventral system may be related to explicit or implicit modes. However, we emphasize that the ventral system—when learning is not characterized as a uni- or multi-dimensional dichotomy—would be more related to explicit learning. Daltrozzo and Conway (2014) discuss three levels of processing: an abstract level storage for higher level goals, followed by an intermediate level encoding of the actions required to reach the goal and a low level acquiring highly specific information related to the exact stimulus and associated final motor action (Clegg et al.,

1998). Our model reflects such a hierarchy by breaking down the actions into a finer set of sub-actions—where the top most abstract actions or goals are decided by a goal-directed, model-based system whereas the more concrete actions are executed by a habitual, model-free system. Walk and Conway (2011) suggest a cascading account where two mechanisms interact with each other in a hierarchical manner—concrete information being encoded in a modality specific format followed by encoding of more domain-general representations. We incorporate such an interleaving phenomenon by suggesting that the actions within a chunk are carried out in a habitual, attention-free manner; the selection of such a chunk being goal-directed and attention-mediated. Thiessen et al. (2013) discuss a dual system involving an extraction and integration framework for general statistical learning. The extraction framework is implicated in conditional statistical learning—formation of chunks or associations between events occurring together. On the other hand, the integration framework is implicated in distributional statistical learning—generalization of the task at hand. We can relate the extraction framework to the implicit, habitual process and the integration framework to a goal-directed mechanism that involves creation of the model of environment using information from potentially multiple sources. Batterink et al. (2015) present evidence suggesting that though there does not seem to be an explicit recognition of statistical regularities, the reaction time task, which is deemed 50% more sensitive to statistical learning, suggests that there is in fact some statistical structure of the presented stimuli learned implicitly. Our framework agrees with the conclusion that implicit and explicit statistical learning occur in parallel, attention deciding the driver process. A similar account has been suggested by Dale et al. (2012) who state that the system initially learns a readiness response to the task in an associative fashion mapping the stimulus to a response and then undergoes a rapid transition into a “predictive mode” as task regularities are decoded. Reber (2013) suggests a role for the hippocampal-medial temporal lobe (MTL) loop in explicit memory whereas implicit memory is said to be distributed in the cortical areas. However, evidence from studies with Parkinsons patients suggests an important role for the basal ganglia in acquiring such implicit knowledge. We posit a similar role for the basal ganglia and corticostriatal loops in implicit learning; the knowledge that follows this learning may be stored throughout the cortex while keeping the role of MTL and hippocampus intact.

6. CONCLUSION AND FUTURE WORK

Sequencing is a fundamental ability that underlies a host of human capacities, especially those related to higher cognitive functions. In this perspective, we suggest a theoretical framework for acquisition and control of hierarchical sequences. We bring together two hitherto unconnected streams of thought in this domain into one framework—the goal-directed and habitual axis on the one hand and the explicit and implicit sequencing paradigms on the other, with the help of model-based

and model-free computational paradigms. We suggest that attentional engagement and disengagement allow the switching between these dichotomies. While goal-directed and habitual behaviors are related to performance of sequences, explicit and implicit paradigms relate to learning and acquisition of sequences. The unified computational framework proposes how the bidirectional flow in this hierarchy implements these two dichotomies. We discuss the neural correlates in light of this synthesis.

One aspect of applicability of our proposed framework could be skill learning. It is well known that skill learning proceeds from initially being slow, attentive and error-prone to finally being fast, automatic and error-free (Fitts and Posner, 1967). Thus, it appears that sequential skill learning starts being explicit and proceeds to be implicit from the point of view of attentional demands. At first sight, this seems to be at odds with the proposed unified framework here where the hierarchy seems to have been set up to proceed from implicit to explicit learning. However, the phase-wise progression of skill learning is consistent with the framework as per the following discussion.

It is pointed out that different aspects of skill are learned in parallel in different systems—while improvements in reaction time are mediated by implicit system, increasing knowledge of the sequential regularities accrues in the explicit system (Hikosaka et al., 1999; Bapi et al., 2000). The proposed unitary network is consistent with these parallel processes, the implicit processes operating from bottom-up and the explicit system in a top-down fashion. Key factor is the engagement and disengagement of attentional system as demarcated in **Figure 1**. One might wonder how this approach can be applied to research in non-human animals, where explicit mechanisms are difficult to be realized. Historically, while SRT research identifying implicit vs. explicit learning systems are largely based on human experiments, that of goal-directed and habitual research is based on animal experiments. The proposed framework is equally applicable for human and non-human participants. What is proposed here is that the lower-level system operates based on associative processes that allow the system to learn implicitly, respond reactively and the computations at this level are compatible with a model-free framework. On the other hand, the upper-level system is based on predictive processes that allow the system to prepare anticipatory responses that sometime cause

errors. Error-evaluation while learning and error-monitoring during control are part of this system that learns using explicit processes, enables goal-directed control of actions and the computations at this level are compatible with a model-based framework. Level of attentional engagement distinguishes these two levels as shown in **Figure 2**. Of course, non-human animals can not give verbal reports of their knowledge. The explicit system in the case of pre-verbal infants and non-human animals needs to be understood in the lines of predictive systems that can elicit anticipatory, predictive responses and learn rules and transfer them to novel tasks (Marcus et al., 1999; Murphy et al., 2008).

Finally based on this theoretical proposal, we make predictions as to how implicit-to-explicit transition might happen in serial reaction time tasks when response-to-stimulus interval (RSI) is systematically manipulated. The mathematical formulation of such a unified mechanism is yet to be established, along with a formalization of the attentional window and its relation to RSI.

AUTHOR CONTRIBUTIONS

TS, AS, and RB conceptualized the framework. TS and AS did literature review and wrote the introduction and review sections. RB added the modeling linkages. TS, AS, and RB all contributed to preparing and finalizing the manuscript.

FUNDING

This work was partially supported by Department of Science and Technology (DST), under both Indo-Trento Program of Advance Research under Cognitive Neuroscience theme (No. INT/ITALY/ITPAR-III/Cog-P(6)/2013(C) dated 08-08-2013) as well as Indo-French CEFIPRA Grant for the project Basal Ganglia at Large (No. DST-INRIA 2013-02/Basal Ganglia dated 13-09-2014)—grants awarded to RB.

ACKNOWLEDGMENTS

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

REFERENCES

- Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.* 33, 109–121. doi: 10.1080/14640748108400816
- Aizenstein, H. J., Stenger, V. A., Cochran, J., Clark, K., Johnson, M., Nebes, R. D., et al. (2004). Regional brain activation during concurrent implicit and explicit sequence learning. *Cereb. Cortex* 14, 199–208. doi: 10.1093/cercor/bhg119
- Albus, J. S. (1991). Outline for a theory of intelligence. *IEEE Trans. Syst. Man Cybern.* 21, 473–509. doi: 10.1109/21.97471
- Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419. doi: 10.1016/S0028-3908(98)00033-1
- Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69. doi: 10.1038/npp.2009.131
- Bapi, R. S., Doya, K., and Harner, A. M. (2000). Evidence for effector independent and dependent representations and their differential time course of acquisition during motor sequence learning. *Exp. Brain Res.* 132, 149–162. doi: 10.1007/s002219900332
- Bapi, R. S., Pammi, V. C., Miyapuram, K. P., and Ahmed. (2005). Investigation of sequence processing: a cognitive and computational neuroscience perspective. *Curr. Sci.* 89, 1690–1698.
- Batterink, L. J., Reber, P. J., Neville, H. J., and Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *J. Mem. Lang.* 83, 62–78. doi: 10.1016/j.jml.2015.04.004

- Botvinick, M. M., Niv, Y., and Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition* 113, 262–280. doi: 10.1016/j.cognition.2008.08.011
- Botvinick, M. M. (2012). Hierarchical reinforcement learning and decision making. *Curr. Opin. Neurobiol.* 22, 956–962. doi: 10.1016/j.conb.2012.05.008
- Cleeremans, A. (2006). “Conscious and unconscious cognition: a graded, dynamic, perspective,” in *Progress in Psychological Science Around the World Vol. I Neural, Cognitive and Developmental Issues*, eds Q. Jing, M. R. Rosenzweig, G. d’Ydewalle, H. Zhang, H. C. Chen, and K. Zhang (Hove: Psychology Press), 401–418.
- Cleeremans, A. (2014). Connecting conscious and unconscious processing. *Cogn. Sci.* 38, 1286–1315. doi: 10.1111/cogs.12149
- Clegg, B. A., DiGirolamo, G. J., and Keele, S. W. (1998). Sequence learning. *Trends Cogn. Sci.* 2, 275–281. doi: 10.1016/S1364-6613(98)01202-9
- Cushman, F., and Morris, A. (2015). Habitual control of goal selection in humans. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13817–13822. doi: 10.1073/pnas.1506367112
- Dale, R., Duran, N. D., and Morehead, J. R. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Adv. Cogn. Psychol.* 8, 196–209. doi: 10.5709/acp-0115-z
- Daltrozzo, J., and Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Front. Hum. Neurosci.* 8:437. doi: 10.3389/fnhum.2014.00437
- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560
- Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., et al. (2005). The neural correlates of implicit and explicit sequence learning: interacting networks revealed by the process dissociation procedure. *Learn. Mem.* 12, 480–490. doi: 10.1101/lm.95605
- Dezfouli, A., and Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* 35, 1036–1051. doi: 10.1111/j.1460-9568.2012.08050.x
- Dezfouli, A., Lingawi, N. W., and Balleine, B. W. (2014). Habits as action sequences: hierarchical action control and changes in outcome value. *Philos. Trans. R. Soc. B* 369:20130482. doi: 10.1098/rstb.2013.0482
- Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. B Biol. Sci.* 308, 67–78. doi: 10.1098/rstb.1985.0010
- Dolan, R. J., and Dayan, P. (2013). Goals and habits in the brain. *Neuron* 80, 312–325. doi: 10.1016/j.neuron.2013.09.007
- Doya, K., Samejima, K., Katagiri, K.-I., and Kawato, M. (2002). Multiple model-based reinforcement learning. *Neural Comput.* 14, 1347–1369. doi: 10.1162/089976602753712972
- Fitts, P. M. and Posner, M. I. (1967). *Human Performance*. Brooks/Cole Publishing Company.
- Gazzaniga, M. S. (2004). *The Cognitive Neurosciences*. New York, NY; Norton, OH: MIT Press.
- Gershman, S. J., and Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* 20, 251–256. doi: 10.1016/j.conb.2010.02.008
- Gooding, P., Mayes, A., and Van Eijk, R. (2000). A meta-analysis of indirect memory tests for novel material in organic amnesics. *Neuropsychologia* 38, 666–676. doi: 10.1016/S0028-3932(99)00119-0
- Grafton, S. T., Hazeltine, E., and Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *J. Cogn. Neurosci.* 7, 497–510. doi: 10.1162/jocn.1995.7.4.497
- Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiol. Learn. Mem.* 70, 119–136. doi: 10.1006/nlme.1998.3843
- Hikosaka, O., Nakahara, H., Rand, M. K., Sakai, K., Lu, X., Nakamura, K., et al. (1999). Parallel neural networks for learning sequential procedures. *Trends Neurosci.* 22, 464–471. doi: 10.1016/S0166-2236(99)01439-3
- Howard, D. V., and Howard, J. H. (1992). Adult age differences in the rate of learning serial patterns: evidence from direct and indirect tests. *Psychol. Aging* 7:232. doi: 10.1037/0882-7974.7.2.232
- Jueptner, M., Frith, C., Brooks, D., Frackowiak, R., and Passingham, R. (1997a). Anatomy of motor learning. II. subcortical structures and learning by trial and error. *J. Neurophysiol.* 77, 1325–1337.
- Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., and Passingham, R. E. (1997b). Anatomy of motor learning. I. frontal cortex and attention to action. *J. Neurophysiol.* 77, 1313–1324.
- Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., and Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychol. Rev.* 110:316. doi: 10.1037/0033-295X.110.2.316
- Keramati, M., Dezfouli, A., and Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7:e1002055. doi: 10.1371/journal.pcbi.1002055
- Khamassi, M., and Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and model-free navigation strategies. *Front. Behav. Neurosci.* 6:79. doi: 10.3389/fnbeh.2012.00079
- Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408. doi: 10.1093/cercor/13.4.400
- Kinder, A., Rolf, M., and Kliegl, R. (2008). Sequence learning at optimal stimulus-response mapping: evidence from a serial reaction time task. *Q. J. Exp. Psychol.* 61, 203–209. doi: 10.1080/17470210701557555
- Knowlton, B. J., Ramus, S. J., and Squire, L. R. (1992). Intact artificial grammar learning in amnesia: dissociation of classification learning and explicit memory for specific instances. *Psychol. Sci.* 3, 172–179. doi: 10.1111/j.1467-9280.1992.tb00021.x
- Knowlton, B. J., and Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *J. Exp. Psychol.* 22:169. doi: 10.1037/0278-7393.22.1.169
- Lee, S. W., Shimojo, S., and O’Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81, 687–699. doi: 10.1016/j.neuron.2013.11.028
- Marcus, G. F., Vijayan, S., Rao, S. B., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80. doi: 10.1126/science.283.5398.77
- Murphy, R. A., Mondragón, E., and Murphy, V. A. (2008). Rule learning by rats. *Science* 319, 1849–1851. doi: 10.1126/science.1151564
- Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.* 19, 1–32. doi: 10.1016/0010-0285(87)90002-8
- Pammi, V. C., Miyapuram, K. P., Ahmed, Samejima, K., Bapi, R. S., Doya, K., et al. (2012). Changing the structure of complex visuo-motor sequences selectively activates the fronto-parietal network. *Neuroimage* 59, 1180–1189. doi: 10.1016/j.neuroimage.2011.08.006
- Perruchet, P., and Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci.* 10, 233–238. doi: 10.1016/j.tics.2006.03.006
- Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia* 51, 2026–2042. doi: 10.1016/j.neuropsychologia.2013.06.019
- Robertson, E. M. (2007). The serial reaction time task: implicit motor skill learning? *J. Neurosci.* 27, 10073–10075. doi: 10.1523/JNEUROSCI.2747-07.2007
- Schendan, H. E., Searl, M. M., Melrose, R. J., and Stern, C. E. (2003). An fmri study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37, 1013–1025. doi: 10.1016/S0896-6273(03)00123-5
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Shea, C. H., Wulf, G., Whitacre, C. A., and Park, J.-H. (2001). Surfing the implicit wave. *Q. J. Exp. Psychol.* A 54, 841–862. doi: 10.1080/713755993
- Shukla, A. (2012). *Investigation of the Influence of Attention on Oculomotor Implicit Learning*. Unpublished M. Phil. Dissertation. Center for Neural and Cognitive Sciences, University of Hyderabad, Hyderabad.
- Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Mach. Learn.* 3, 9–44. doi: 10.1007/BF00115009
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*, Vol. 1. Cambridge: MIT Press.
- Sutton, R. S., Precup, D., and Singh, S. (1999). Between mdps and semi-mdps: a framework for temporal abstraction in reinforcement learning. *Artif. Intell.* 112, 181–211. doi: 10.1016/S0004-3702(99)00052-1
- Thiessen, E. D., Kronstein, A. T., and Hufnagle, D. G. (2013). The extraction and integration framework: a two-process account of statistical learning. *Psychol. Bull.* 139:792. doi: 10.1037/a0030801

- Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232. doi: 10.1111/j.1460-9568.2009.06796.x
- Tricomi, E. M., Delgado, M. R., and Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292. doi: 10.1016/S0896-6273(03)00848-1
- Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026. doi: 10.1523/JNEUROSCI.0564-07.2007
- Van Tilborg, I. A., Kessels, R. P., Kruijt, P., Wester, A. J., and Hulstijn, W. (2011). Spatial and nonspatial implicit motor learning in korsakoffs amnesia: evidence for selective deficits. *Exp. Brain Res.* 214, 427–435. doi: 10.1007/s00221-011-2841-6
- Walk, A. M., and Conway, C. M. (2011). "Multisensory statistical learning: can associations between perceptual categories be acquired," in *Paper Session Presented at the Meeting of Cognitive Science Society* (Boston, MA).
- Willingham, D. B., Salidis, J., and Gabrieli, J. D. (2002). Direct comparison of neural systems mediating conscious and unconscious skill learning. *J. Neurophysiol.* 88, 1451–1460. doi: 10.1152/jn.00461.2001
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Savalia, Shukla and Bapi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



No Apparent Influence of Reward upon Visual Statistical Learning

Leeland L. Rogers*, Kyle G. Friedman and Timothy J. Vickery

Department of Psychological and Brain Sciences, University of Delaware, Newark, DE, USA

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Steve Majerus,
University of Liège, Belgium
Maximilien Chaumon,
Humboldt University of Berlin,
Germany

*Correspondence:

Leeland L. Rogers
llrogers@psych.udel.edu

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 20 July 2016

Accepted: 13 October 2016

Published: 02 November 2016

Citation:

Rogers LL, Friedman KG and
Vickery TJ (2016) No Apparent
Influence of Reward upon Visual
Statistical Learning.
Front. Psychol. 7:1687.
doi: 10.3389/fpsyg.2016.01687

Humans are capable of detecting and exploiting a variety of environmental regularities, including stimulus–stimulus contingencies (e.g., visual statistical learning) and stimulus–reward contingencies. However, the relationship between these two types of learning is poorly understood. In two experiments, we sought evidence that the occurrence of rewarding events enhances or impairs visual statistical learning. Across all of our attempts to find such evidence, we employed a training stage during which we grouped shapes into triplets and presented triplets one shape at a time in an undifferentiated stream. Participants subsequently performed a surprise recognition task in which they were tested on their knowledge of the underlying structure of the triplets. Unbeknownst to participants, triplets were also assigned no-, low-, or high-reward status. In Experiments 1A and 1B, participants viewed shape streams while low and high rewards were “randomly” given, presented as low- and high-pitched tones played through headphones. Rewards were always given on the third shape of a triplet (Experiment 1A) or the first shape of a triplet (Experiment 1B), and high- and low-reward sounds were always consistently paired with the same triplets. Experiment 2 was similar to Experiment 1, except that participants were required to learn value associations of a subset of shapes before viewing the shape stream. Across all experiments, we observed significant visual statistical learning effects, but the strength of learning did not differ amongst no-, low-, or high-reward conditions for any of the experiments. Thus, our experiments failed to find any influence of rewards on statistical learning, implying that visual statistical learning may be unaffected by the occurrence of reward. The system that detects basic stimulus–stimulus regularities may operate independently of the system that detects reward contingencies.

Keywords: statistical learning, reward processing, reward learning, visual attention, associative learning, implicit learning

INTRODUCTION

At every moment, human cognition faces the complex task of interpreting and responding to an overwhelming amount of stimulation. One important means by which humans may cope with this constant stream of information in the world is by learning and exploiting statistical regularities ubiquitous in natural environments. Many laboratory studies have demonstrated the potential for human learning to pick up such regularities in an unsupervised fashion. For example, repeatedly experiencing one phoneme that reliably predicts another (Saffran et al., 1996), or particular visual items that reliably co-occur in time or space with others (Fiser and Aslin, 2001, 2002),

can lead to above-chance recognition rates of those regularities. This kind of statistical learning is available to us from a time shortly after birth and throughout adulthood (Saffran et al., 1996; Saffran et al., 1999), and such learning spans perceptual systems (Glicksohn and Cohen, 2013), allowing humans to automatically detect and learn rich probabilistic relationships common within real-world environments. Although statistical learning can be more complex, simple stimulus–stimulus associative relationships are an important (and most commonly studied) component of statistical learning, and these relationships can apparently be detected and learned without observers' intentions or awareness (Turk-Browne et al., 2005).

There is no shortage of evidence that visual statistical learning is a powerful and ubiquitous ability in humans (Turk-Browne and Scholl, 2009; Turk-Browne et al., 2009; Glicksohn and Cohen, 2011) as measured by recognition or familiarity with contingent stimuli. However, the consequences of experiencing statistical regularities in our environment is not limited to simply making sense of streams of information by segmenting and chunking stimuli. For instance, performance is improved for target items that are predicted by preceding elements in a visual statistical learning stream (Turk-Browne et al., 2010). Evidence also suggests that statistical regularities bias attention, with attention being drawn to regions in which statistical regularities occur (Zhao et al., 2013). Attention can also be guided to locations based upon implicitly learned associations between distractor and target positions (Chun and Jiang, 1998). Thus, visual statistical learning is considered useful in new environments for both recognition and for the guidance of attention.

Another type of associative learning that also possesses the capacity to guide attention is stimulus–reward learning. Reliable associations between stimuli and rewards have been shown to influence performance in many different contexts, and there is a rich history of animal studies showing strong influences of primary reward associations (Schultz et al., 1997; Berridge, 2007; Haber and Knutson, 2009) guiding behavior and driving brain activity, where primary rewards are water, juice, or food rewards that are directly registered as rewards by brain circuits as a function of states such as thirst and hunger. However, secondary rewards (e.g., money, or stimuli indicating monetary value or simply “positive” outcomes) can also be highly effective at driving performance and brain activity, and are effective as stimuli which, when reliably paired with a previously non-rewarding stimulus, imbue that previously non-rewarding stimulus with value (Daw and Doya, 2006; Haber and Knutson, 2009).

A large and growing literature using human subjects has employed such secondary cues in order to imbue previously non-rewarding stimuli with value, leading to striking differences in performance related to differences in stimulus–value associations. Higher (explicitly learned) associative value, based on secondary reward in terms of monetary value, leads to better explicit recognition memory, and high value associations can even lead to stimuli escaping the attentional blink (Raymond and O'Brien, 2009), suggesting that such

associations drive low-level attentional biases. Even in cases where participants are not consciously aware of the association between stimulus characteristics and rewarding outcomes, evidence suggests a clear attentional bias toward stimuli that are consistently paired with higher secondary rewards (Anderson et al., 2011; Sha and Jiang, 2015). Stimulus–reward learning may allow for the optimization of behavior by automatically orienting attention towards reward-predicting elements of a scene, and thus help optimize choice behavior to seek reward and avoid punishment (Engelmann et al., 2009; Hickey et al., 2010; Theeuwes and Belopolsky, 2012; Chelazzi et al., 2013; Sali et al., 2014; Pessoa, 2015). Mounting evidence suggests that even secondary cues to value can serve as associative markers that drive value-based differences in low-level performance.

These two types of learning, statistical learning and stimulus–reward learning based upon secondary reward cues, bear some obvious similarities. Visual statistical and reward learning mechanisms incorporate similar associative mechanisms. Indeed, in many published cases in which stimulus–reward learning plays a significant role, a visual stimulus is typically repeatedly paired with another specific sensory stimulus that indicates reward value – thus, both statistical learning and reward-related learning could play a role in such studies. As reviewed above, both types of learning appear to play a role in biasing selective attention. Studies of the neural bases of these mechanisms provide further reason to suspect that they may be interrelated. Visual statistical learning may be supported by some of the same neural structures that support reward learning; correlates of both reward learning and visual statistical learning have been noted in striatum and medial temporal lobe structures (Delgado et al., 2000; Aron, 2004; Wittmann et al., 2007), and there is increasing evidence that the hippocampus plays a role in reward learning as well as trial-and-error learning (Lansink et al., 2009; Dickerson and Delgado, 2015). Thus, the visual statistical learning system bears some resemblance to the value-learning system, in terms of the importance of prediction and deviations from predicted events in generating surprise signals, but the relationship between the systems is currently not well characterized.

Given the similar nature of these two types of learning and how they appear to contribute to our ability to learn about and navigate environments, an intuitive question arises – how do these two mechanisms interact? To our knowledge, potential relationships between reward and visual statistical learning remain unexplored. Even if they do not depend upon shared mechanisms, there is reason to believe that they may interact. To wit, evidence suggests that statistical learning is dependent upon selective attention to the constituent, related items (Turk-Browne et al., 2005). If rewarding events drive attention (Jiang et al., 2013), then stimulus–stimulus learning might reasonably be expected to show a dependence upon co-occurrence of constituent stimuli with rewarding events, with more-rewarding events drawing greater attention and leading to stronger memory traces. On the other hand, rewarding events and contingencies

might draw attention away from stimulus–stimulus relationships, or occupy resources otherwise required for stimulus–stimulus associative learning, thus impairing such learning.

The current set of experiments seeks to identify whether statistical learning operates independently or not from reward. Specifically, by introducing rewarding events and stimulus–reward associations while simultaneously establishing stimulus–stimulus statistical associations (Experiment 1), or by establishing reward associations immediately before establishing statistical associations (Experiment 2), we sought evidence that reward associations either enhance or impair the ability to detect statistical regularities across time.

GENERAL MATERIALS AND METHODS

Ethics Statement

These studies were carried out with full review and approval by the Institutional Review Board at the University of Delaware with written informed consent from all participants.

Participants

A total of 136 University of Delaware students took part in the study in partial fulfillment of course credit. Experiment 1A included 32 participants and Experiment 1B included 43 participants. Experiment 3 included a total of 61 participants divided into three groups: a first position, second position, and third position reward-associate groups. The first position group contained 22 participants, the second position group contained 18 participants, and the third position group contained 21 participants.

Stimulus Materials

Visual stimuli were 27 symbols that were novel and unfamiliar to our sample. These symbols, derived from the African Ndjuká syllabary and unfamiliar to our Western subjects, were adopted based upon recent research that successfully utilized them to explore visual statistical learning (Turk-Browne et al., 2009; Zhao et al., 2013; Yu and Zhao, 2015). For every participant, all 27 symbols were randomly assigned to 9 different triplet sets (see **Figure 1** for an example). Triplet sets were then randomly assigned to high-value, low-value, and no-value association conditions (i.e., three triplet sets were assigned to each condition).

Apparatus

All experiments were run using a computer running Ubuntu Linux and attached to a 17-inch CRT monitor. Experiment 1 was written in Python, using the PsychoPy package (Peirce, 2007) while Experiment 2 was written in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007).

Procedure

Participants were given written and oral instructions before each experiment. Critically, participants were only given explicit

instructions related to the familiarization phase of the experiment prior to beginning. We intentionally avoided providing any information about the underlying structure (i.e., triplets) within the familiarization phase. Participants were not informed that there would be a test phase following the familiarization phase. After an explanation of the familiarization phase cover task, participants were seated in front of a computer within an isolated and dimly lit testing space. All stages of the experiment were accompanied by a full set of instructions for the participant to read on-screen. Participants were also told to ask the experimenter for clarification on any set of instructions while completing the experiment, as needed.

Familiarization Phase

Figure 2A provides an example of the familiarization phase. Participants viewed a stream of symbol triplets on the computer screen. The presentation of the triplets were randomized. Stimuli sequentially appeared at the center of the screen for 800 ms each, with a 200 ms blank screen inter-trial interval. Each triplet appeared 24 times throughout the familiarization phase, for a total of 648 symbol presentations (not counting 24 1-back repetitions that occurred in Experiment 1). Additionally, participants were required to make responses by pressing the spacebar whenever there was a 1-back repetition (Experiment 1) or if a shape quickly moved back and forth (“jiggle” in Experiment 2). These events occurred 24 times for each individual. Despite the fact that the stream was composed of repetitions of structured triplets containing three symbols, there was no explicit indication that this structure was present. Rather, participants viewed a steady stream of characters throughout the experiment and were expected to implicitly learn the statistical regularities hidden within the stream over the course of the experiment.

In Experiment 1, tones indicating reward status (low-value or high-value), were played through headphones, and occurred on 75% of trials in which a low-value or high-value triplet appeared. In Experiment 1A, the tone always co-occurred with the third item of the triplet sequence, while in Experiment 1B, the tone always co-occurred with the first item of the triplet sequence. In Experiment 2, reward status was established through an explicit reward-learning phase that preceded the familiarization phase (described in section “Experiment 2”).

Test Phase

Figure 2B provides an example of the test phase. After the familiarization phase and before the test phase, participants were informed that the stream of characters they had just viewed contained structured triplets. The instructions continued on to explain the test phase. On each trial of the test phase, participants were shown two sequences (i.e., triplets) and had to choose the sequence that appeared more familiar to them. At the beginning of the trial, participants viewed the words “Sequence 1” on the screen for 1000 ms, followed by a central fixation cross presented for 500 ms. Three stimuli then appeared on screen with identical timing to the familiarization phase. After the first sequence had completed, a second sequence with the preceding label “Sequence 2” appeared on screen.

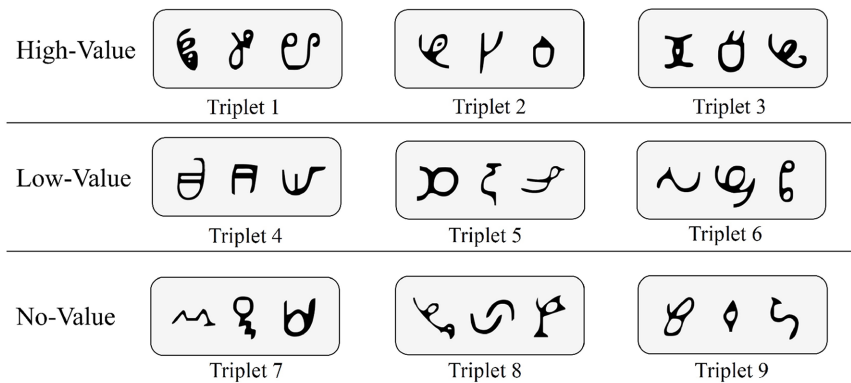


FIGURE 1 | Sample base triplets of Ndjuká symbols with random assignment to high-value, low-value, and no-value association conditions.

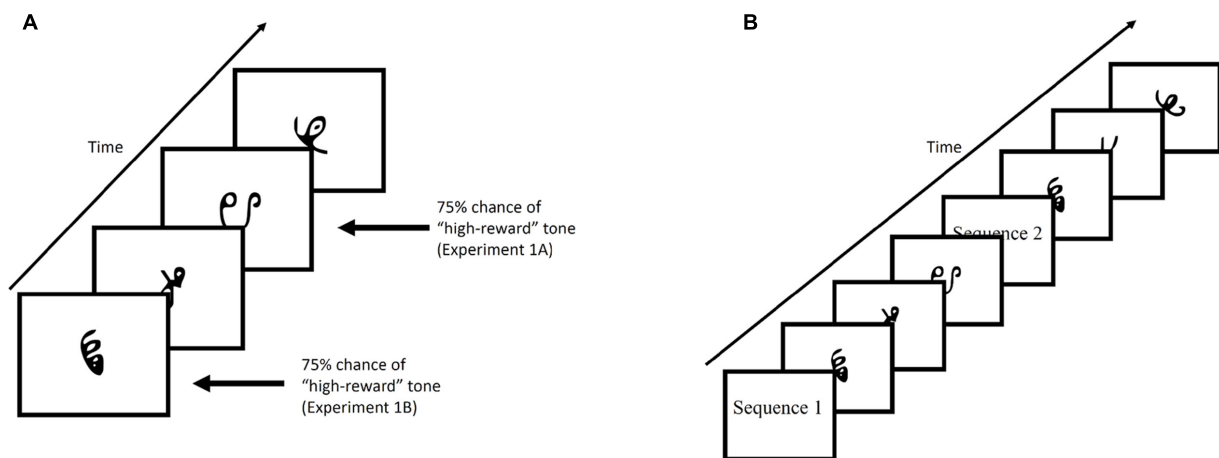


FIGURE 2 | (A) Example of the familiarization phase of Experiment 1. Stimuli were presented for 800 ms with a 200 ms blank ITI. **(B)** Example of the test phase. Participants were shown, sequentially, a characters comprised of a target triplet and characters comprised of a foil triplet. Each triplet was preceded by either "Sequence 1" or "Sequence 2". Participants were then required to select the triplet they viewed during the familiarization phase by pressing "1" or "2" on the keyboard.

Participants chose the sequence that appeared more familiar to them by pressing "1" or "2" on the keyboard in front of them.

Test phase trials employed one of the nine original triplets and one of nine foil triplets. Foil triplets were constructed using the same symbols exposed during training, recombined into new triplets such that each shape appeared in the same position within both the original and foil triplet (e.g., first, second, or third item in the triplet), but in novel combinations. For example, given the assignment shown in **Figure 1**, a foil triplet could contain the first character from Triplet 1, the second character from Triplet 2, and the third character from Triplet 3. Foil triplets were constructed exclusively from the same "value" triplets (i.e., we did not intermingle low-reward, high-reward, and neutral triplet constituents in foil triplets). Each trial in the test phase included one original triplet paired with one foil triplet. The order by which an original triplet or a foil triplet appeared was randomized, and participants were again required to choose

the triplet that they had observed during the familiarization phase. Experiments 1A and 1B each contained 162 test trials (each triplet was matched with each possible foil exactly twice) while Experiment 2 contained 54 test trials (each triplet was paired against each same-value triplet exactly once). In order to determine if visual statistical learning occurred and reward associations had an impact, proportion correct scores were calculated for each triplet value (e.g., low, high, and neutral), which were compared to chance performance.

EXPERIMENT 1A

Participants were instructed that they would be earning points during the familiarization phase, with the total number of points they earned converted into a cash reward at the conclusion of the experiment. While participants were viewing the stream of symbols, they were told to listen for an occasional "beep".

Every time they heard the high-pitched tone, they would gain 10 points toward their total. Alternatively, if they heard a low-pitched tone, they would not gain any points at all. Critically, and unbeknownst to the participants, these tones could only occur simultaneously with the third item in a triplet (i.e., reward was paired with the triplet a participant had just viewed) 75% of the time. Out of our nine original triplets, we associated a high-value reward with three triplets (+10 each time the high-pitch tone plays), a low-value reward with another three triplets (+0 each time the low-pitch tone plays, and no association with the remaining three triplets. With each triplet being presented 24 times and the high-pitched tone played on 75% of high-value triplet occurrences, each participant earned 540 points ($3 \times (0.75 \times 24) \times 10 = 540$). Points were converted into cents at the end of the experiment, and all participants won a total of \$5.40.

Participants were also given an attention-check task to ensure they were paying attention during the familiarization phase of Experiments 1A and 1B. Specifically, whenever a symbol was repeated, participants had to press the spacebar. The third item of each triplet was randomly selected to occasionally repeat throughout the familiarization phase. While reward tones may have occurred with the third item in the triplet, it never occurred with this fourth “repeat” item.

Should the high-pitched tone carrying the high-value association enhance visual statistical learning, then we may expect enhanced recognition of those triplets it had been paired with during the test phase. Alternatively, interference between reward learning and visual statistical learning may be evident if participants are more successful at selecting the low-value or value-absent triplets during the test phase. Finally, results may indicate that reward associations have no impact on visual statistical learning. In other words, despite the co-occurrence of these two powerful types of learning, participants may correctly identify triplets from the familiarization phase evenly across reward conditions.

Results

Figure 3A displays mean accuracies in selecting target triplets over foil triplets for Experiment 1A, as a function of both target and foil triplet value. Employing a 3×3 (triplet value \times foil value) repeated-measures ANOVA, we found no significant interaction of target triplet value and foil triplet value, $F(4,124) = 0.65$, $p = 0.63$, $\eta_p^2 = 0.02$, and no main effect of target value association, $F(2,62) = 0.36$, $p = 0.70$, $\eta_p^2 = 0.012$, or foil value association, $F(2,62) = 0.38$, $p = 0.69$, $\eta_p^2 = 0.012$. Visual statistical learning, however, was robust as measured by a one-sample t -test comparing performance to chance (50% correct), with participants correctly identifying more target triplets overall than foil triplets for high-value triplets, $t(31) = 3.08$, $p = 0.004$, Cohen's $d = 0.55$, low-value triplets, $t(31) = 2.25$, $p = 0.032$, $d = 0.39$, and value-absent triplets, $t(31) = 2.74$, $p = 0.01$, $d = 0.48$.

To examine strength of evidence favoring the null hypothesis, we applied a Bayesian repeated-measures ANOVA these data using the JASP software project (Love et al., 2015), with default

priors (Rouder et al., 2012). This analysis compares models that include versus do not include each factor and interaction, producing a Bayes Factor (BF) ratio that indicates the evidence in favor of the null model compared with evidence favoring a model that includes the factor or interaction in question. This analysis was used to produce a BF_{01} statistic for each main effect and the interaction. BF_{01} is an inverted Bayes Factor, with values greater than 1 indicating that the null model is favored, and with higher BF_{01} values indicating stronger evidence for a model that does not include the factor or interaction than one which does include the factor/interaction. In the base of both main effects and the interaction, evidence strongly favored the null model (target value main effect, $BF_{01} = 17.4$; foil value main effect, $BF_{01} = 16.9$; interaction, $BF_{01} = 294.4$), indicating strong evidence against the possibility that value meaningfully altered performance in the context of this experiment, either in terms of triplet or foil value.

EXPERIMENT 1B

While we had chosen to pair the reward tone with the third item in every triplet with the intention of establishing a retroactive association to the triplet, it could be the case that the reward association enhances visual statistical learning for *subsequent* characters in the stream. In this case, any effect of reward would be washed out across randomized triplet orderings. We examined this possibility in Experiment 1B by instead providing a reward association with the first item in some triplets rather than the third. Other than this change, all other aspects of Experiment 1B were identical to Experiment 1A.

Results

Figure 3 displays mean accuracies in selecting target triplets over foil triplets for Experiment 1B. 3×3 repeated measures ANOVA, we found no significant interaction of target triplet value and foil triplet value, $F(4,168) = 0.39$, $p = 0.82$, $\eta_p^2 = 0.009$. No significant main effect of target value association was observed, $F(2,84) = 0.57$, $p = 0.57$, $\eta_p^2 = 0.013$, and no significant main effect of foil value association was observed, $F(2,84) = 0.95$, $p = 0.39$, $\eta_p^2 = 0.022$. Visual statistical learning, was again robust with participants correctly identifying more target triplets overall than foil triplets, $t(42) = 4.97$, $p < 0.001$, $d = 0.76$. We again applied a Bayesian repeated-measures ANOVA to assess strength of evidence favoring the null hypothesis, BF_{01} . In the base of both main effects and the interaction, evidence continued to strongly favor the null model (target value main effect, $BF_{01} = 16.8$; foil value main effect, $BF_{01} = 9.7$; interaction, $BF_{01} = 169.4$).

EXPERIMENT 2

Experiments 1A and 1B demonstrated no evidence that visual statistical learning processes are influenced by ongoing reward signals paired consistently with constituent

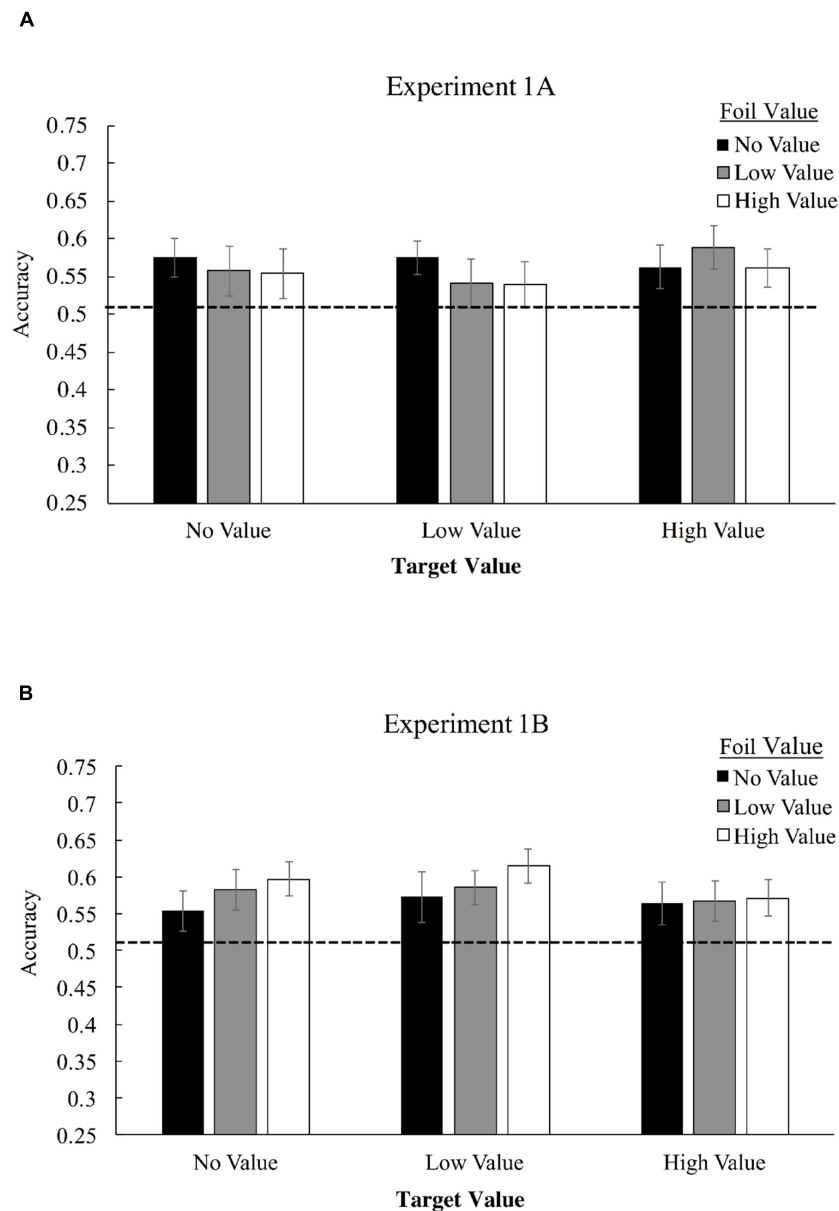


FIGURE 3 | Mean accuracies in identifying target triplets over foil triplets in Experiment 1A (A) and Experiment 1B (B). High-value triplets (+10 points) consists of shapes which were associated with the high-pitched tone, low value triplets (+0 points) were associated with a low-pitched tone, and value-absent triplets did not possess any association. During Familiarization, only one shape was consistently paired with the tone – the last shape in the triplet for Experiment 1A and the first shape for 1B, but all shapes in the triplet were considered no-, low-, or high-value for purposes of determining foil composition.

items, with no significant differences observed in identification accuracy according to value association. Experiment 2 was designed to explicitly introduce stimulus–reward learning prior to stimulus–stimulus associative learning, rather than including both on-going reward signals and stimulus–stimulus contingencies simultaneously. Participants were first required to learn the values of six specific symbols (half low-value, half high-value) at the start of the experiment. Instead of pairing a reward-tone with the symbols during the familiarization phase, participants in Experiment 2 were simply required to

commit value associations of specific symbols to memory before beginning the familiarization phase.

In Experiment 2, stimulus–reward learning was induced by first showing participants all six symbols alongside their corresponding value (e.g., +1 or +9). This initial presentation occurred twice. Participants were then shown all six symbols sequentially, in a random order, and were required to press the “1” or the “9” key on the keyboard to indicate its value. Shuffled presentation of all six symbols comprised a single block, and before moving on to the familiarization phase participants were

required to complete five consecutive blocks of the value-learning phase with 100% accuracy. At the beginning of this value-learning phase, participants were told that they would have to identify the value of a symbol at the conclusion of the experiment, and that they would be awarded the full value of that symbol if they are correct. Therefore, the value associations of these symbols was real, and participants' ability to memorize their value would dictate whether or not they could win \$1 or \$9 at the conclusion of the study.

An attention-check task was also implemented in Experiment 2. However, instead of having participants press a spacebar whenever an item repeated, they were required to press the spacebar whenever an item "jiggled" from left to right during the familiarization phase. Additionally, because participants had already learned the value associations, the presentation of a value-associated tone was obviated. Instead, the three high-value (+9) and low-value (+1) symbols were placed as the first, second, or third item in six of the nine triplets, with position manipulated between groups. The remaining three triplets did not possess a symbol with a learned value association (i.e., none of the three neutral triplets' shapes had been observed prior to the training phase).

During the test phase, the original triplet and the foil triplet were always matched by value. For example, a low-value original triplet was never paired with a high-value or a value-absent triplet. This logical restriction left us with a total of 54 test trials. Following the test phase, and congruent with what participants had been told in the initial value-learning phase, participants were required to recall the value of a symbol they had learned during the initial value-learning phase. One of the six value-associated symbols was presented on screen for people to explicitly recall the value of. If a participant was correct in recalling the value of this symbol, they were rewarded with a corresponding dollar amount of either \$1 or \$9. Accuracy at this stage was 100% – all subjects responded correctly, verifying success of the reward-training regimen.

Results and Discussion

Figure 4 displays mean accuracies in selecting target triplets over foil triplets for Experiment 2. With regard to where the value-associated symbols were placed, there was no difference in recall between value in the first, $F(2,42) = 0.66$, $p = 0.52$, $\eta_p^2 = 0.03$, second, $F(2,34) = 0.45$, $p = 0.64$, $\eta_p^2 = 0.026$, or third positions, $F(2,40) = 0.48$, $p = 0.62$, $\eta_p^2 = 0.02$. Nonetheless, participants continued to display visual statistical learning selecting learned triplets at above-chance levels, regardless of whether the learned symbol appeared in either the first, $t(21) = 4.04$, $p = 0.001$, $d = 0.86$, second, $t(17) = 4.10$, $p = 0.001$, $d = 0.97$ or third position $t(20) = 7.02$, $p < 0.001$, $d = 1.53$. Applying a Bayesian repeated-measures ANOVA for each value-associated placement again produced moderate evidence favoring the null model (first position, $BF_{01} = 4.2$; second position, $BF_{01} = 4.8$; and third position, $BF_{01} = 5.0$). Viewed as a mixed repeated-measures (value condition) and between-subjects (position) Bayesian ANOVA, the omnibus BF_{01} for the effect of target triplet value was 9.2.

GENERAL DISCUSSION

Based upon similar associative characteristics and impacts upon performance, we argued that stimulus–stimulus and stimulus–reward learning might interact. The shared neural basis of these two systems (Delgado et al., 2000; Aron, 2004; Wittmann et al., 2007; Lansink et al., 2009; Dickerson and Delgado, 2015) further bolstered our motivation to explore this possibility. Finally, the dependence of visual statistical learning on attention (Turk-Browne et al., 2005) in conjunction with the attention-modulating effects of reward (Engelmann et al., 2009; Hickey et al., 2010; Theeuwes and Belopolsky, 2012; Chelazzi et al., 2013; Sali et al., 2014; Pessoa, 2015), further suggest that reward learning may have the potential to enhance or disrupt statistical learning. However, results from the present study were unable to identify any influence of reward learning on visual statistical learning.

Despite reliable evidence that visual statistical learning successfully occurred throughout our experiments, we failed to find a reliable influence of reward. This may be most surprising in Experiment 1, where presentation of a sound stimulus probabilistically paired with either the first or third item provided additional subtle information about the presence of structure, as well as serving as a reward cue. Even in this case, where an additional cue to structure was present in value-associated triplets but absent from no-value triplets, participants performed comparably in identifying no-, low-, and high-value reward triplets. Thus, in an environment where reward learning and visual statistical learning are concurrently active, our results suggest that visual statistical learning is unaffected by the occurrence of rewarding events.

While visual statistical learning appears to be unaffected by concurrently active learning mechanisms, it also appears to be unaffected by previously learned reward. In Experiment 2, participants first learn to strongly associate values with symbols before engaging in the familiarization phase. Similar to our findings from the first experiment, visual statistical learning was generally evident across all conditions, but there appeared to be no clear effect of the value association upon learning, nor was there any effect of pre-exposure of some constituent shapes from the reward-learning phase of the experiment.

These results suggest that concurrently presented rewards and previously learned stimulus–reward associations have no impact on visual statistical learning. That is, whether stimulus–reward associations were introduced at the same time as stimulus–stimulus associations, or whether stimulus–reward associations were established before stimulus–stimulus associations were learned, participants' ability to accurately identify familiar structured triplets of symbols remained unaffected. However, it is important to acknowledge the potential shortcomings of the present work.

It is possible that our attempt to create strong stimulus–reward contingencies was not powerful enough to influence visual statistical learning. In other words, it is feasible that one could have participants engage in a more rigorous reward-learning tasks before or during visual statistical learning. Some evidence suggests that the impact of reward on attention

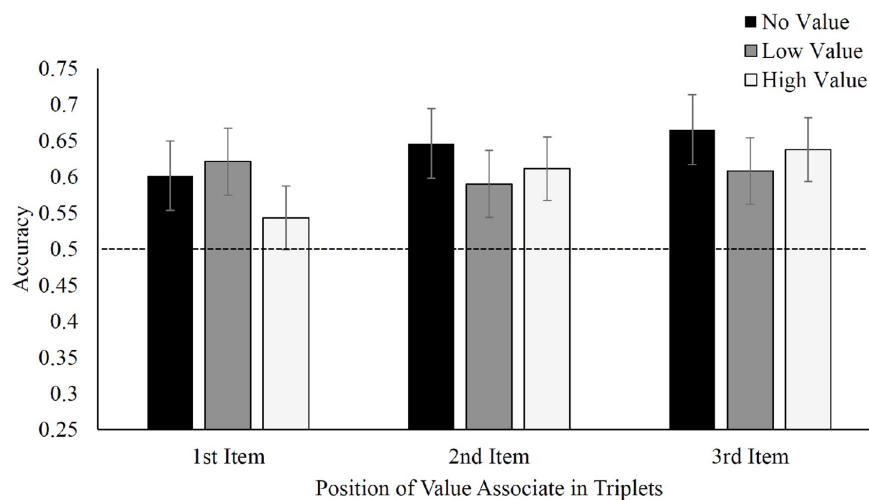


FIGURE 4 | Mean accuracy in identifying target triplets over foil triplets in Experiment 2, separated by position of value associate within the triplet.

scales as rewards increase (Anderson et al., 2013). Given our understanding that visual statistical learning is dependent upon attention (Turk-Browne et al., 2005), one could argue that larger rewards may have produced a larger effect of reward learning by amplifying attention during high-reward experiences. However, potent and well-documented effects of reward on various cognitive processes have been demonstrated using similar reward values and reward delivery strategies (Cohen et al., 2007; Hickey et al., 2010; Chelazzi et al., 2013; Marx and Einhauser, 2015); thus, it seems unlikely that an interaction between reward learning and visual statistical learning was missed due to our choice of reward quantity, which is within the reasonable range of incentives provided in efforts that demonstrate clear effects of rewards.

In terms of reward's effects on attention, one possible explanation for our null result is that attention actually has limited effect on visual statistical learning. Although earlier work suggested that statistical learning is "gated" by selective attention (Turk-Browne et al., 2005), recent work has challenged the robustness of this finding (Musz et al., 2015). Thus, it is possible that statistical learning is non-existent, context-dependent, or immune to variations in selective attention, or at least those variations likely evoked by the kinds of cues to reward used in our studies.

Regardless of the degree of attentional variation induced by the value manipulations used here, and of the role of attention in visual statistical learning, the well-established, strong role of value associations in driving variations in performance (possibly without necessitating a role of attention as a mediating variable) suggests potential for value associations to influence statistical learning. Two important considerations are the possibility of distinctions between primary and secondary reinforcers, and between transient- and state-based effects. Regarding the former, here we only tested secondary reinforcements (signals to monetary value). Primary rewards might be more powerful and effective at inducing changes to visual statistical learning.

Regarding the latter, we have demonstrated cases in which reward-related stimuli selectively paired with constituent visual images does not impact visual statistical learning, but the random interleaving of low- and high-reward events leaves open the possibility that reward may influence statistical learning in a manner that depends upon cognitive or emotional state. An example of a well-known effect of emotional states on learning and memory are contextual effects of mood, in which better recall is experienced in a mood state congruent to encoding (Bower, 1981). Effects of reward on associative learning between stimuli could be state-based, require longer and more powerful periods of induction by repeated or otherwise more potent reward, and/or such effects may act over longer learning timescales.

The present works provides insight into the integrity of visual statistical learning, with evidence suggesting isolation from the effects of rewarding events. Additionally, while these two systems share similar neural correlates, the functional role of these neural structures in each type of learning may differ. However, we did not test the opposite relationship here – that statistical learning may impact reward learning, even if rewarding associations do not impact statistical learning. Statistical learning may precede reward learning and influence inferences about reward value (e.g., by supporting transitive inference of reward from one item to its statistical associates), or facilitate or impair reward learning in other ways. Further work is needed to explore these possibilities.

CONCLUSION

Interjecting events that vary in reward significance appears to leave visual statistical learning unchanged. While this finding depends upon a lack of statistical significance, we saw no clear trend of any effect, implying limits to any such interference or facilitation. In environments that feature both statistical regularity amongst stimuli, as well as contingencies between those stimuli and rewarding events or history, our

evidence suggests that statistical learning is unaffected and possibly independent of reward.

AUTHOR CONTRIBUTIONS

KF and TV conceived of, implemented, and collected data for the experiments. LR, KF, and TV analyzed and interpreted data. LR wrote the manuscript. LR, KF, and TV revised the manuscript.

REFERENCES

- Anderson, B. A., Laurent, P. A., and Yantis, S. (2011). Value-driven attentional capture. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10367–10371. doi: 10.1073/pnas.1104047108
- Anderson, B. A., Laurent, P. A., and Yantis, S. (2013). Reward predictions bias attentional selection. *Front Hum. Neurosci.* 7:262. doi: 10.3389/fnhum.2013.00262
- Aron, A. R. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J. Neurophysiol.* 92, 1144–1152. doi: 10.1152/jn.01209.2003
- Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology* 191, 391–431. doi: 10.1007/s00213-006-0578-x
- Bower, G. H. (1981). Mood and memory. *Am. Psychol.* 36, 129–148. doi: 10.1037/0003-066X.36.2.129
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Chelazzi, L., Perlato, A., Santandrea, E., and Della Libera, C. (2013). Rewards teach visual selective attention. *Vis. Res.* 85, 58–62. doi: 10.1016/j.visres.2012.12.005
- Chun, M. M., and Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognit. Psychol.* 36, 28–71. doi: 10.1006/cogp.1998.0681
- Cohen, J. D., McClure, S. M., and Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 933–942. doi: 10.1098/rstb.2007.2098
- Daw, N. D., and Doya, K. (2006). The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 16, 199–204. doi: 10.1016/j.conb.2006.03.006
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., and Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072–3077.
- Dickerson, K. C., and Delgado, M. R. (2015). Contributions of the hippocampus to feedback learning. *Cogn. Affect. Behav. Neurosci.* 15, 861–877. doi: 10.3758/s13415-015-0364-5
- Engelmann, J. B., Damaraju, E., Padmala, S., and Pessoa, L. (2009). Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Front. Hum. Neurosci.* 3:4. doi: 10.3389/fnhum.2009.004.2009
- Fiser, J., and Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* 12, 499–504. doi: 10.1111/1467-9280.00392
- Fiser, J., and Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol.* 28, 458–467. doi: 10.1037/0278-7393.28.3.458
- Glicksohn, A., and Cohen, A. (2011). The role of Gestalt grouping principles in visual statistical learning. *Atten. Percept. Psychophys.* 73, 708–713. doi: 10.3758/s13414-010-0084-4
- Glicksohn, A., and Cohen, A. (2013). The role of cross-modal associations in statistical learning. *Psychon. Bull. Rev.* 20, 1161–1169. doi: 10.3758/s13423-013-0458-4
- Haber, S. N., and Knutson, B. (2009). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26. doi: 10.1038/npp.2009.129

FUNDING

This work was supported by NSF grant 1558535 to TV and a University of Delaware Research Foundation grant to TV.

ACKNOWLEDGMENT

We thank Nicholas Angelides, Jared Beneroff, and Sarah Sweigart for assistance with data collection.

- Hickey, C., Chelazzi, L., and Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *J. Neurosci.* 30, 11096–11103. doi: 10.1523/JNEUROSCI.1026-10.2010
- Jiang, Y. V., Sigstad, H. M., and Swallow, K. M. (2013). The time course of attentional deployment in contextual cueing. *Psychon. Bull. Rev.* 20, 282–288. doi: 10.3758/s13423-012-0338-3
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception* 36:1.
- Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., and Pennartz, C. M. A. (2009). Hippocampus leads ventral striatum in replay of place-reward information. *PLoS Biol.* 7:e1000173. doi: 10.1371/journal.pbio.1000173
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., et al. (2015). *JASP (Version 0.7)* [Computer software]. Amsterdam: JASP Project.
- Marx, S., and Einhauser, W. (2015). Reward modulates perception in binocular rivalry. *J. Vis.* 15, 11–11. doi: 10.1167/15.1.11
- Musz, E., Weber, M. J., and Thompson-Schill, S. L. (2015). Visual statistical learning is not reliably modulated by selective attention to isolated events. *Atten. Percept. Psychophys.* 77, 78–96. doi: 10.3758/s13414-014-0757-5
- Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017
- Pessoa, L. (2015). Multiple influences of reward on perception and attention. *Vis. Cogn.* 23, 272–290. doi: 10.1080/13506285.2014.974729
- Raymond, J. E., and O'Brien, J. L. (2009). Selective visual attention and motivation: the consequences of value learning in an attentional blink task. *Psychol. Sci.* 20, 981–988. doi: 10.1111/j.1467-9280.2009.02391.x
- Rouder, J. N., Morey, R. D., Speckman, P. L., and Province, J. M. (2012). Default Bayes factors for ANOVA designs. *J. Math. Psychol.* 56, 356–374. doi: 10.1016/j.jmp.2012.08.001
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926
- Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52. doi: 10.1016/S0010-0277(98)00075-4
- Sali, A. W., Anderson, B. A., and Yantis, S. (2014). The role of reward prediction in the control of attention. *J. Exp. Psychol.* 40, 1654–1664. doi: 10.1037/a0037267
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Sha, L. Z., and Jiang, Y. V. (2015). Components of reward-driven attentional capture. *Atten. Percept. Psychophys.* 78, 403–414. doi: 10.3758/s13414-015-1038-7
- Theeuwes, J., and Belopolsky, A. V. (2012). Reward grabs the eye: oculomotor capture by rewarding stimuli. *Vis. Res.* 74, 80–85. doi: 10.1016/j.visres.2012.07.024
- Turk-Browne, N. B., Jungé, J., and Scholl, B. J. (2005). The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* 134, 552–564. doi: 10.1037/0096-3445.134.4.552
- Turk-Browne, N. B., and Scholl, B. J. (2009). Flexible visual statistical learning: transfer across space and time. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 195–202. doi: 10.1037/a0015233
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., and Johnson, M. K. (2009). Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J. Cogn. Neurosci.* 21, 1934–1945. doi: 10.1162/jocn.2009.21131
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., and Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning.

- J. Neurosci.* 30, 11177–11187. doi: 10.1523/JNEUROSCI.0858-10.2010
- Wittmann, B. C., Bunzeck, N., Dolan, R. J., and Düzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage* 38, 194–202. doi: 10.1016/j.neuroimage.2007.06.038
- Yu, R. Q., and Zhao, J. (2015). The persistence of the attentional bias to regularities in a changing environment. *Atten. Percept. Psychophys.* 77, 2217–2228. doi: 10.3758/s13414-015-0930-5
- Zhao, J., Al-Aidroos, N., and Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychol. Sci.* 24, 667–677. doi: 10.1177/0956797612460407

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Rogers, Friedman and Vickery. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Episodic Short-Term Recognition Requires Encoding into Visual Working Memory: Evidence from Probe Recognition after Letter Report

Christian H. Poth* and Werner X. Schneider

Neuro-Cognitive Psychology, Department of Psychology and Cluster of Excellence Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Louise A. Brown,
University of Strathclyde, UK
Thomas Alrik Sørensen,
Aalborg University, Denmark

*Correspondence:

Christian H. Poth
c.poth@uni-bielefeld.de

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 06 June 2016

Accepted: 08 September 2016

Published: 22 September 2016

Citation:

Poth CH and Schneider WX (2016)
Episodic Short-Term Recognition
Requires Encoding into Visual
Working Memory: Evidence from
Probe Recognition after Letter Report.
Front. Psychol. 7:1440.
doi: 10.3389/fpsyg.2016.01440

Human vision is organized in discrete processing episodes (e.g., eye fixations or task-steps). Object information must be transmitted across episodes to enable episodic short-term recognition: recognizing whether a current object has been seen in a previous episode. We ask whether episodic short-term recognition presupposes that objects have been encoded into capacity-limited visual working memory (VWM), which retains visual information for report. Alternatively, it could rely on the activation of visual features or categories that occurs before encoding into VWM. We assessed the dependence of episodic short-term recognition on VWM by a new paradigm combining letter report and probe recognition. Participants viewed displays of 10 letters and reported as many as possible after a retention interval (whole report). Next, participants viewed a probe letter and indicated whether it had been one of the 10 letters (probe recognition). In Experiment 1, probe recognition was more accurate for letters that had been encoded into VWM (reported letters) compared with non-encoded letters (non-reported letters). Interestingly, those letters that participants reported in their whole report had been near to one another within the letter displays. This suggests that the encoding into VWM proceeded in a spatially clustered manner. In Experiment 2, participants reported only one of 10 letters (partial report) and probes either referred to this letter, to letters that had been near to it, or far from it. Probe recognition was more accurate for near than for far letters, although none of these letters had to be reported. These findings indicate that episodic short-term recognition is constrained to a small number of simultaneously presented objects that have been encoded into VWM.

Keywords: visual working memory, visual attention, episodic memory, object recognition, short-term memory

INTRODUCTION

Visual information processing is organized in discrete episodes. This is most evident from the fact that the uptake of visual information is largely limited to eye fixations, discrete periods of stable eye position that are interrupted by fast saccadic eye movements (e.g., Krock and Moore, 2015). However, on a greater time scale, processing episodes can also be defined by steps of sensorimotor

actions, other task-demands, and changes in the visual environment (Petersen et al., 2012; Duncan, 2013; Schneider, 2013; Herwig, 2015; Poth et al., 2015; Poth and Schneider, 2016). To remain oriented in time and space and to act guided by vision, visual information from consecutive processing episodes must be linked. This is particularly evident from tasks requiring to recognize that objects (or subjects) have been viewed recently (e.g., Sternberg, 1966; Wickelgren, 1970; Kahana and Sekuler, 2002; Zhou et al., 2004; Donkin and Nosofsky, 2012). For example, imagine you are standing at a busy inner-city intersection and someone shows you a picture of a dog that just went missing and asks if you have seen it. To answer this question, you must be able to recognize if the dog appeared in one of the many recent processing episodes that consisted of your eye fixations, steps of your actions, and periods of cars passing by. Such tasks require *episodic short-term recognition*: the cognitive function of recognizing whether a now-present object has been contained in a recently passed visual processing episode¹ (cf. Kahana and Sekuler, 2002; Zhou et al., 2004; Donkin and Nosofsky, 2012).

How is episodic short-term recognition accomplished? What are its underlying mechanisms? First of all, to recognize that an object has been present before, the object must be represented internally. Several views on visual processing posit that initially, objects are represented by activating their corresponding feature or category representations in visual long-term memory (Cowan, 1988; Bundesen, 1990; Henderson, 1994; Henderson and Anes, 1994; Eriksson et al., 2015; cf. Oberauer, 2002; LaRocque et al., 2014; for a more general overview, see Palmeri and Tarr, 2008). These representations code for visual features and categories of objects that have been acquired through past visual experience and are often called *visual types* (e.g., Kanwisher, 1987; Kahneman et al., 1992; although other terms are in use as well, e.g., Duncan and Humphreys, 1989; Bundesen, 1990). Visual types represent objects in a multidimensional feature and category space and they may also represent exemplars of certain objects (cf. Kahana and Sekuler, 2002; Nosofsky et al., 2011; Donkin and Nosofsky, 2012).

Critically, activating an object's visual type (feature, category) is only considered an initial step of processing (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Kyllingsbæk, 2014). This activation does neither suffice to act upon the object nor to consciously perceive the object in the sense that it can be reported. Importantly, the activation is "pre-attentive" in the sense of being unselective: it proceeds likewise for all objects in the visual field (or parts of the visual field, depending on pre-existing spatial biases, Bundesen and Habekost, 2008, p. 117, and retinal inhomogeneity, Strasburger et al., 2011). That is, it proceeds before mechanisms of visual attention select task-relevant objects for further processing at the expense of task-irrelevant ones (e.g., Duncan and Humphreys, 1989; Bundesen,

1990; Bundesen et al., 2005; Duncan, 2006; Poth et al., 2014). For action and report, objects must be attended, processed further, and eventually encoded into visual working memory (VWM; Duncan and Humphreys, 1989; Bundesen, 1990; Cowan, 2001; Bundesen et al., 2005; Schneider, 2013; note that we use VWM synonymously to the also common term of visual short-term memory).

Visual working memory consists of a mechanism for retaining visual object representations accessible over short time-windows (for reviews, see Luck, 2008; Bundesen et al., 2011; Luck and Vogel, 2013; LaRocque et al., 2014; Ma et al., 2014). In this way, VWM may provide an essential basis for further processing these representations, as recoding them into other representational formats (e.g., the verbal format) so that they can be retained and used by non-visual mechanisms of working memory (e.g., Logie, 2011). The capacity of VWM is limited so that it can only hold about three to four objects (e.g., Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997; Dyrholm et al., 2011; Poth et al., 2014; note that capacity is also limited in the number of object features, Wheeler and Treisman, 2002; Oberauer and Eichenberger, 2013, and the precision of object features, Wilken and Ma, 2004; Bays and Husain, 2008). Which of all available objects are encoded into VWM depends on selection by visual attention (e.g., Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Duncan, 2006; Poth et al., 2014). Because of the limited capacity of VWM, all visually available objects may initially and (pre-attentively) activate visual types in visual long-term memory, but only a limited number of objects is (attentively) processed up to the level of VWM (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005). Encoding objects into VWM is a core requirement of visually controlled behavior, because objects can only be reported and used for action when they are represented in VWM (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005). This paper focuses on the open question of whether encoding into VWM is also necessary for episodic short-term recognition.

Episodic short-term recognition requires comparisons of object representations of a recently preceding processing episode with representations of objects of the current episode. This can be conceptualized as a decision process (e.g., Pearson et al., 2014) which is driven by the degree of similarity between these two kinds of representations (e.g., Ratcliff, 1978; Donkin and Nosofsky, 2012; cf. Kahana and Sekuler, 2002). Two rival hypotheses can be advanced regarding the role of VWM in this comparison process (based on the literature covered above). According to the *VWM-encoding* hypothesis, episodic short-term recognition of an object from a previous episode requires that the object has been encoded into VWM. Consequently, objects that have not been processed up to the level of VWM cannot be used for episodic short-term recognition. Alternatively, the *type-activation* hypothesis states that episodic short-term recognition is also possible for objects which have not been encoded into VWM but whose mere presentation has activated their visual types in visual long-term memory. This means that episodic short-term recognition is possible for all external objects that have been visually available within recent eye fixations. In such a case, activations of visual types could extend into the

¹Note that the term episodic short-term recognition refers to the described cognitive function (in the sense of a cognitive task-requirement). In this way, the concept of episodic short-term recognition does not include any assumptions about the cognitive mechanisms enabling to fulfill this function (such as for example, interacting mechanisms of episodic long-term memory or working memory).

next processing episode. These remaining activations could be matched against activations elicited by objects of this episode. A resulting signal could then allow the comparison of object representations from the previous episode and from the actual environment underlying episodic short-term recognition (e.g., Ratcliff, 1978; Donkin and Nosofsky, 2012). Such a mechanism could be similar to mechanisms assumed to produce attention-independent priming effects, where the presentation of objects facilitates their subsequent object recognition (e.g., Kahneman et al., 1992; Henderson, 1994; Henderson and Anes, 1994; Jensen and Lisman, 1998) or affects motor responses to other stimuli (even if the objects are not discriminable, Klotz and Neumann, 1999, and hence not in VWM, Bundesen, 1990).

Here, we aimed at deciding between the two hypotheses. In two experiments, we asked whether episodic short-term recognition of an object requires that this object has previously been encoded into capacity-limited VWM. To approach this question, we introduced a new paradigm combining letter report with probe recognition.

EXPERIMENT 1

In Experiment 1, participants performed a whole report task (e.g., Sperling, 1960; Shibuya and Bundesen, 1988) which was combined with a probe recognition task. They briefly viewed displays of to-be-memorized letters (memory letters) and then, after a retention interval, reported as many letters as they could. The retention interval outlasted early sensory memory (e.g., Sperling, 1960; Phillips, 1974; Irwin and Thomas, 2008) so that letter reports should have required retention in VWM (followed by a recoding into a verbal format on which the actual report was based, e.g., Logie, 2011; Baddeley, 2012). Memory letters were always 10 different ones, exceeding VWM capacity and thus ensuring participants could never report all letters (Sperling, 1960; Shibuya and Bundesen, 1988). After reporting the letters, a single probe letter appeared within the same trial and participants indicated whether or not the probe had been shown as one of the previous memory letters. Importantly, the probe was either one of the memory letters and reported (*reported condition*), or one of the memory letters but not reported (*non-reported condition*), or it was a letter not contained in the set of memory letters (*not shown condition*).

Here, episodic short-term recognition was assessed as performance in probe recognition, that is, in indicating whether or not the probe letter had been shown as one of the memory letters. Which memory letters were encoded into VWM was assessed by preceding letter reports. Since VWM is defined by the accessibility of its content (e.g., Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013; but see, Soto et al., 2011), reported letters must have been in VWM by definition. Following a number of theories (e.g., Bundesen, 1990; Bundesen et al., 2005; Martens and Wyble, 2010; Schneider, 2013), we assume that letters which were not reported did not enter VWM. Consequently, the VWM-encoding hypothesis predicts higher probe recognition performance in the reported than in the non-reported and not shown conditions. In contrast, no such performance differences

are expected based on the type-activation hypothesis. According to this hypothesis, performance should be equal in the reported and non-reported conditions. More specifically, episodic short-term recognition should be possible for all presented memory letters, irrespective of their encoding into VWM. That is because all presented memory letters should have activated their visual types in visual long-term memory as part of the initial processing of the letters (e.g., Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Kyllingsbæk, 2014; see above). Besides testing these hypotheses, Experiment 1 explored whether memory letters in the whole report task were encoded in a spatially clustered manner. That is, whether letters in close spatial proximity were encoded with preference over letters that were farther apart. Such a spatial clustering may reveal attentional selection strategies and this will become important in Experiment 2.

Method

Participants

Fourteen participants were paid to take part in the experiment. They were between 18 and 30 years old ($Mdn = 20$ years), nine were male, five female, 13 were right-handed and one left-handed, and all reported normal or corrected-to-normal visual acuity and color vision. All participants gave written informed consent before performing the experiments that were conducted according with the ethical standards of the German Psychological Association (Deutsche Gesellschaft für Psychologie, DGPs), and were approved by Bielefeld University's ethics committee. One additional participant was excluded from data analysis because of an experimentation error.

Apparatus and Stimuli

The experiment took place in a dimly lit room. Stimuli were presented on a 19" CRT-screen (Trinitron MultiScan G420, Sony, Park Ridge, NJ, using a graphics card of type Quadro NVS 290, NVIDIA, Santa Clara, CA, USA) with a refresh rate of 85 Hz and a resolution of 1280×1024 pixels at physical dimensions of $36 \text{ cm} \times 27 \text{ cm}$. The participant's head was stabilized by a chin rest positioned 71.8 cm from the screen. Responses were collected using a standard computer keyboard with German layout. Labels indicating "yes" (by the German word "Ja") and "no" (by the German word "Nein") were placed above the F1 and F9 keys of the keyboard. The experiment was controlled by the Psychophysics Toolbox 3.0.12 extension (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MATLAB R2013b (The MathWorks, Natick, MA, USA).

A MAVOLUX-digital luminance meter (Gossen, Nuremberg, Germany) was used to measure stimulus luminance. Black letter stimuli (0.32° of visual angle $\times 0.48^\circ$; $< 1 \text{ cd} \times \text{m}^{-2}$) from the set [ABDEFGHJKLMNOPRSTVXZ] (this set of letters was chosen to avoid highly confusable letters, as e.g., by Poth et al., 2015) were located equally spaced on an imaginary circle with a radius of 2° around screen center. Fixation cross ($0.32^\circ \times 0.32^\circ$) and response screen text were white ($108 \text{ cd} \times \text{m}^{-2}$). The response screen showed the German text "Buchstaben?", which

means “Letters?” in English. Stimuli were shown against a gray background ($21 \text{ cd} \times \text{m}^{-2}$).

Procedure and Design

Before the experiment, participants read instructions on the screen and reported them to the experimenter in their own words. The experimenter repeated the instructions again, if participants had reported them incorrectly. **Figure 1** illustrates the experimental paradigm. Participants initiated each trial by pressing the space-bar. In the beginning of a trial, a fixation cross was shown for 400 ms. Next, 10 memory letters were presented for 200 ms. The letters were randomly drawn without replacement from the set of used letters. The memory letters were followed by a blank interstimulus interval (ISI) lasting for 1000 ms (this duration ensures that early sensory (iconic) memory representations of the letters have been decayed, e.g., Sperling, 1960; Phillips, 1974; Irwin and Thomas, 2008), after which a response screen prompted participants to enter letters. Participants should report as many from the preceding memory letters as they could (without being required to report as many as 10 letters). A maximum of 10 letters could be entered (but this never happened). After confirming that they had finished reporting letters by pressing the enter-key, another ISI of 94 ms followed. Then a single probe letter was presented. Participants indicated whether or not this probe was one of the preceding memory letters by pressing the F1 or F9 key, respectively.

The probe was manipulated in three conditions of a within-subjects design. In the reported condition, the probe was randomly chosen from the letters which were shown and reported by the participant on this trial. In the non-reported condition, the probe was one of the letters that were shown on this trial but that the participant did not report. In both of these two conditions, probes appeared at their locations in the display of the memory letters. In the not shown condition, the probe was randomly chosen from the set of all letters excluding the memory letters of the trial (irrespective of whether participants had entered these letters). In this condition, the probe appeared at a random location.

Participants performed three blocks of 100 trials, each comprising 25 trials of the reported, 25 trials of the non-reported, and 50 trials of the not shown condition. Twice as many trials of the not shown as of the other two conditions were included to equate the number of trials in which a previously shown (correct answer “yes”) or a not shown letter (correct answer “no”) was probed. Within each block, trials of the three conditions were administered in random order. Participants performed twelve training trials prior to the experiment.

Results and Discussion

A significance criterion of $p < 0.05$ was used for all statistical analyses. Performance in the three conditions was compared using one-way repeated-measures analyses of variance with type II sums-of-squares for which η^2_G (Bakeman, 2005) is reported as effect size. Where the assumption of sphericity was violated, p -values are based on Greenhouse–Geisser-corrected degrees of freedom and the correction factor ϵ is reported alongside

the uncorrected degrees of freedom. Paired t -tests (two-tailed) with Bonferroni-corrected p -values (p_B) were used for pairwise comparisons for which d_z (Cohen, 1988) is reported as effect size. These t -tests were supplemented with corresponding Bayes factors (BF ; Rouder et al., 2009), of which values greater one favor the null hypothesis and values smaller one favor the alternative hypothesis. All analyses were performed using R (3.0.3; R Development Core Team, 2016).

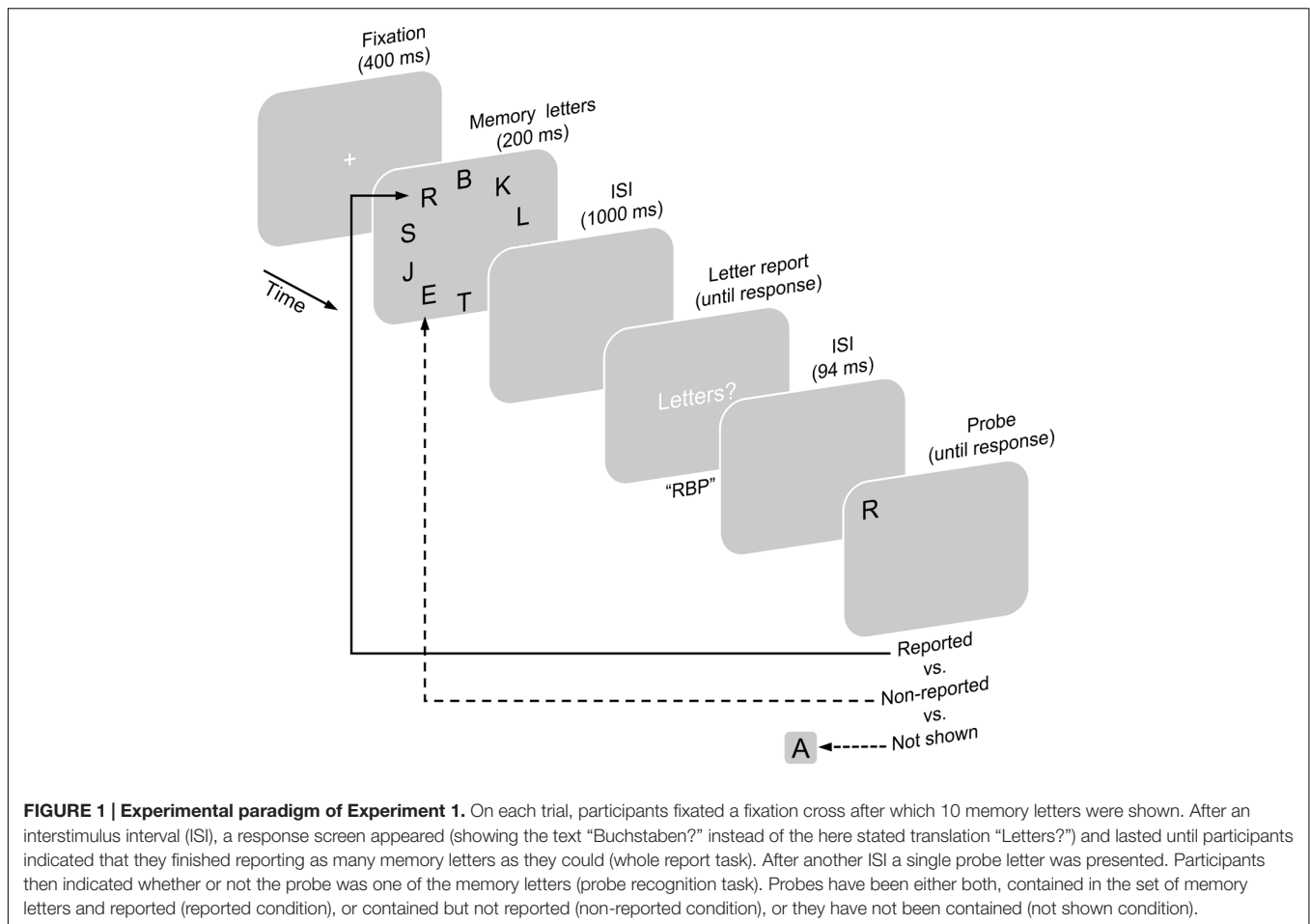
A total of 3.3% of all trials were discarded before analysis because either, (1) none of the memory letters was reported (0.57%), or (2) duplicate letters were contained in the letter report (2.76%).

Letter Report Performance

Letter report performance was assessed as participants’ mean number of correctly reported letters, that is, for each individual participant the mean number of typed-in letters matching one of the memory letters across trials. There were no significant differences regarding letter report performance in the three conditions, $F(2,26) = 2.231$, $p = 0.128$, $\eta^2_G = 0.002$. In addition, mean letter report performance was in the range of three to four letters in all three conditions (reported: $M = 3.62$, $SD = 0.59$, $\min = 2.41$, $\max = 4.60$; non-reported: $M = 3.56$, $SD = 0.61$, $\min = 2.35$, $\max = 4.41$; not shown: $M = 3.56$, $SD = 0.59$, $\min = 2.44$, $\max = 4.5$), consistent with previous estimates of VWM capacity in letter report tasks (Sperling, 1960; Shibuya and Bundesen, 1988).

Spatial Clustering of Reported Letters

Whether letters were encoded into VWM in a spatially clustered manner was assessed as follows. For each trial, the extent to which reported letters were spatially clustered within the original display of memory letters (i.e., their spatial proximity in this display) was quantified. The data was collapsed across conditions, since trials in the three conditions did not differ until after letters had been reported. Each correctly reported letter was selected for one step of the analysis. For this selected letter, it was determined whether or not the memory letters at the 10 positions relative to it were correctly reported (**Figure 2A**). This must be always the case for relative position zero, as this is the position of the selected letter itself. The procedure resulted in a matrix with the dimensions number of reported letters (rows) \times 10 letter positions (columns) and with entries coding for whether or not a given letter has been reported. Now, spatial clustering of letter reports was assessed as the proportions of reported letters for each letter position (i.e., for each column) across all reported letters (i.e., across all rows). If participants reported letters in a spatially random manner, then these proportions should be equal with the exception of a proportion of 1 for the selected letters (see **Figure 2B** for a computer simulation). In contrast, spatial clustering in encoding letters would become manifest in higher proportions for letters at positions more proximal compared with positions more distant to the selected letter (**Figure 2C** for a computer simulation). Note that these analyses require that the number of presented letters clearly exceeds participants’ VWM capacity because otherwise there would be no clear differences between proportions. This condition is assumed to be met because participants reported



between three and four of the 10 presented letters (see the letter report performance above).

As can be seen in **Figure 2D**, the mean proportions of reported letters monotonically decreased with increasing distance to selected letters and this pattern was present in all participants. Page's trend test was used to test whether monotonic decreases from closer to more distant positions were statistically significant. To this end, Page's trend test was applied to the participants' proportions at relative positions -1 to -4 and, separately, at relative positions 1 to 5 (**Figure 2A**). Results revealed monotonic decreases for both of these subsets of the data, locations -1 to -4 : $L = 420$, $p < 0.001$, locations 1 to 5 : $L = 768$, $p < 0.001$ (and these monotonic decreases were present in all of the three blocks of trials, all $Ls \geq 420$, all $ps < 0.001$).

Selective encoding of letters into VWM was not spatially random. Instead, all participants encoded subsets of the memory letters into VWM that were in close spatial proximity in the letter display. This spatial clustering may reflect an attentional encoding strategy. Participants learned over trials that always more memory letters were shown than they could report. Thus, participants learned they had to select subsets of the memory letters for report. Spatial clustering may be a means to accomplish such a selection from equally task-relevant objects by restricting encoding to objects in close spatial proximity. In this way, spatial

clustering may reflect the distribution of spatial attention (e.g., Posner, 1980; Bundesen, 1990), which in this specific case selects objects at or close to a strategically and internally specified location.

Probe Recognition Performance

Probe recognition performance was assessed as the proportion of trials on which probe letters were correctly recognized as having been shown or not shown on the trial. **Figure 3** depicts the participants' probe recognition performance, both at the sample and individual level. Probe recognition performance differed significantly between the three conditions, $F(2,26) = 44.912$, $\epsilon = 0.522$, $p < 0.001$, $\eta_G^2 = 0.771$. Probe recognition performance was significantly higher in the reported ($M = 0.96$, $SD = 0.03$) compared with the non-reported ($M = 0.29$, $SD = 0.19$), $t(13) = 12.774$, $p_B < 0.001$, $d_z = 3.41$, $BF = 8.8 \times 10^{-7}$, and the not shown condition ($M = 0.74$, $SD = 0.20$), $t(13) = 4.170$, $p_B = 0.003$, $d_z = 1.11$, $BF = 0.028$. Moreover, performance was significantly lower in the non-reported than in the not shown condition, $t(13) = -4.498$, $p_B = 0.002$, $d_z = -1.20$, $BF = 0.016$. One-sample t -tests (two-sided) revealed that performance was significantly below the chance level of 0.5 in the non-reported condition, $t(13) = -4.243$, $p < 0.001$, $BF = 0.025$, whereas it was significantly above

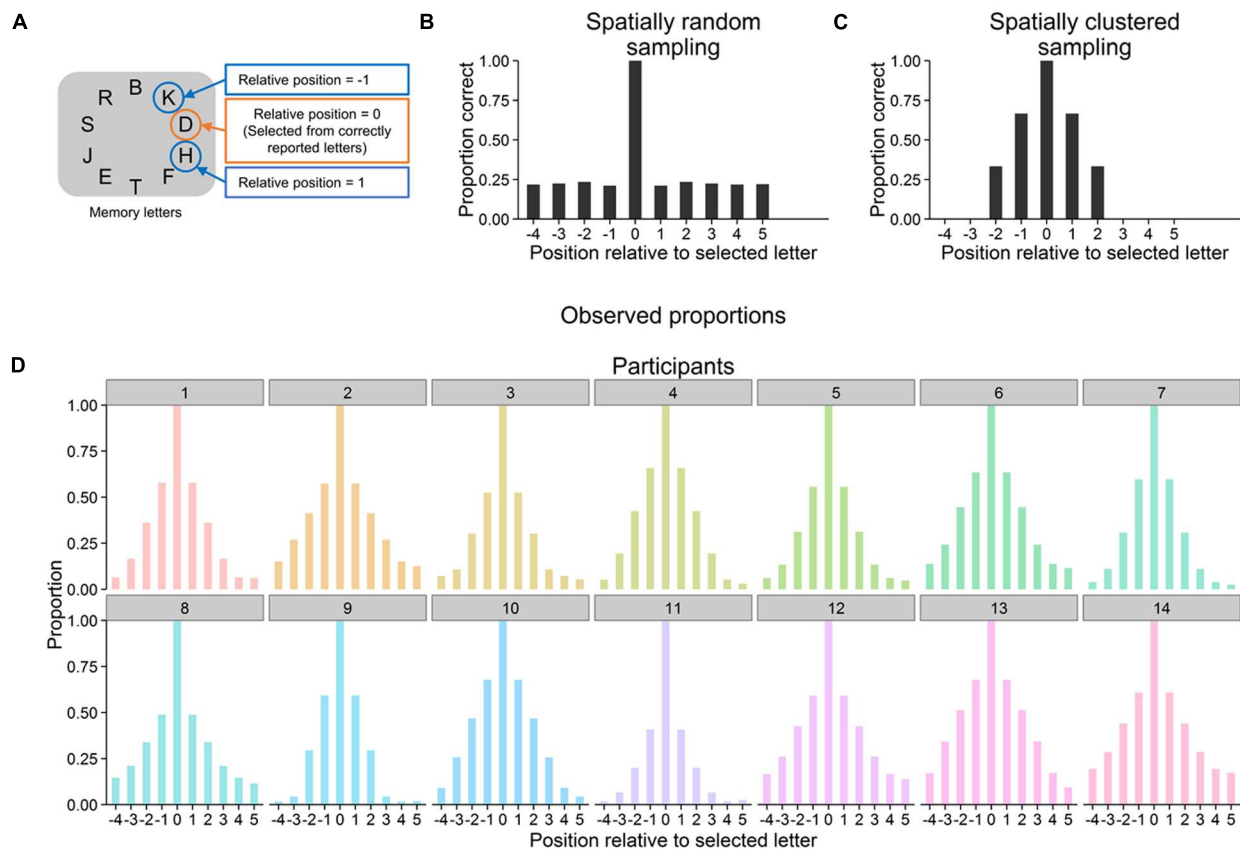


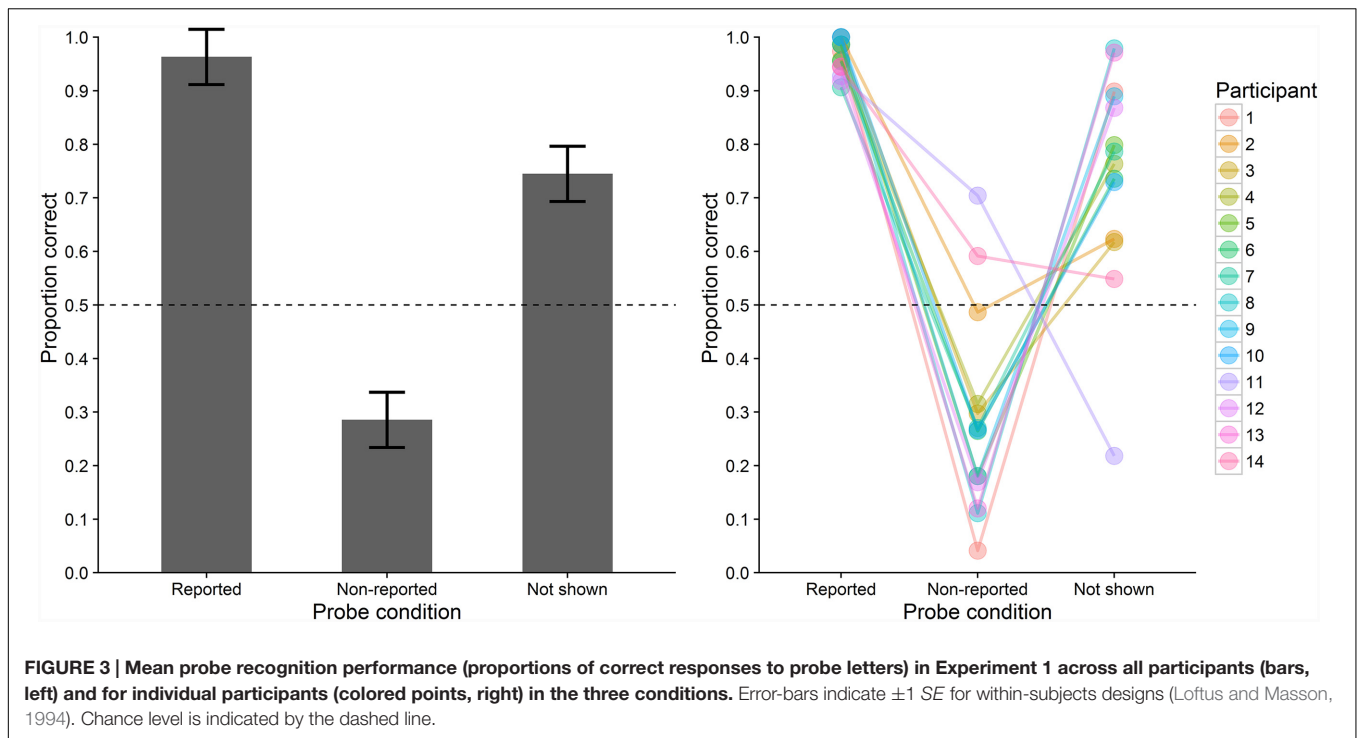
FIGURE 2 | (A) For each correctly reported letter, it was determined whether letters at the 10 relative positions in the display of memory letters were reported as well. Averaged across all correctly reported letters, this resulted in a proportion of correctly reported letters for each relative position. **(B)** Simulation of 1000 trials in each of which three letters were randomly sampled from the display of memory letters. Proportions of correctly reported letters are shown as a function of position relative to a selected correctly reported letter (averaged across trials). Proportions are approximately equal for all relative positions except for the position of the selected letter (which must always equal one). **(C)** Simulation of 1000 trials in each of which three letters were sampled that were next to each other in the display of memory letters. Proportions monotonically decrease from relative positions closer to selected letters to relative positions farther away from them. **(D)** For each individual participant, observed proportions of correctly reported letters as a function of position relative to a selected reported letter (across all correctly reported letters and averaged over all trials).

chance in the not shown condition, $t(13) = 4.589$, $p < 0.001$, $BF = 0.014$.

Whether probe recognition depended on how many letters participants entered for the whole report (irrespective of whether letters were correct) was assessed as the point-biserial correlation between the number of entered letters and probe recognition performance, separately for each participant and each condition. Values of three participants in the reported condition had to be excluded from this analysis because probe recognition was correct in all trials so that no correlation could be computed. One-sample t -tests (two-sided) indicated that the correlations of the 11 remaining participants did not significantly depart from zero in any of the three conditions, all $|ts| (10) < 1.713$, all $ps > 0.110$, all $BFs > 1.149$.

Probe recognition performance was close to ceiling in the reported condition but it was substantially lower in the non-reported and not shown conditions. These findings clearly argue against the type-activation hypothesis which predicts

equal performance for all presented memory letters and hence equal performance in the reported and non-reported condition. Instead, the findings seem to support the VWM-encoding hypothesis which predicts higher performance in the reported condition, in which probe letters were encoded into VWM. However, before arriving at these conclusions, several issues should be considered. According to the VWM-encoding hypothesis, performance should have been at chance level in the non-reported condition but it was below chance level. This may indicate that participants based their probe responses not only on the letters they remembered having viewed on this trial. Rather, they may have partly based their responses on the letters they remembered having reported on this trial. This would have biased them away from responding those probes had been contained in the memory letters when they had not reported the letters of these probes. This bias might also have contributed to the above-chance performance in the not shown condition. Besides biasing responses, reporting the letters itself



might also have improved their subsequent episodic short-term recognition compared to non-reported letters. Similarly, reporting memory letters might have interfered with retaining non-reported letters. In addition, reporting the letters may have prolonged the interval that the non-reported letters had to be retained. In all of these cases, letters that were inaccessible for report might have been available for later episodic short-term recognition if intervening report requirements were controlled for. Therefore, the aim of Experiment 2 was to control for all effects reporting letters might have on probe recognition performance.

EXPERIMENT 2

Experiment 2 was designed to investigate episodic short-term recognition performance for letters that were more likely to be encoded into VWM compared with letters whose encoding was less likely. To manipulate the likelihood of encoding specific letters into VWM, we made use of the spatial clustering of VWM encoding found in Experiment 1. Participants briefly viewed a display of 10 letters in which a colored frame identified one letter as report-target and frames in a different color identified the nine other letters as non-targets *regarding report*. Participants' task was to report the single report-target after a retention interval. After reporting, a single probe letter was shown and participants were to indicate whether or not it had been presented as one of the preceding letters (Figure 4). There were three conditions. In the *report-target condition*, the probe tested recognition of the report-target. In the *near non-target condition*, the probe tested recognition of a letter that has been located directly beside the

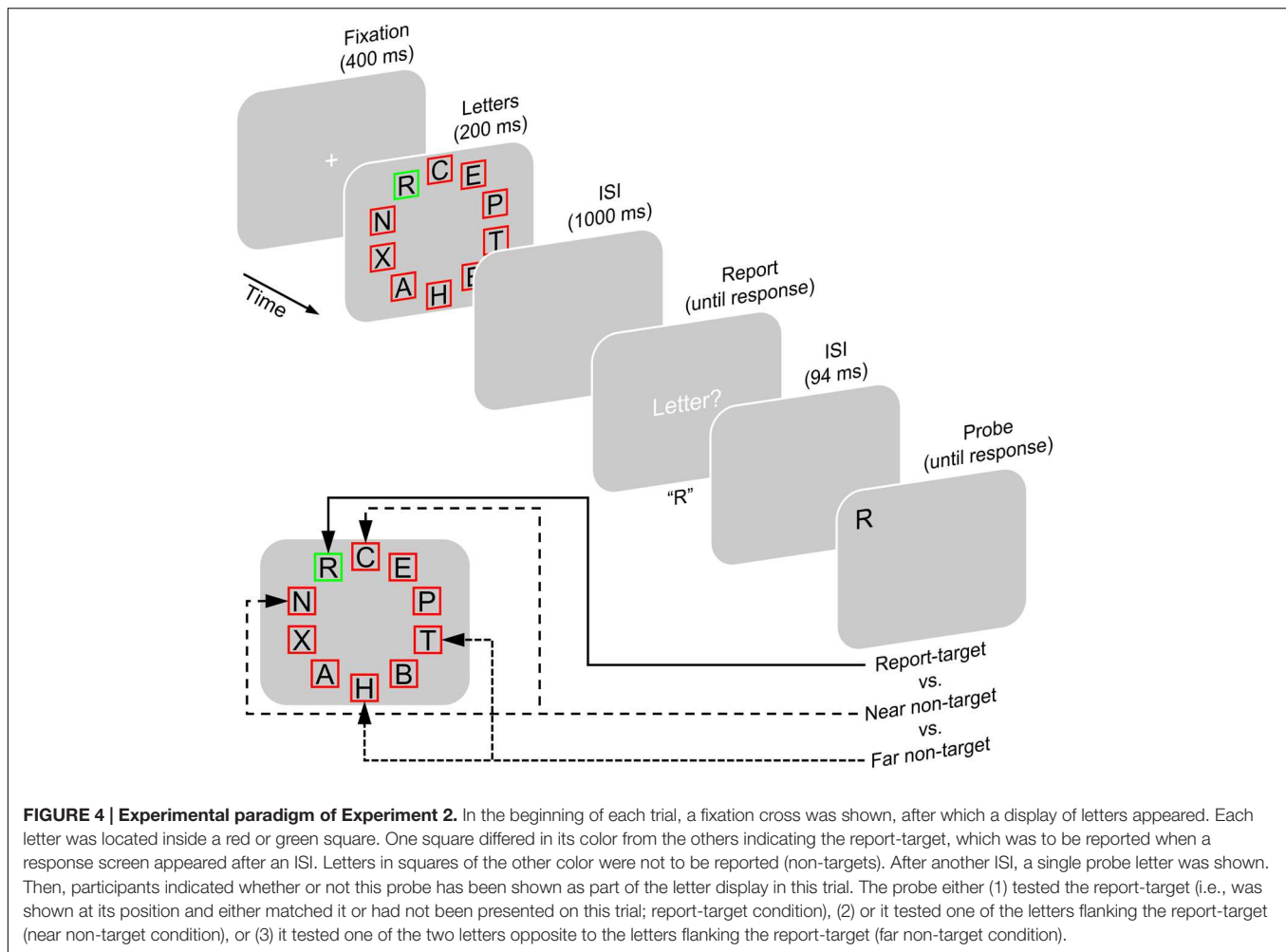
report-target. In the *far non-target condition*, the probe tested recognition of a letter that has been located far away from the report-target, on the other side of the letter display.

The report-target has to be encoded into VWM, in order to be accessible for being reported (e.g., Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013). Because of the spatial clustering of letter reports in Experiment 1, we assumed that while participants aimed at encoding the report-target, they were more likely to encode near non-targets selectively compared with far non-targets. This is compatible with the view that spatial attention was primarily directed at the report-target (e.g., Kim and Cave, 1995; Gaspelin et al., 2015), but was secondarily directed more at near non-targets than at far non-targets or was secondarily directed at near non-targets only. According to the VWM-encoding hypothesis, probe recognition performance should be highest for report-targets, followed by near non-targets, and lowest for far non-targets because of their lowest likelihood of being encoded into VWM. In contrast, according to the type-activation hypothesis probe recognition performance should be equal for all presented letters and thus equal in all three conditions. Importantly, the near and far non-targets were not subject to report requirements.

Method

Participants

Ten paid participants took part in Experiment 2. They were between 22 and 30 years old ($Mdn = 25$). Four of them were male, six female, nine were right, and one left-handed. All participants reported normal or corrected-to-normal visual acuity and color vision. They gave written informed consent before performing the experiments that were conducted according to the ethical



standards of the German Psychological Association (DGPs), and were approved by Bielefeld University's ethics committee.

Apparatus and Stimuli

The apparatus and experimental setup of Experiment 2 were the same as those of Experiment 1. The stimuli of Experiment 2 were identical to those of Experiment 1 with the following exceptions. All letters were placed inside a square frame ($0.72^\circ \times 0.72^\circ$). Frames of the nine non-targets were either all red ($20 \text{ cd} \times \text{m}^{-2}$; RGB: 255, 0, 0) or green ($76 \text{ cd} \times \text{m}^{-2}$; RGB: 0, 255, 0). The frame of the report-target was in the other color (i.e., green when the others were red or red when the others were green). The colors of report-target and non-targets remained the same throughout the experiment. Whether red or green indicated the report-target was counterbalanced across the sample. The text of the response screen was identical to that in Experiment 1, except that it prompted participants to enter only one instead of several letters (by the German text "Buchstabe?", which means "Letter?" in English).

Procedure and Design

As illustrated in **Figure 4**, the experimental paradigm of Experiment 2 was identical to that of Experiment 1 except for

the following aspects. Instead of all 10 letters, participants were to report only the one report-target (partial report). On each trial, the position of the report-target was randomly chosen. No confirmation of this report was required, instead the trial proceeded as soon as a letter-key had been pressed. As in Experiment 1, at the end of each trial a single probe letter was shown and participants were required to indicate whether or not it was shown within the letter display of this trial. Participants performed three conditions of a within-subjects design. In the *report-target condition*, the probe appeared at the location of the report-target and either matched the report-target or consisted in a letter not presented on this trial. In the *near non-target condition*, the probe appeared at the location of one of the two letters that flanked the report-target and either matched this letter or had not been presented on this trial. In the *far non-target condition*, the probe appeared at the location of one of the two letters opposite to the two flanking letters, on the other side of the letter display than the report-target and either matched this letter or had not been shown on this trial.

Participants performed four blocks of 72 trials each comprising 24 trials of the report-target, near non-target, and far non-target condition. For the two non-target conditions,

probes appeared equally often at positions in clockwise or counter-clockwise direction of the report-target. In each of the three conditions and for each possible probe location, trials with probes matching the former letter at the probe's location (correct answer "yes") and probes not shown (correct answer "no") occurred equally often. Participants performed 24 training trials prior to the experiment.

Results and Discussion

The same statistical procedures were used as in Experiment 1. Two trials were excluded from analysis because participants entered more than one letter in their letter report (which could happen only if participants pressed two keys close to simultaneously). Whether report-targets were in red or green frames did not interact with any of the below described dependent variables, all $F_s < 1.64$, all $p_s > 0.227$ (revealed by a repeated-measures ANOVA with type III sums-of-squares). Therefore, data of participants with report-targets in red and green frames was collapsed for the following analyses.

Letter Report Performance

Letter report performance was assessed as participants' proportion of trials on which the report-target was correctly reported. Unsurprisingly, there were no significant differences between letter report performance in the three experimental probe conditions (report-target condition: $M = 0.94$, $SD = 0.07$; near non-target condition: $M = 0.94$, $SD = 0.06$; far non-target condition: $M = 0.93$, $SD = 0.07$), $F(2,18) = 0.545$, $p = 0.589$, $\eta^2_G < 0.004$. In addition, Friedman's test was applied, because the assumption of normal distribution of the repeated-measures analysis of variance was not met. This test yielded a non-significant effect as well, $\chi^2(2) = 1.316$, $p = 0.518$.

Participants' letter report performance did not differ reliably between the three conditions. Participants achieved close-to-ceiling performance in all three conditions, as could be expected since only one letter had to be reported which should not touch the capacity limit of VWM (Sperling, 1960; Shibuya and Bundesen, 1988).

Probe Recognition Performance

Different from Experiment 1, each condition contained trials in which probes did and trials in which probes did not match the letters they referred to. Therefore, probe recognition performance could be quantified as d' , the difference between the z -transformed rate of correct responses to probes shown on this trial, z ("hit rate"), and the z -transformed rate of false responses

to probes not shown on this trial, z ("false alarm rate"; for an overview, see Macmillan and Creelman, 2005). Performance at chance level leads to a d' of zero and close to perfect performance to values of 4.65 (or higher and 0.5 was added to all data cells on which hit and false alarm rates were based to avoid infinite values for d' , Macmillan and Creelman, 2005, pp. 8–9). To facilitate comparison with the results of Experiment 1, in **Table 1** we also report the probe recognition performance assessed as the proportion of trials on which probe letters were correctly recognized as having been shown or not shown on the trial.

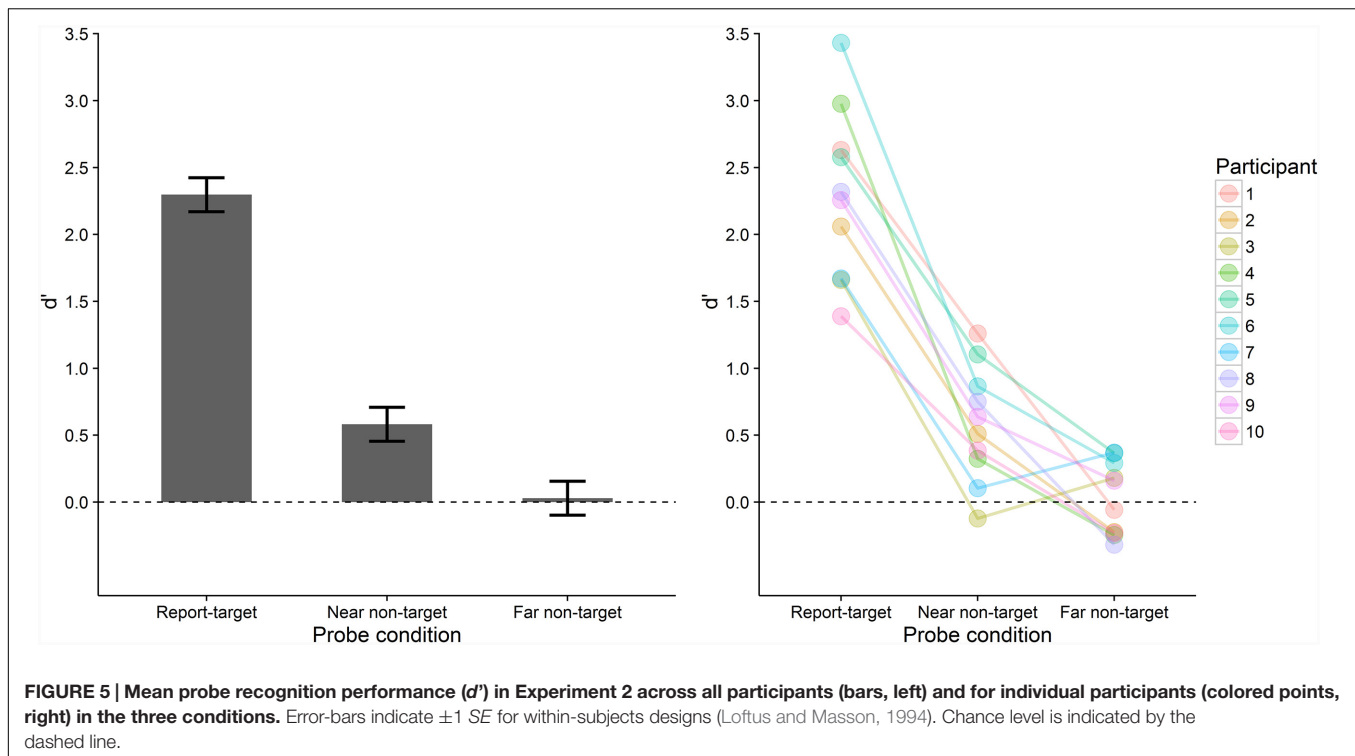
Figure 5 depicts participants' probe recognition performance in the three conditions at the sample and individual level. Performance differed significantly between the three conditions, $F(2, 18) = 86.859$, $p < .001$, $\eta^2_G = 0.824$. That is, performance was significantly higher in the report-target ($M = 2.30$, $SD = 0.63$) compared with the near non-target ($M = 0.58$, $SD = 0.43$), $t(9) = 10.562$, $p_B < 0.001$, $d_z = 3.34$, $BF = 1.3 \times 10^{-4}$, and far non-target condition ($M = 0.03$, $SD = 0.27$), $t(9) = 10.770$, $p_B < 0.001$, $d_z = 3.41$, $BF = 1.2 \times 10^{-4}$. Performance was also significantly higher in the near than in the far non-target condition, $t(9) = 3.435$, $p_B = 0.022$, $d_z = 1.09$, $BF = 0.127$. This data pattern was present in all except two participants whose performance was slightly higher in the far compared with the near non-target condition (**Figure 5**, right). In addition, two one-sample t -tests revealed that performance was significantly above chance in the near non-target condition, $t(9) = 4.262$, $p = 0.002$, $BF = 0.045$, but did not differ from chance level (i.e., a d' of zero) in the far non-target condition, $t(9) = 0.335$, $p > 0.745$, $BF = 3.086$.

Probe recognition performance was highest when the probe letter tested the former report-target which had been encoded into VWM, as evident from the near-ceiling performance in reporting its identity. Importantly, performance was higher for near non-targets than for far non-targets. This indicates that episodic short-term recognition was better for letters that were more likely to be encoded into VWM compared with letters less likely to be encoded (given that encoding into VWM seems to proceed in a spatially clustered manner, see Experiment 1). In fact, performance for far non-targets was at chance level which suggests that episodic short-term recognition was not possible for these letters. Furthermore, Experiment 2 controlled for potential alternative explanations of the findings of Experiment 1. These alternative explanations stated that differences between the conditions did not stem from whether letters were encoded into VWM but from whether letters were reported. In Experiment

TABLE 1 | Probe recognition performance in the three conditions of Experiment 2 assessed as the proportion of correct responses to the probe.

	<i>M (SD)</i>	vs. report-target	vs. near non-target
Report-target	0.79 (0.09)	–	–
Near non-target	0.61 (0.08)	$t(9) = -6.515$, $p_B < 0.001$, $d_z = -2.06$, $BF = 3.79 \times 10^{-3}$	–
Far non-target	0.51 (0.05)	$t(9) = -9.313$, $p_B < 0.001$, $d_z = -2.94$, $BF = 3.30 \times 10^{-4}$	$t(9) = -3.451$, $p_B = 0.022$, $d_z = -1.09$, $BF = 0.125$

Provided are means (*M*) with standard deviations (*SD*) in parentheses, and for each of the three conditions (rows), the results of the pairwise comparisons to the other two conditions (last two columns).



2, near and far non-targets both did not have to be reported and differed only in their distance from the report-target. Hence, the performance difference between these two conditions cannot be attributed to effects reporting letters itself might have on performance. Therefore, we interpret the higher performance for near non-targets compared with the performance at chance level for far non-targets as strong evidence for the VWM-encoding hypothesis. Conversely, we interpret this finding as evidence against the type-activation hypothesis.

GENERAL DISCUSSION

We investigated whether episodic short-term recognition of objects from a previous processing episode requires that these objects have been encoded into VWM. For this purpose, we introduced a new paradigm combining letter report with probe recognition. In two experiments, episodic short-term recognition was assessed as performance in recognizing whether a probe letter was presented in the preceding letter display of the current trial. In Experiment 1, probe recognition performance was higher for letters that had been encoded into VWM compared with letters that had not been encoded. In Experiment 2, only a single letter had to be reported in the letter report task. This controlled for effects reporting letters itself might have on probe recognition. In Experiment 2, probe recognition performance was higher for non-target letters that were near to a report-target letter, and hence more likely to be encoded into VWM, compared with non-target letters far from the report-target, whose encoding was less likely. Crucially, this difference in probe recognition refers to non-target letters which did not have to be reported. Strikingly,

performance was at chance level for letters far from the report-target which were unlikely to enter VWM. Therefore, we interpret the present findings as strong evidence for the VWM-encoding hypothesis which states that episodic short-term recognition presupposes that visual objects have been encoded into VWM. Conversely, we interpret these findings as evidence against the type-activation hypothesis. Note that one might distinguish a strong form of the type-activation hypothesis, the one that we have put forward so far, and a weaker form. The strong form states that episodic short-term recognition can be accomplished perfectly (at least in principle) for all objects of the current visual field. In contrast, the weaker form states that episodic short-term recognition can be accomplished for all objects of the visual field, but not perfectly, and that recognition performance may be improved by additional encoding into VWM. The results of Experiment 2 provide evidence against both forms of the type-activation hypothesis. The finding that probe recognition performance was higher for near than for far non-targets argues against the strong form. The finding that performance was at chance level in the far non-target condition argues against the weak form. That is, episodic short-term recognition seemed impossible in this condition. Thus, taken together, the present findings indicate that type-activation is not sufficient for later episodic short-term recognition but that encoding into VWM is required instead.

Visual Working Memory as a Basis of Episodic Short-Term Recognition

Encoding an object into VWM seems to be necessary for its later episodic short-term recognition. This means that the functional

basis of episodic short-term recognition emerges at a level of processing after the activation of visual types in visual long-term memory (e.g., Kanwisher, 1987; Kahneman et al., 1992; cf. Schneider, 1995) and after visual attention has mediated selective encoding into VWM (Duncan and Humphreys, 1989; Bundesen, 1990; Bundesen et al., 2005; Schneider, 2013). In the present study, letters were used as visual objects. After successful visually based recognition, letters can be processed verbally, which makes it likely that their episodic short-term recognition also involved verbal processing in addition to visual processing. However, because the letters had to be acquired visually, they had to be encoded into VWM first, before such a verbal processing could take place. After their encoding into VWM, they may have been recoded into a verbal format. Such a verbal format may have provided the advantage of verbal rehearsal by verbal working memory, which may have prolonged and secured their retention (e.g., Logie, 2011; Baddeley, 2012). Thus, importantly, even though episodic short-term recognition may rely on several different (working) memory mechanisms (such as visual and verbal ones), encoding into VWM seems to be a necessary processing step for these mechanisms to operate.

Why may encoding into VWM be necessary for episodic short-term recognition? Several theories assume that by encoding into VWM, information about visual objects is transformed into a special representational state (e.g., Cowan, 1988; Oberauer, 2002; LaRocque et al., 2014; cf. Olivers et al., 2011). We suggest that it is this representational state that makes encoding into VWM a requirement of episodic short-term recognition. Specifically, we propose that two characteristics of this representational state are necessary for episodic short-term recognition: *binding* and *robustness*.

Binding means that different visual features of an object are integrated which yields representations of objects as a whole, with all their features (e.g., Treisman and Gelade, 1980). The mere presentation of objects activates visual types (features) in visual long-term memory but this happens in isolation (cf. Bundesen, 1990; Schneider, 1995). Episodic short-term recognition requires binding of activated visual types because otherwise objects that share visual features cannot be distinguished. VWM is assumed to mark the first level in the course of visual processing at which the visual types (or features) activated by an object are bound to integrated object representations (Bundesen, 1990; Luck, 2008; Schneider, 2013; Kyllingsbæk, 2014). This point is illustrated by referring to integrated object representations as VWM objects (Schneider, 2013), which have also been called object files (Kahneman et al., 1992) and visual tokens (Schneider, 1995). In sum, the binding of visual types within object representations in VWM may be one reason for that episodic short-term recognition requires encoding into VWM.

Robustness means that object representations in VWM are protected against so-called proactive interference (Keppel and Underwood, 1962). Proactive interference arises when the same visual objects occur repeatedly (e.g., Endress and Potter, 2014). It describes an impairment in recognizing if an object has been viewed in the very recent past as opposed to having been encountered before at all (e.g., Endress and Potter, 2014). Episodic short-term recognition clearly requires

to assess whether an object has been viewed in a recently passed episode rather than at some unspecified point in the past. Hence, successful episodic short-term recognition presupposes that proactive interference is eliminated. Robustness against proactive interference is assumed to be a hallmark of VWM representations and providing it is considered a core function of VWM (Endress and Potter, 2014). Thus, taken together, episodic short-term recognition may presuppose encoding of objects into VWM because this might establish representations of objects as bound units (cf. Luck and Vogel, 1997) which are robust against proactive interference (cf. Endress and Potter, 2014).

Episodic Short-Term Recognition Might Be Constrained by an Encoding-Limitation but Not a Retention-Limitation of Visual Working Memory

As we have argued, the present findings indicate that episodic short-term recognition presupposes encoding into VWM but this seems to conflict with earlier findings. Specifically, Sternberg (1966) presented participants with series of up to six digits followed by a probe digit. Participants indicated whether the probe was contained in a given series. The six presented digits exceed the number of about three to four objects that VWM can hold (e.g., Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997). Thus, when the last two digits were shown, VWM should have already been filled up so that the digits could not be encoded into VWM. Nevertheless, Sternberg found that probe recognition performance was close to ceiling even for six digits. One might attribute this result to the relatively long presentation durations of digits (1.2 s) that could have allowed verbal rehearsal (e.g., Sternberg, 1975). However, congruent to Sternberg's findings, later experiments revealed high levels of probe recognition performance for objects that were presented more briefly and thus difficult to rehearse verbally (Endress and Potter, 2014). Taken together, these findings are compatible with the type-activation hypothesis in that they suggest episodic short-term recognition is possible also for objects that have not reached VWM.

How may the conflict between the present and Sternberg's (1966; cf. Endress and Potter, 2014) findings be resolved? One solution is provided by Schneider's (2013) recent "theory of task-driven visual attention and working memory" (TRAM) which offers an account of how visual information processing might be accomplished within and across processing episodes. According to TRAM, a new processing episode is started with each onset of visual objects (e.g., after a saccadic eye movement). A processing episode comprises three phases. Premising upon Bundesen's (1990) theory of visual attention (a model of biased competition, Desimone and Duncan, 1995), TRAM's first two phases describe how visual attention mediates selective encoding of visual objects into capacity-limited VWM. In TRAM's third phase, objects that have been encoded into initial activation-based VWM (i.e., VWM based on persistent neural activity) are consolidated which results in *passive* VWM representations

(which do not require neural activity but may rely on short-term changes in synaptic connectivity, as reviewed by Eriksson et al., 2015; Postle, 2015; and Stokes, 2015). Critically, according to TRAM, the number of passive VWM representations is not constrained by the traditionally assumed capacity-limitation of VWM. With this in mind, one may interpret classical estimates of VWM capacity (Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997) as reflecting an encoding limitation but not a retention limitation. In other words, classical VWM capacity may constrain the amount of object information that can be acquired within one processing episode but not the amount of information that can be retained across episodes. In Sternberg's (1966) paradigm, each of the serially presented digits should have started a new processing episode. Within each of these episodes, a passive VWM representation of the digit should have emerged. Probe recognition should then have been based on a comparison of these passive VWM representations with actual probe digits (which could involve retrieving passive representations again into classical activation-based VWM; Schneider, 2013). In this vein, episodic short-term recognition becomes possible for more serially presented objects than classical VWM can retain. In contrast, TRAM posits that if several objects are presented simultaneously, as in the present experiments, then this can reach the encoding limit of VWM. All simultaneously presented objects are processed within the same processing episode. Therefore, encoding further objects becomes impossible if activation-based VWM is filled up. Critically, creating passive VWM representations of objects presupposes that the objects have been encoded into VWM. Thus, in a given processing episode, only as many objects as VWM can hold can be consolidated into passive VWM representations. As a consequence, episodic short-term recognition across successive processing episodes should be limited with respect to the number of simultaneously shown objects that can be encoded into VWM. In contrast, episodic short-term recognition should not be restricted with respect to the number of retained objects in VWM because this includes also passive VWM representations that have arisen over the course of several episodes, as in Sternberg's experiments. Interestingly, recent findings might suggest that in such situations of serial object presentations (RSVP), the capacity of passive VWM can be extended beyond "magical number four" by eliminating proactive interference (Endress and Potter, 2014). As an alternative to consolidation in passive VWM, representations of objects in classical VWM could also be recoded into a different representational format (Petersen et al., 2012) which might

then be used for later episodic short-term recognition. The objects of the present experiments consisted of letters which may have been recoded into the verbal format (that is open to verbal rehearsal, e.g., Sternberg, 1975, and may allow retention by working memory systems dedicated to verbal information, e.g., Baddeley, 2012). However, since the to-be-recoded object information is acquired visually, recoding would still presuppose encoding into VWM (Petersen et al., 2012). Hence, episodic short-term recognition would still be constrained by the encoding limitation of VWM but not by a retention limit. However, testing this hypothesis is left for further experimental studies.

CONCLUSION

The present study shows that episodic short-term recognition of objects from previous episodes presupposes that the objects have been processed up to the level of VWM. In this way, VWM not only provides bound visual objects for online perception and action within a processing episode but also paves the way for episodic short-term recognition across episodes. However, this also implies that episodic short-term recognition is only possible for a limited number of simultaneously presented objects due to the encoding limitation of VWM (Schneider, 2013; cf. Sperling, 1960; Shibuya and Bundesen, 1988; Luck and Vogel, 1997).

AUTHOR CONTRIBUTIONS

CP and WS designed the research. CP programmed the experiments and analyzed the data. CP and WS wrote the paper.

FUNDING

This research was funded by the DFG, Cluster of Excellence 277 "Cognitive Interaction Technology (CITEC)". We acknowledge support for the Article Processing Charge by the Deutsche Forschungsgemeinschaft and the Open Access Publication Fund of Bielefeld University.

ACKNOWLEDGMENT

We thank Katharina Weiß for helpful comments on an earlier draft of this article.

REFERENCES

- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi: 10.3758/BF03192707
- Bays, P. M., and Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science* 321, 851–854. doi: 10.1126/science.1158023
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Bundesden, C. (1990). A theory of visual attention. *Psychol. Rev.* 97, 523–547. doi: 10.1037/0033-295X.97.4.523
- Bundesden, C., and Habekost, T. (2008). *Principles of Visual Attention: Linking Mind and Brain*. Oxford: Oxford University Press.
- Bundesden, C., Habekost, T., and Kyllingsbæk, S. (2005). A neural theory of visual attention: bridging cognition and neurophysiology. *Psychol. Rev.* 112, 291–328. doi: 10.1037/0033-295X.112.2.291

- Bundesden, C., Habekost, T., and Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia* 49, 1446–1457. doi: 10.1016/j.neuropsychologia.2010.12.006
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Psychology Press.
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. doi: 10.1037/0033-2909.104.2.163
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. doi: 10.1017/S0140525X01003922
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205
- Donkin, C., and Nosofsky, R. M. (2012). The structure of short-term memory scanning: an investigation using response time distribution models. *Psychon. Bull. Rev.* 19, 363–394. doi: 10.3758/s13423-012-0236-8
- Duncan, J. (2006). EPS Mid-Career Award 2004: brain mechanisms of attention. *Q. J. Exp. Psychol.* 59, 2–27. doi: 10.1080/17470210500260674
- Duncan, J. (2013). The structure of cognition: attentional episodes in mind and brain. *Neuron* 80, 35–50. doi: 10.1016/j.neuron.2013.09.015
- Duncan, J., and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychol. Rev.* 96, 433–458. doi: 10.1037/0033-295X.96.3.433
- Dyrholm, M., Kyllingsbæk, S., Espeseth, T., and Bundesden, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: the case of TVA. *J. Math. Psychol.* 55, 416–429. doi: 10.1016/j.jmp.2011.08.005
- Endress, A. D., and Potter, M. C. (2014). Large capacity temporary visual memory. *J. Exp. Psychol. Gen.* 143, 548–565. doi: 10.1037/a0033934
- Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., and Nyberg, L. (2015). Neurocognitive architecture of working memory. *Neuron* 88, 33–45. doi: 10.1016/j.neuron.2015.09.020
- Gaspelin, N., Leonard, C. J., and Luck, S. J. (2015). Direct evidence for active suppression of salient-but-irrelevant sensory inputs. *Psychol. Sci.* 26, 1740–1750. doi: 10.1177/0956797615597913
- Henderson, J. M. (1994). Two representational systems in dynamic visual identification. *J. Exp. Psychol. Gen.* 123, 410–426. doi: 10.1037/0096-3445.123.4.410
- Henderson, J. M., and Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 826–839. doi: 10.1037/0096-1523.20.4.826
- Herwig, A. (2015). Linking perception and action by structure or process? Toward an integrative perspective. *Neurosci. Biobehav. Rev.* 52, 105–116. doi: 10.1016/j.neubiorev.2015.02.013
- Irwin, D. E., and Thomas, L. E. (2008). “Visual sensory memory,” in *Visual Memory*, eds S. J. Luck and A. Hollingworth (Oxford: Oxford University Press), 9–41.
- Jensen, O., and Lisman, J. E. (1998). An oscillatory short-term memory buffer model can account for data on the Sternberg task. *J. Neurosci.* 18, 10688–10699.
- Kahana, M. J., and Sekuler, R. (2002). Recognizing spatial patterns: a noisy exemplar approach. *Vision Res.* 42, 2177–2192. doi: 10.1016/S0042-6989(02)00118-9
- Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219. doi: 10.1016/0010-0285(92)90007-O
- Kanwisher, N. G. (1987). Repetition blindness: type recognition without token individuation. *Cognition* 27, 117–143. doi: 10.1016/0010-0277(87)90016-3
- Keppel, G., and Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *J. Verbal Learning Verbal Behav.* 1, 153–161. doi: 10.1016/S0022-5371(62)80023-1
- Kim, M.-S., and Cave, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychol. Sci.* 6, 376–380. doi: 10.1111/j.1467-9280.1995.tb00529.x
- Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36, 1–16.
- Klotz, W., and Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 976–992. doi: 10.1037/0096-1523.25.4.976
- Krock, R. M., and Moore, T. (2015). The influence of gaze control on visual perception: eye movements and visual stability. *Cold Spring Harb. Symp. Quant. Biol.* 79, 123–130. doi: 10.1101/sqb.2014.79.024836
- Kyllingsbæk, S. (2014). *Modeling Visual Attention: Encoding, Attention, and Short-Term Memory*. Doctoral dissertation, University of Copenhagen, Copenhagen.
- LaRocque, J. J., Lewis-Peacock, J. A., and Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Front. Hum. Neurosci.* 8:5. doi: 10.3389/fnhum.2014.00005
- Loftus, G. R., and Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* 1, 476–490. doi: 10.3758/BF03210951
- Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Curr. Dir. Psychol. Sci.* 20, 240–245. doi: 10.1177/0963721411415340
- Luck, S. J. (2008). “Visual short-term memory,” in *Visual Memory*, eds S. J. Luck and A. Hollingworth (Oxford: Oxford University Press), 43–83.
- Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281. doi: 10.1038/36846
- Luck, S. J., and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400. doi: 10.1016/j.tics.2013.06.006
- Ma, W. J., Husain, M., and Bays, P. M. (2014). Changing concepts of working memory. *Nat. Neurosci.* 17, 347–356. doi: 10.1038/nn.3655
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*. New York, NY: Psychology Press.
- Martens, S., and Wyble, B. (2010). The attentional blink: past, present, and future of a blind spot in perceptual awareness. *Neurosci. Biobehav. Rev.* 34, 947–957. doi: 10.1016/j.neubiorev.2009.12.005
- Nosofsky, R. M., Little, D. R., Donkin, C., and Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychol. Rev.* 118, 280–315. doi: 10.1037/a0022494
- Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 411–421. doi: 10.1037/0278-7393.28.3.411
- Oberauer, K., and Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Mem. Cogn.* 41, 1212–1227. doi: 10.3758/s13421-013-0333-6
- Olivers, C. N. L., Peters, J., Houtkamp, R., and Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15, 327–334. doi: 10.1016/j.tics.2011.05.004
- Palmeri, T. J., and Tarr, M. J. (2008). “Visual object perception and long-term memory,” in *Visual Memory*, eds S. J. Luck and A. Hollingworth (Oxford: Oxford University Press), 163–207.
- Pearson, B., Raskevicius, J., Bays, P. M., Pertzov, Y., and Husain, M. (2014). Working memory retrieval as a decision process. *J. Vis.* 14, 2. doi: 10.1167/14.2.2
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Petersen, A., Kyllingsbæk, S., and Bundesden, C. (2012). Measuring and modeling attentional dwell time. *Psychon. Bull. Rev.* 19, 1029–1046. doi: 10.3758/s13423-012-0286-y
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Percept. Psychophys.* 16, 283–290. doi: 10.3758/BF03203943
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/0033558008248231
- Postle, B. R. (2015). The cognitive neuroscience of visual short-term memory. *Curr. Opin. Behav. Sci.* 1, 40–46. doi: 10.1016/j.cobeha.2014.08.004
- Poth, C. H., Herwig, A., and Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Front. Syst. Neurosci.* 9:176. doi: 10.3389/fnsys.2015.00176
- Poth, C. H., Petersen, A., Bundesden, C., and Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Front. Psychol.* 5:930. doi: 10.3389/fpsyg.2014.00930
- Poth, C. H., and Schneider, W. X. (2016). Breaking object correspondence across saccades impairs object recognition: the role of color and luminance. *J. Vis.* 16, 1–12. doi: 10.1167/16.11.1
- R Development Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R foundation for statistical computing.

- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225
- Schneider, W. X. (1995). VAM: a neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Vis. Cogn.* 2, 331–376. doi: 10.1080/13506289508401737
- Schneider, W. X. (2013). Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20130060. doi: 10.1098/rstb.2013.0060
- Shibuya, H., and Bundesen, C. (1988). Visual selection from multielement displays: measuring and modeling effects of exposure duration. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 591–600. doi: 10.1037/0096-1523.14.4.591
- Soto, D., Mäntylä, T., and Silvanto, J. (2011). Working memory without consciousness. *Curr. Biol.* 21, R912–R913. doi: 10.1016/j.cub.2011.09.049
- Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* 74, 1–29. doi: 10.1037/h0093759
- Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652
- Sternberg, S. (1975). Memory scanning: new Findings and current controversies. *Q. J. Exp. Psychol.* 27, 1–32. doi: 10.1080/14640747508400459
- Stokes, M. G. (2015). ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405. doi: 10.1016/j.tics.2015.05.004
- Strasburger, H., Rentschler, I., and Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *J. Vis.* 11, 13. doi: 10.1167/11.5.13
- Treisman, A., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5
- Wheeler, M. E., and Treisman, A. M. (2002). Binding in short-term visual memory. *J. Exp. Psychol. Gen.* 131, 48–64. doi: 10.1037/0096-3445.131.1.48
- Wickelgren, W. A. (1970). Time, interference, and rate of presentation in short-term recognition memory for items. *J. Math. Psychol.* 7, 219–235. doi: 10.1016/0022-2496(70)90045-3
- Wilken, P., and Ma, W. J. (2004). A detection theory account of change detection. *J. Vis.* 4, 1120–1135. doi: 10.1167/4.12.11
- Zhou, F., Kahana, M. J., and Sekuler, R. (2004). Short-term episodic memory for visual textures: a roving probe gathers some memory. *Psychol. Sci.* 15, 112–118. doi: 10.1111/j.0963-7214.2004.01502007.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Poth and Schneider. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Spatio-Temporal Structure, Path Characteristics, and Perceptual Grouping in Immediate Serial Spatial Recall

Carlo De Lillo*, Melissa Kirby and Daniel Poole†

Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester, UK

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology Jodhpur,
India

Reviewed by:

Mark John Hurlstone,
University of Western Australia,
Australia

Steven Trawley,
Deakin University, Australia

*Correspondence:

Carlo De Lillo
CDL2@LE.AC.UK

†Present Address:

Daniel Poole,
Division of Neuroscience and
Experimental Psychology, University of
Manchester, Manchester, UK

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 30 June 2016

Accepted: 13 October 2016

Published: 11 November 2016

Citation:

De Lillo C, Kirby M and Poole D (2016)
Spatio-Temporal Structure, Path
Characteristics, and Perceptual
Grouping in Immediate Serial Spatial
Recall. *Front. Psychol.* 7:1686.
doi: 10.3389/fpsyg.2016.01686

Immediate serial spatial recall measures the ability to retain sequences of locations in short-term memory and is considered the spatial equivalent of digit span. It is tested by requiring participants to reproduce sequences of movements performed by an experimenter or displayed on a monitor. Different organizational factors dramatically affect serial spatial recall but they are often confounded or underspecified. Untangling them is crucial for the characterization of working-memory models and for establishing the contribution of structure and memory capacity to spatial span. We report five experiments assessing the relative role and independence of factors that have been reported in the literature. Experiment 1 disentangled the effects of spatial clustering and path-length by manipulating the distance of items displayed on a touchscreen monitor. Long-path sequences segregated by spatial clusters were compared with short-path sequences not segregated by clusters. Recall was more accurate for sequences segregated by clusters independently from path-length. Experiment 2 featured conditions where temporal pauses were introduced between or within cluster boundaries during the presentation of sequences with the same paths. Thus, the temporal structure of the sequences was either consistent or inconsistent with a hierarchical representation based on segmentation by spatial clusters but the effect of structure could not be confounded with effects of path-characteristics. Pauses at cluster boundaries yielded more accurate recall, as predicted by a hierarchical model. In Experiment 3, the systematic manipulation of sequence structure, path-length, and presence of path-crossings of sequences showed that structure explained most of the variance, followed by the presence/absence of path-crossings, and path-length. Experiments 4 and 5 replicated the results of the previous experiments in immersive virtual reality navigation tasks where the viewpoint of the observer changed dynamically during encoding and recall. This suggested that the effects of structure in spatial span are not dependent on perceptual grouping processes induced by the aerial view of the stimulus array typically afforded by spatial recall tasks. These results demonstrate the independence of coding strategies based on structure from effects of path characteristics and perceptual grouping in immediate serial spatial recall.

Keywords: Corsi test, grouping, spatial span, serial recall, spatial memory, working memory, chunking, virtual reality

INTRODUCTION

One of the most enduring problems in psychology and the neurosciences is the characterization of the mechanisms supporting the representation of serial order information (Lashley, 1951; Rosenbaum et al., 2007; Hurlstone et al., 2014). Serial Spatial Recall (SSR) refers to the ability to temporarily retain a sequence of spatial locations in a prescribed order and is one of the most common instantiations of the problem of serial order in short-term and working memory. The assessment of SSR is of central importance in several areas of psychological research. It has been used to evaluate the extent to which the processing of serial order in the verbal and visuo-spatial domain rests on similar mechanisms (Baddeley, 1992; Smyth and Scholey, 1992; Jones et al., 1995; Hurlstone et al., 2014), a crucial issue for the characterization of human cognitive architecture. SSR is one of the most widespread neuropsychological measures (Berch et al., 1998; Kessels et al., 2000) and is included as a test in widely used batteries (e.g., WAIS-R, Kaplan et al., 1991; Wechsler, 1997a; Wechsler Memory Scale, WMS-III, Wechsler, 1997b; Cantab, Cambridge Cognition, 2006). SSR has been extensively employed in the study of individual differences in working-memory (Cornoldi and Vecchi, 2003) and as a predictor of scholastic achievement (Jarvis and Gathercole, 2003; St Clair-Thompson, 2007). Because of its non-verbal nature, the assessment of SSR has been used for the comparison of memory skills in monkeys and humans, with important implications for the evaluation of primate models of human memory (Botvinick et al., 2009; Fagot and De Lillo, 2011).

Despite the popularity of SSR as a psychological measure and its suitability for addressing the problem of serial order from a cognitive, comparative, and neuropsychological perspective, its cognitive bases are still poorly understood and, as argued below, a number of central constructs for its description are often confounded. One of the most important issues to address in relation to SSR, as identified by a recent eminent review (Hurlstone et al., 2014) and as further elaborated below, is the characterization of the organizational factors that can contribute to accurate SSR (e.g., Kemps, 1999, 2001; Bor et al., 2003; De Lillo, 2004; Busch et al., 2005; Parmentier et al., 2005; Rossi-Arnaud et al., 2005; Parmentier and Andrés, 2006; Parmentier et al., 2006; Ridgeway, 2006; Imbo et al., 2009; De Lillo and Lesk, 2010).

SSR is typically measured by assessing spatial span with the Corsi test (Milner, 1971; Corsi, 1972), allegedly the most widely used non-verbal neuropsychological test (Berch et al., 1998; Kessels et al., 2000). In the Corsi test participants observe a sequence of spatial items, such as a series of finger tapping movements across an array of wooden blocks, or a series of flashing icons presented on a touch-screen. Then, they are required to reproduce the series by tapping the items in the same order. Because the items are all identical in shape and color, they need to be identified by their spatial position. For this reason, the Corsi test is considered one of the purest measures of spatial memory span (see Baddeley, 2001, for a review).

Traditionally, the Corsi test has featured irregular arrays of items and random sequences as recall material (Milner, 1971). However, it was realized soon that that not all random sequences

are recalled at the same level of accuracy (Smirni et al., 1983) and attempts to standardize the test ensued with important applied implications for the use of these tests for clinical diagnosis (Kessels et al., 2000; Busch et al., 2005).

The complexity of Corsi sequences has been manipulated in order to assess the relative autonomy of short and long-term memory structures (Kemps, 1999). The results of ingenious experiments have clarified that items in spatial working-memory are coded configurationally, using allocentric frames of reference (Avons, 2007; Avons and Oswald, 2008; Boduroglu and Shah, 2014), thus highlighting the role of relational properties of items in the display in recall.

Some studies (De Lillo, 2004, 2012; De Lillo and Lesk, 2010) have emphasized the notion that the understanding of the effects of organizational factors in SSR is of interest apart from the assessment of memory span *per se*. They proposed that with the irregular spatial arrangement of the items in Corsi-type tasks and randomly selected sequences of block tapping (see Berch et al., 1998 for examples of Corsi displays and criteria for selecting sequences that have been used in the literature), it is impossible to isolate the effect of particular organizational factors and interpret them in relation to the memory representation that they afford.

In order to assess the contribution of a specific type of organizational factor on spatial span De Lillo (2004) used a Corsi display, presented on a touch-screen, where 9 squares were arranged spatially to form 3 clusters of 3 items each, so that the separation of the items within clusters was inferior to that between clusters. The use of a configuration of items grouped in spatial clusters was motivated by different considerations. It seemed the appropriate way to convey in a Corsi task the fact that space can be divided in different sub-regions. It provided a spatial analogy of forms of semantic clustering and chunking observed in non-spatial domains. Finally, a configuration of items grouped in spatial clusters resembles a “patchy” foraging environment that according to foraging theories of cognitive evolution provided the pressures for the emergence of large brain and working memory skills in humans and other primates (e.g., Milton, 1993). Importantly, the use of a clustered Corsi display with items arranged in spatial clusters enables the manipulation of the serial organization of sequences so that they can be made either compatible or not with chunking by spatial proximity. De Lillo (2004) used different types of sequences. Some sequences were segregated by clusters, so that consecutive items were always in the same cluster and a transition to a different cluster occurred only after all the items within a cluster had been selected. Clustered sequences were deemed to afford a hierarchical representation because the order of the clusters, into which the sequence was segregated, could be stored independently from the order of the items within a given cluster. Other sequences were designed to be incompatible with such hierarchical organization because consecutive items were always in different clusters.

When recall for the two types of sequences was compared a beneficial “clustering effect” emerged; sequences segregated by clusters were reported at a higher level of accuracy. Consistently with a hierarchical model, in sequences segregated by clusters, longer Response Times (RTs) emerged at cluster boundaries.

This suggested that the retention of spatially clustered sequences could be supported by a hierarchical representation similar to that observed for chunking in non-spatial domains (Miller, 1956; Klahr et al., 1983). By contrast, non-clustered sequences showed longer reaction times for the items at intermediate ordinal positions within the sequences. This is a pattern of RT that resembles the serial position curve typically observed for lists of unrelated items in other domains, such as nonsense words, where items at intermediate ordinal positions are the most difficult to recall.

The recall of Corsi sequences typically shows a long initial RT which is indicative of the processing of serial order just before recall (see Fischer, 2001). Further evidence for the hierarchical representation of clustered Corsi sequences has been provided by showing that a component of this initial RT is proportional to the number of clusters in which sequences are segregated and RTs at cluster boundaries that are proportional to the number of items within each cluster (De Lillo and Lesk, 2010).

The neural correlates of the processing of organizational factors in SSR have been highlighted in an f-MRI study (Bor et al., 2003) where participants faced an array of items arranged as a square matrix and were presented with “structured” and “unstructured” sequences. “Structured” sequences were operationally defined as those sequences where consecutive items were within the same row, column or diagonal. “Unstructured” sequences were defined as those violating this constraint (i.e., with consecutive items never within the same row, column, or diagonal). Behavioral results confirmed that structured sequences were reported at a higher level of accuracy than unstructured sequences. Moreover, f-MRI data indicated a higher activation of the dorsolateral prefrontal cortex (DLPFC) during the encoding of structured sequences than during the encoding of unstructured sequences.

Other important factors have been reported to affect the reproduction of spatial sequences. These include the length of the path of the trajectory necessary to connect all the items in the sequence and the number of times the path crosses itself (Orsini et al., 2001, 2004; Parmentier et al., 2005, 2006). These factors sometimes confound the effects of the structure of the representation underpinning performance. For example, it has been proposed that the clustering effect as observed by De Lillo (2004) could be explained by the fact that clustered sequences can have on average a shorter path length than non-clustered sequences (Parmentier et al., 2006). Similarly, the effect of structure observed by Bor et al. (2003) could be due to the fact that unstructured sequences can contain more crossings.

The RT patterns reported for structured sequences (Bor et al., 2003; De Lillo, 2004; De Lillo and Lesk, 2010) and the fMRI results of Bor et al. (2003) suggest that the detection and use of structure in Corsi sequences determines the formation of specific forms of hierarchical representation that contribute to efficient recall quite apart from other effects of path characteristics. Nevertheless, considering the possible contribution of all these factors, it is important to assess their relative role in SSR. We attempted to do so with the present study. The approach we took in the first experiment was to dissociate path-length and organization in a clustered array. With this experiment we

tested the notion proposed by Parmentier et al. (2006) that path-length can be the sole explanation of the clustering effect in spatial span. We manipulated display size so that clustered sequences in a large display had a longer path-length than non-clustered sequences in a small display. Thus, if the benefits of clustering are explained by the shorter path that is normally associated with clustered sequences, then we should expect a more accurate recall for the non-clustered sequences with a short path when compared with the recall for structured sequences with a longer path. In the second experiment, we manipulated the timing structure of the sequence leaving its path-length and any other characteristics of the sequences unchanged. Using the same clustered sequences, we imposed pauses in the sequence presentation either at transitions between items within a cluster or at cluster boundary. An effect of timing in this experiment would indicate that the clustering effect is more likely to be related to the way in which the sequence is represented, rather than to mere effects of path characteristics, such as path-length or number of path-crossings.

In a third experiment we used a square matrix of locations which allowed a fully factorial manipulation of path length, presence of crossings, and structure as defined by Bor et al. (2003). By doing so we aimed to disentangle the effects of path length, presence of crossings, and structure. We then determined which of these factors explained most of the variance in the recall score of the participants.

The aim of the fourth and fifth experiments was to evaluate the importance of perceptual grouping for the emergence of beneficial effects of structure in SSR. In fact, the use of terms such as perceptual grouping, perceptual organization and gestalt principles is so widespread in the literature in relation to the explanation of the benefits of organizational factors in SSR and so often used interchangeably with that of efficient memory coding (e.g., Kemps, 1999; Bor et al., 2003; Rossi-Arnaud et al., 2005; Ridgeway, 2006; Bor, 2012; Hurlstone et al., 2014) to warrant an explicit assessment of the extent to which perceptual grouping is actually required for the benefits of organization in SSR to emerge.

In Experiments 4 and 5 we used immersive virtual reality to implement a navigational version of the Corsi task. In this task the order in which the sequence items had to be reproduced could not be apprehended from the same viewpoint. Having identified the first item in the sequence participants were required to move toward it and select it. Only then would the next item be presented at a different location within the environment. The presentation of the sequence was a lengthy process that involved continuous changes of directions and viewpoints. We reasoned that the observation of beneficial effects of structure in these conditions would have made the hypothesis that perceptual grouping processes are necessary for their emergence implausible.

EXPERIMENT 1

In SSR, path length refers to the length of the trajectory that connects the items that need to be reported in the

prescribed order. It can be manipulated independently from other characteristics of the spatial sequences by altering the size of the item display so that the relative distance between the items is different in the two displays (Smyth and Scholey, 1994). The critical variable in this experiment was the relative distance of the items in the small and the large display. The experiment was designed so that clustered sequences, presented in the large set, had a longer path-length than non-clustered sequences presented in the small set. If the short path that typically accompanies clustered sequences is the sole explanation of the clustering effect, as proposed by Parmentier et al. (2006), then non-clustered sequences with a shorter path should be recalled more accurately than clustered sequences with a longer path.

Methods

Participants

Twenty five volunteers (10 male and 15 females), with a mean age of 26 years ($SD = 7.76$), were recruited from a participant panel at the University of Leicester and paid a small fee to take part in the experiment. They all reported normal or corrected-to-normal vision.

Materials, Design, and Procedure

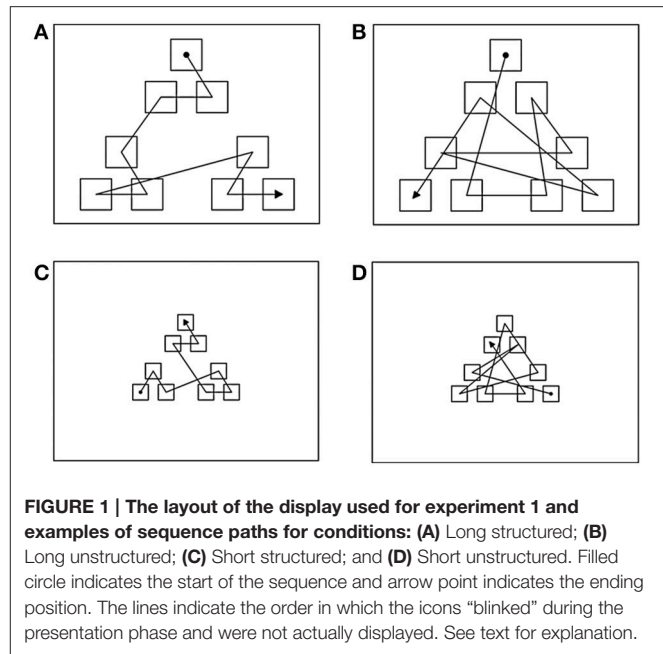
The experiment was presented using a PC equipped with a 17" Elo, IntelliTouch sensitive monitor (1024×768 pixels). A visual display consisting of a black background and nine identical gray squares arranged in three spatial clusters of three icons each was presented on each trial. Two displays were used: a large display, composed of squares 120 pixels wide (**Figures 1A,B**); and a small display, composed of squares 40 pixels wide (**Figures 1C,D**). In the small display, there was a 6 pixel-wide invisible active border area surrounding each square to ensure that the touch of a square was accurately registered even by participants with larger finger tips.

In the large display, squares within a cluster were separated by a distance between 76 and 103 pixels, whereas the distance between clusters was of 152–189 pixels. In the small display, squares within a cluster had a distance between 19 and 27 pixels and the distance between clusters was of 38–53 pixels.

In each trial, participants were first presented with the full display of 9 squares for 700 ms. One square then turned to black for 500 ms, before the full display was represented for another 700 ms. This produced the impression that the icon would “blink.” Another square would then turn black for 500 ms, and so on until a sequence of 9 items was presented in this way. After the ninth square had blinked, the screen turned black for 1 s. The full display was then presented again and participants were required to reproduce the sequence that they had previously observed. To confirm that the touch had been registered, each square turned to black for 50 ms when touched.

The design featured the manipulation of sequence type that could be either “Structured” or “Unstructured” and path-length that could be “Long” or “Short.”

“Structured” sequences were segregated by spatial clusters, so that all the items of each cluster were presented before the sequence moved to a different cluster. By contrast “Unstructured” sequences were not segregated by clusters, so that consecutive



items were always presented in different clusters. “Long” sequences were displayed in the large display and “Short” sequences were presented in the small display. The average path-length of long sequences was of 3093.36 pixels ($SE = 170.37$) and that of short sequences was 883.69 pixels ($SE = 50.14$).

Four experimental conditions were obtained combining these two factors in a 2 (structure) \times 2 (path-length) repeated measures design: Long Structured (L-S), Long Unstructured (L-U), Short Structured (S-S), and Short Unstructured (S-U). Importantly, the path-length of L-S sequences (mean = 2381.32; $SE = 40.09$) was significantly longer than that of S-U sequences (mean = 1096.32; $SE = 17.42$). Examples of each sequence type are also provided in **Figure 1**. Ten sequences of each type were used and presented in random order within a testing session of 40 trials.

Participants were tested in a quiet laboratory with dim lighting. The height of their chair was adjusted so their eyes were at the same level as the center of the screen and they could comfortably touch any point of the display with the index finger of their dominant hand. Participants were informed that they had to use that finger when selecting the squares during the experiment, which took about 15 min to complete.

This experiment and all the other experiments reported in this article were carried out in accordance with the Code of Ethics and Conduct of the British Psychological Society and approved by the University of Leicester Ethics Committee for research involving human participants (Psychology sub-committee). All subjects gave written informed consent in accordance with the Declaration of Helsinki.

Results

Accuracy

An item recalled correctly was defined as a square touched in the correct serial position. Accuracy scores were the frequency

of items correctly recalled by each participant in each condition. The mean accuracy score of each of the four conditions is presented in **Figure 2A**.

A 2 (structure: structured/unstructured) \times 2 (path-length: long/short) repeated measures ANOVA was carried out on the frequency of correct items reported in the different conditions. It revealed a significant main effect for structure [$F_{(1, 24)} = 8.747$, $p < 0.01$, $\eta_p^2 = 0.267$] with a higher level of recall for structured sequences and path-length [$F_{(1, 24)} = 198.965$, $p < 0.001$, $\eta_p^2 = 0.892$], with a higher level of recall for long sequences. No interaction between path-length and structure was found.

Paired sample *t*-tests with Bonferroni correction (alpha of 0.05 corrected to 0.01 and alpha of 0.01 corrected to 0.002) were carried out to further clarify these results. Long structured sequences produced a significantly higher level of accuracy on recall than short unstructured sequences [$t_{(24)} = 12.404$, $p < 0.01$], demonstrating that structured sequences were recalled at a higher level of accuracy than unstructured sequences even when they had a longer path-length. Moreover, the effect of structure was very robust as it was maintained in both short [$t_{(24)} = 11.681$, $p < 0.01$] and long sequences [$t_{(24)} = 11.208$, $p < 0.01$]. The effect of path-length proved less robust as it did not emerge when sequences with long and short path-length were

compared within the structured and unstructured conditions separately.

Serial Position Analysis

Serial position effects were observed in each condition, as can be observed for the serial position curves presented in **Figure 2B**. A 2 (path-length: long/short) \times 2 (structure: structured/unstructured) \times 9 (serial position: 1/2/3/4/5/6/7/8/9) ANOVA for repeated measures was carried out. A significant main effect emerged for path-length [$F_{(1, 24)} = 8.747$, $p < 0.01$, $\eta_p^2 = 0.267$], structure [$F_{(1, 24)} = 198.965$, $p < 0.001$, $\eta_p^2 = 0.892$], and serial position [$F_{(8, 192)} = 26.326$, $p < 0.001$, $\eta_p^2 = 0.523$], confirming the main results reported above. A particularly strong interaction emerged between structure and serial position [$F_{(8, 192)} = 13.082$, $p < 0.001$, $\eta_p^2 = 0.345$]. As can be observed from **Figure 2B** this can be easily accounted for by the different shape of the curves of the structured sequences on the one hand, and the unstructured sequences on the other. The latter curves resemble typical serial position curves for unstructured material with some indications of possible primacy and recency effects (see Crowder, 1969). Such effects are absent in the structured sequences. A significant interaction between path-length and serial position was also found, [$F_{(8, 192)} = 5.510$, $p < 0.01$, $\eta_p^2 = 0.114$]. This was not as conspicuous as the interaction between serial position and cluster type. It is likely to be explained by small differences occurring at different serial positions, which are more difficult to pinpoint. The third order interaction between path-length, structure and serial position was not significant.

Discussion

In this experiment we observed beneficial effects of clustering similar to those observed in other studies (De Lillo, 2004; De Lillo and Lesk, 2010). We found that clustering had a beneficial effect in both sequences with a long and with a short path. This suggests that path-length alone is unlikely to explain the effects of clustering in SSR. It has previously been suggested that clustered sequences afford a hierarchical coding of the sequence with spatial clusters forming the superordinate level and the items within each cluster forming the subordinate level (e.g., De Lillo, 2004; De Lillo and Lesk, 2010). However, Experiment 1 was not designed to assess this possibility. An attempt at gaining a better insight on the type of memory coding supported by clustering in this study was made in Experiment 2 by manipulating the temporal pattern of the presentation of clustered sequences.

EXPERIMENT 2

Experiment 2 aimed to provide additional support for the independence of the effects of structure and path-length in SSR and some indication concerning the nature of the representation underlying sequences segregated by spatial clusters. In order to do so, we used an approach previously used in the study of chunking and hierarchical representation in recall in the spatial (Bor et al., 2003) and other domains (Farrell and Lelievre, 2012). The temporal structure of the presentation of the sequences was manipulated by inserting temporal pauses

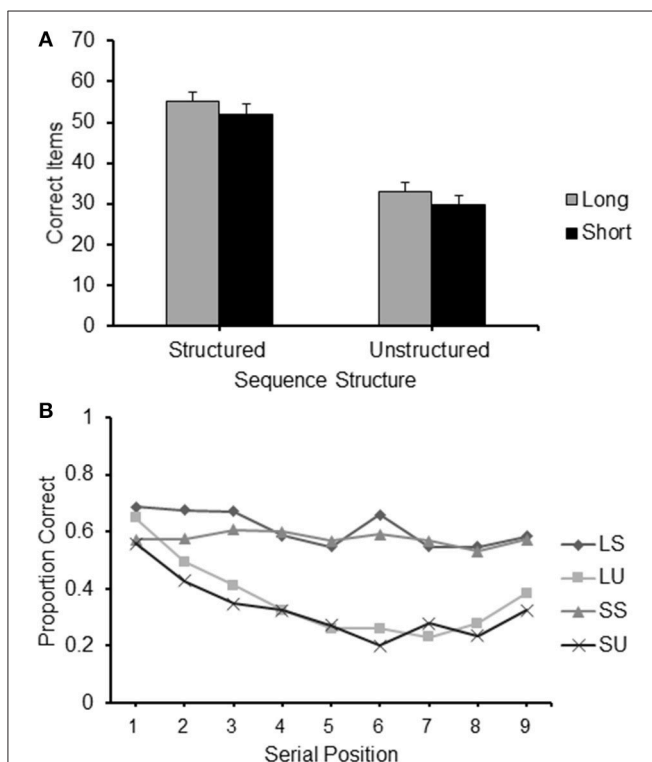


FIGURE 2 | (A) Frequency of items correct recalled for the different conditions (Long Structured; Short Structured; Long Unstructured and Short Unstructured) of Experiment 1; **(B)** Proportion of correct items recalled at each serial position for the four different conditions of Experiment 1: LS, Long Structured; LU, Long Unstructured; SS, Short Structured; SU, Short Unstructured.

during the presentation of clustered sequences. Only clustered sequences were used in this experiment. For some sequences the pause was inserted at transitions between items within a cluster. For other sequences, the pause was inserted at transitions between clusters. As such, sequences could be either consistent or inconsistent with a hierarchical representation based on segmentation by spatial clusters. Other path characteristics of the sequences remained the same so that effects of the temporal structure of the sequence could not be confounded with other effects of path characteristics.

Methods

Participants

Twenty five undergraduate psychology students (23 females and 2 males, age range 18–25 years) from the University of Leicester took part in this experiment as part of their course requirement. All participants reported normal or corrected-to-normal vision.

Materials, Design, and Procedure

The same PC used in Experiment 1 was used here. The static display was the same as the large display of Experiment 1. The presentation of sequences followed the same general procedure of Experiment 1, with the following exception. In Experiment 2, a 3 s pause was introduced at critical points during the presentation of the sequence according to different conditions as described below. To discourage the participants from fixating the last item that blinked before the pause, thus minimizing any effects of the temporal pauses, during the pauses all the items disappeared (turned the same color of the background). In the Between Cluster Pause (BCP) condition, the pause was inserted when the sequence reached a cluster boundary and a transition between clusters was required. In the Within Clusters Pause (WCP) condition, the pause was inserted at a transition between items in the same cluster (see **Figure 3** for a visual representation of a BCP and a WCP sequence, illustrating this procedure).

Eighteen different sequences were used. Each sequence was displayed twice, once as part of the BCP condition and once as part of the WCP condition. Two pauses were presented in each sequence of both conditions.

Thus, a total of 36 sequences were presented to each participant in random order. As exactly the same sequences were presented for the BCP and of the WCP condition, we ensured that the path length and any other path characteristic were the same in the sequences of both conditions.

Results

The mean frequency of items correctly recalled for sequences with BCP and WCP are presented in **Figure 4A**. It can be observed from there that BCP sequences were recalled more accurately than WCP sequences.

A paired-sample *t*-test confirmed that this difference was significant [$t_{(24)} = 2.727, p < 0.05$].

Serial Position Analysis

A 2 (pause location: between/within) \times 9 (serial position: 1/2/3/4/5/6/7/8/9) repeated measure ANOVA carried out on the proportion of items correctly recalled confirmed an effect of

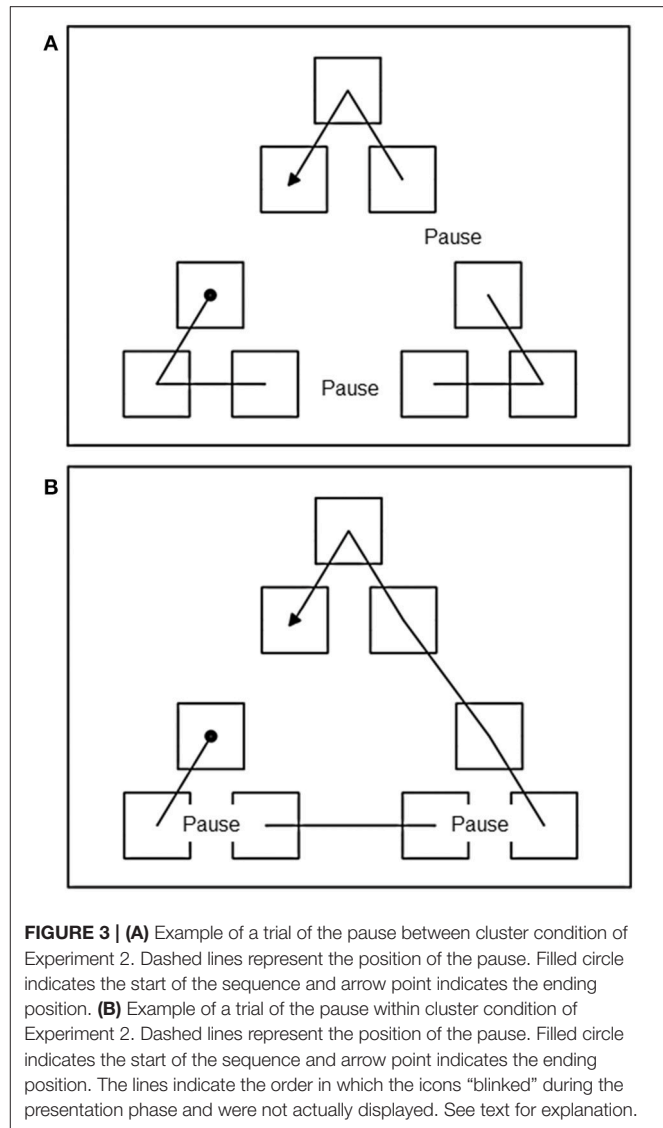


FIGURE 3 | (A) Example of a trial of the pause between cluster condition of Experiment 2. Dashed lines represent the position of the pause. Filled circle indicates the start of the sequence and arrow point indicates the ending position. **(B)** Example of a trial of the pause within cluster condition of Experiment 2. Dashed lines represent the position of the pause. Filled circle indicates the start of the sequence and arrow point indicates the ending position. The lines indicate the order in which the icons “blinked” during the presentation phase and were not actually displayed. See text for explanation.

pause location [$F_{(1, 24)} = 7.334, p < 0.05, \eta_p^2 = 0.234$] and revealed a significant effect of serial position [$F_{(8, 192)} = 5.007, p < 0.001, \eta_p^2 = 0.173$] (see **Figure 4B**).

Discussion

Experiment 2 clarifies the possible causes of the beneficial effects of spatial clustering in SSR. The pattern of results of this experiment is consistent with the pattern that is usually considered evidence for chunking in other domains. During the presentation of the to-be-recalled material, pauses at the predicted locations of chunking boundaries (which were compatible with the formation of a hierarchical representation based on grouping by spatial proximity) produced a higher level of recall accuracy than pauses imposed within the predicted chunk boundaries (which presumably hindered the hierarchical organization of the sequence based on spatial proximity). Moreover, in this experiment exactly the same sequences, and as such the same sequence paths, were used for BCP and for

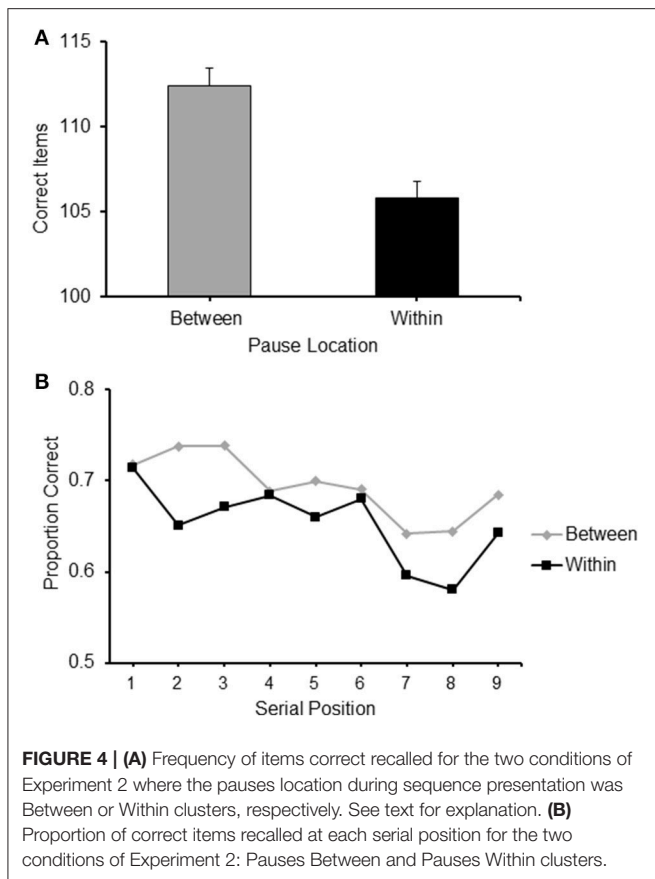


FIGURE 4 | (A) Frequency of items correct recalled for the two conditions of Experiment 2 where the pauses location during sequence presentation was Between or Within clusters, respectively. See text for explanation. **(B)** Proportion of correct items recalled at each serial position for the two conditions of Experiment 2: Pauses Between and Pauses Within clusters.

the WCP. Only the temporal pattern of the presentation of the sequences was manipulated. Therefore, the effects observed here cannot be accounted for by factors related to the path characteristics of the sequence such as path-length or number of crossings featured in a given sequence (Orsini et al., 2001, 2004; Parmentier et al., 2005; Parmentier and Andrés, 2006).

Taken together, the results of the first two experiments support the notion that participants use spatio-temporal structure in immediate serial recall tasks to form representations, such as a hierarchical representation of the sequence based on spatial clusters, to enhance recall (De Lillo, 2004; De Lillo and Lesk, 2010).

In Experiments 1 and 2 we used a clustered arrangement of items. The use of such an arrangement is important for several reasons. It provides a spatial analogy of class formation in other domains. It conforms to the most intuitive hierarchical organization of space, which is in regions and sub-regions. Moreover, the study of recall of clustered locations is particularly relevant in relation to the potential role played by the pressure to search efficiently patchy resources in the evolution of sophisticated memory skills in humans and other primates as advocated by foraging theories of primate cognitive evolution (Milton, 1993).

Nevertheless, the use of spatially clustered items does not offer the flexibility necessary for the manipulation of path-length and number of path-crossings in sequences which are

either structured or unstructured, in a fully nested factorial design. Therefore, in order to explore further the relative role of effects of structure and effects of different path characteristics, in Experiment 3 we used a large matrix of locations which allowed such experimental design.

EXPERIMENT 3

Sequence structure (Bor et al., 2003; De Lillo, 2004; De Lillo and Lesk, 2010; Fagot and De Lillo, 2011) and the path characteristics of sequences have been shown to affect SSR (Kemps, 2001; Orsini et al., 2001, 2004; Parmentier et al., 2005; Parmentier and Andrés, 2006; Fagot and De Lillo, 2011). Among path characteristics, the presence of path-crossings in the sequence has proved an important variable that can substantially affect recall. Path-crossings refers to the occasions where the imaginary line traced by connecting all the items in the sequence in the prescribed order, crosses itself. Whereas, effects of path-length on SSR have failed to emerge in some studies on human participants (Smyth and Scholey, 1994; Fagot and De Lillo, 2011), negative effects of the presence of path-crossings have emerged consistently in studies where this variable has been manipulated (Orsini et al., 2001, 2004; Parmentier et al., 2005; Parmentier and Andrés, 2006; Fagot and De Lillo, 2011). Path-length and path-crossings are often confounded because sequences containing more crossings will on average be longer than sequences with less crossings. Thus, some studies have used the strategy of keeping sequence length constant when assessing the effect of path crossing (Parmentier and Andrés, 2006). However, the relative contribution of path-length, path-crossings, and any other possible residual effects of structure that cannot be explained by these two path characteristics has not been evaluated yet in a single factorial experiment. We aimed to do so in experiment 3. Clustered arrays, such as those used in Experiments 1 and 2 offered us the opportunity to study effects of hierarchical organization based on spatial proximity on SSR. Another form of organization that has been reported to have a beneficial effect on SSR is operationally defined using arrays of items arranged as a 4×4 square matrix of locations (Bor et al., 2003). Structured sequences, defined as those where consecutive items are within the same row, column or diagonal are recalled with a higher level of accuracy than sequences that violate this rule. These beneficial effects of structure have been considered to be related to chunking (Bor et al., 2003). However, here too effects of structure, path-length and crossings can be confounded. A systematic assessment of the relative role of structure and path characteristics within these arrays has not been attempted yet, possibly also because a 4×4 grid is not large enough to allow the generation of sequences that would allow the systematic manipulation of all these factors in the same experiment. In Experiment 3 we attempted to identify the contribution of each of these factors using a larger 5×5 matrix of items (see also Kemps, 2001; Rossi-Arnaud et al., 2012; Cestari et al., 2013) that provided the flexibility for generating a sufficient number of sequences for a fully nested 2 (structure) \times 2 (path-length) \times 2 (crossings) factorial design.

Methods

Participants

Twenty seven (20 female and 7 male, age range 19–27) undergraduate psychology students from the University of Leicester took part in Experiment 3 and received course credits for their participation. All participants reported normal or corrected to normal vision.

Materials, Design, and Procedure

We used the same PC as in the previous two experiments. Software developed in-house allowed the presentation of 25 identical white squares arranged as a 5×5 matrix (see **Figure 5**). Each square had a side of 116 pixels (2.5 cm).

In each trial, participants were first presented with the 5×5 matrix of squares for 700 ms, before the sequence presentation started and the matrix of items remained present throughout the presentation and the recall phase of each trial. In this experiment the sequences contained 7 items. The timing of the item presentation was the same as in the first two experiments. A within-participant design was used. The independent variable was the type of sequence, defined according to its path characteristics (henceforth referred to as “sequence type”), with three factors: structure (structured/unstructured), path-length (long/short), and crossings (with/without crossings). Eight experimental conditions were obtained by nesting the two levels of each of the above factors: Structured Long with Crossings (SLC); Structured Short with Crossings (SSC); Structured Long No-crossings (SLN); Structured Short no Crossings (SSN);

Unstructured Long with Crossings (ULC); Unstructured Short with Crossings (USC); Unstructured Long No crossings (ULN); and Unstructured Short No Crossings (USN). Examples of different types of sequences are presented in **Figure 5**.

An operational definition of the factors determining each type of sequence is provided below.

Structure

This was defined following Bor et al. (2003): “Structured sequences” had consecutive items within the same row, column or diagonal of the matrix; “Unstructured sequences” systematically violated these constraints (i.e., consecutive items were never within the same row, column, or diagonal).

Path-length

This was defined in terms of the total number of squares intersected by transitions between consecutive items in the sequence. Long sequences intersected a total of at least 10 items. Short sequences intersected a maximum of 6 squares.

Path-crossings

This variable referred to transitions in the sequence crossing over an imagined line of the previously completed sequence that the participants gaze or finger would travel through, as defined by Kemps (2001). Sequences without crossings did not contain any instance of the above and sequences with crossings contained at least three such crossings. The number of crossings in the sequence was chosen on the basis of previous research

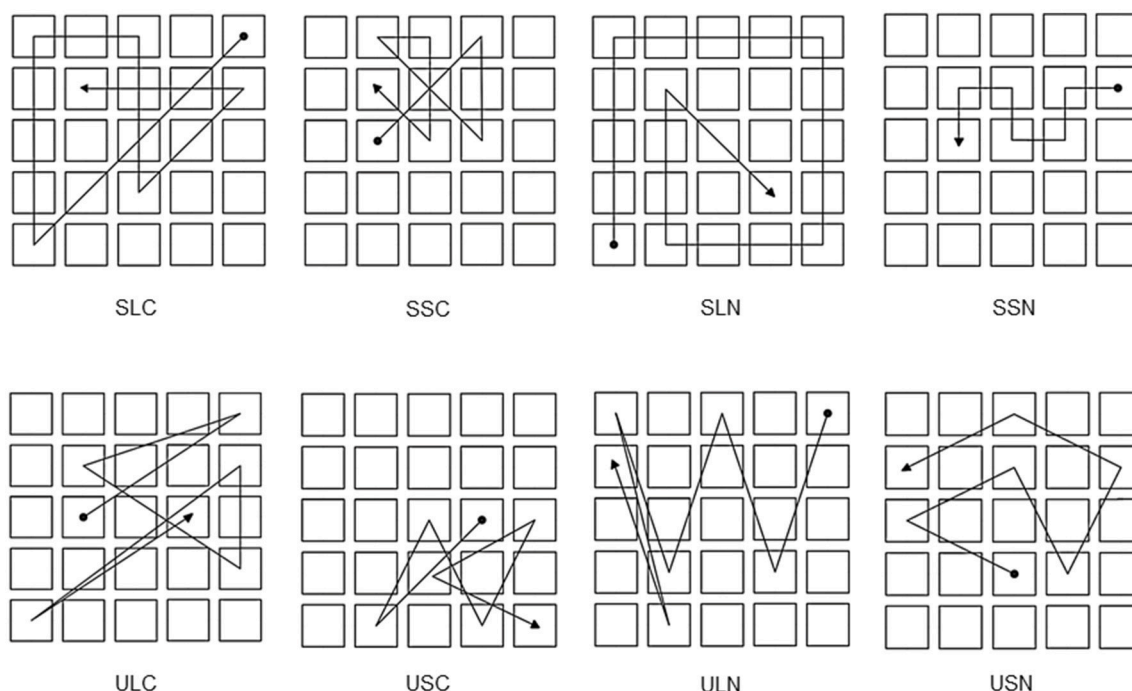


FIGURE 5 | Example of trials for the eight experimental conditions administered in Experiment 3: SLC, Structured Long with Crossings; SSC, Structured Short with Crossings; SLN, Structured Long No-crossings; SSN, Structured Short with No-crossings; ULC, Unstructured Long with Crossings; USC, Unstructured Short with Crossings; ULN, Unstructured Long No-crossings; and USN, Unstructured Short No-crossings.

that reported sequences presented with three crossings having a significant negative effect on recall compared to sequences with no crossings. This effect did not continue linearly when more crossings were included (Parmentier et al., 2005).

The testing session comprised a total of 64 trials featuring 8 sequences for each of the 8 conditions, interspersed in random order.

Results

The main effects of the different conditions featured in Experiment 3 are presented in **Figure 6**. From the figure it can be observed that recall accuracy was better for sequences which were structured, or with a short path or without crossings.

A 2 (structure: structured/unstructured) \times 2 (crossings: with/without crossings) \times 2 (path-length: long/short) repeated measures ANOVA carried out on the frequency of items correctly recalled in the different conditions revealed significant main effects for all the three factors: structure [$F_{(1, 26)} = 193.20$, $p < 0.001$, $\eta_p^2 = 0.881$], with structured sequences recalled with a higher level of accuracy than non-structured sequences (see **Figure 6A**); crossings [$F_{(1, 26)} = 132.60$, $p < 0.0005$, $\eta_p^2 = 0.836$] with sequences without crossings recalled at a higher level of accuracy than sequences with crossing (see **Figure 6B**); and path-length [$F_{(1, 26)} = 54.69$, $p < 0.001$, $\eta_p^2 = 0.678$], with short sequences reported at a higher level of accuracy than long sequences (see **Figure 6C**).

The only significant second order interaction was between the factors structure and crossings [$F_{(1, 26)} = 18.807$, $p < 0.01$, $\eta_p^2 = 0.42$]. However, t -tests comparing accuracy for sequences

with and without crossings, with the two levels of the factor path-length combined, all revealed highly significant differences [Structured-No crossings vs. Structured-with-Crossings, $t_{(26)} = 9.503$, $p < 0.01$; Unstructured-No crossings vs. Unstructured-with-Crossings, $t_{(26)} = 6.25$, $p < 0.001$]. Therefore, this interaction cannot be explained by the lack of an effect of crossings at some level of the factor structure. Instead, the interaction is likely to be due to a slight variation of the effects of crossings at the two levels of sequence structure. In fact, whereas the difference in the accuracy of recall in the sequence with and without crossings was of 21.15 in the structured condition, it was of only 9.37 in the unstructured condition. The accuracy values for the sequences with crossings and without crossing for structured and unstructured sequences are reported in **Table 1**. A significant third order interaction also emerged [$F_{(1, 26)} = 4.745$, $p < 0.05$, $\eta_p^2 = 0.154$]. Planned comparisons (t -tests with Bonferroni correction, $\alpha = 0.0042$) of the effects of each of the factors with each level of the other factors kept constant, revealed all significant effects [$4.21 < t_{(26)} < 12.03$, $p < 0.05$] with the exception of the comparison between sequences with and without crossing in the unstructured short sequences [$t_{(26)} = 2.59$, $p = \text{n.s.}$] and the difference between long and short path-length in the unstructured sequences with no crossings [$t_{(26)} = 1.05$, $p = \text{n.s.}$].

Eta Squared

In order to illustrate the strength of the experimental effect of each of the three factors, in **Figure 6D**, we report the value of the η^2 which indicates the proportion of variance that is explained by

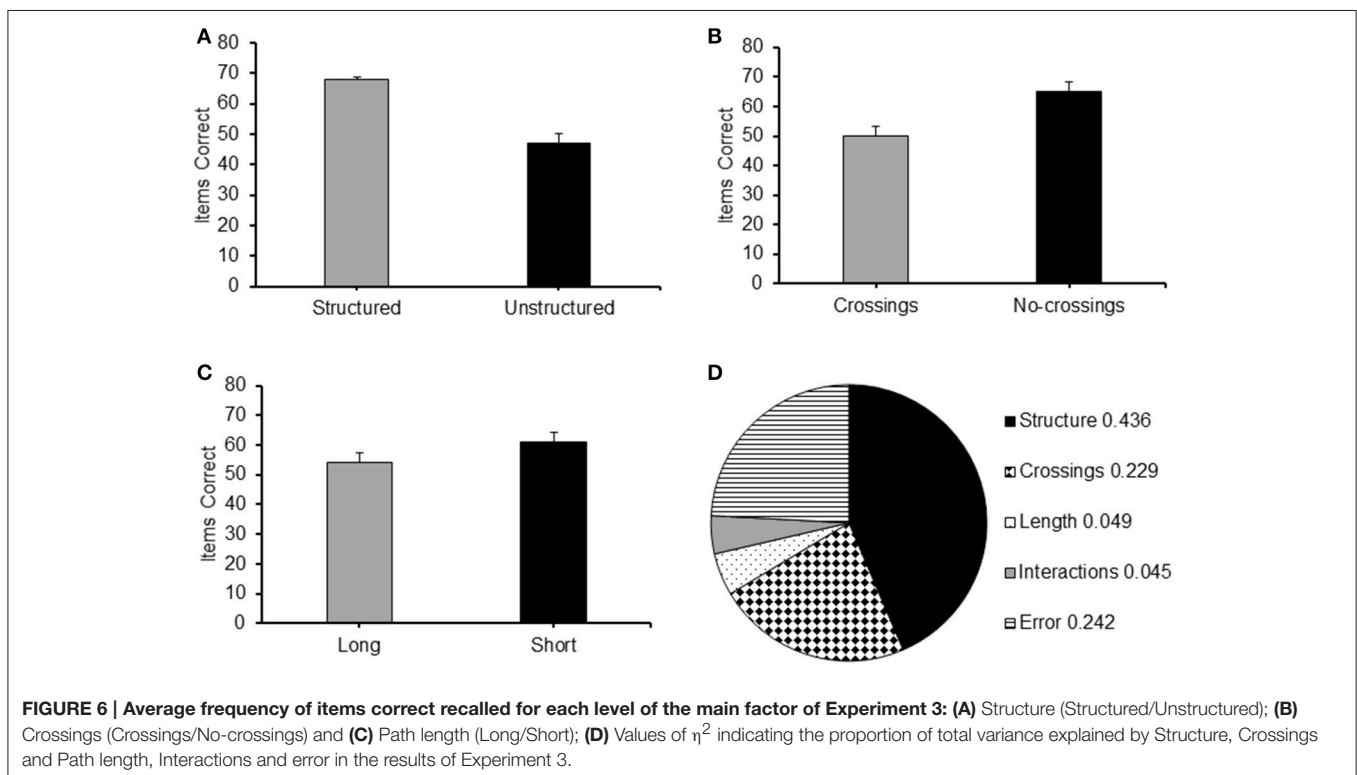


TABLE 1 | Experiment 3: participants mean frequency of correct items and Standard Deviation for on each type of sequence.

	Structured		Unstructured	
	Crossings	No crossings	Crossings	No crossings
Long	55.56 (SD 21.12)	74.11 (SD 18.75)	37.67 (SD 12.72)	51.03 (SD 19.37)
Short	61.63 (SD 18.60)	83.37 (SD 18.97)	47.19 (SD 18.47)	52.56 (SD 16.78)

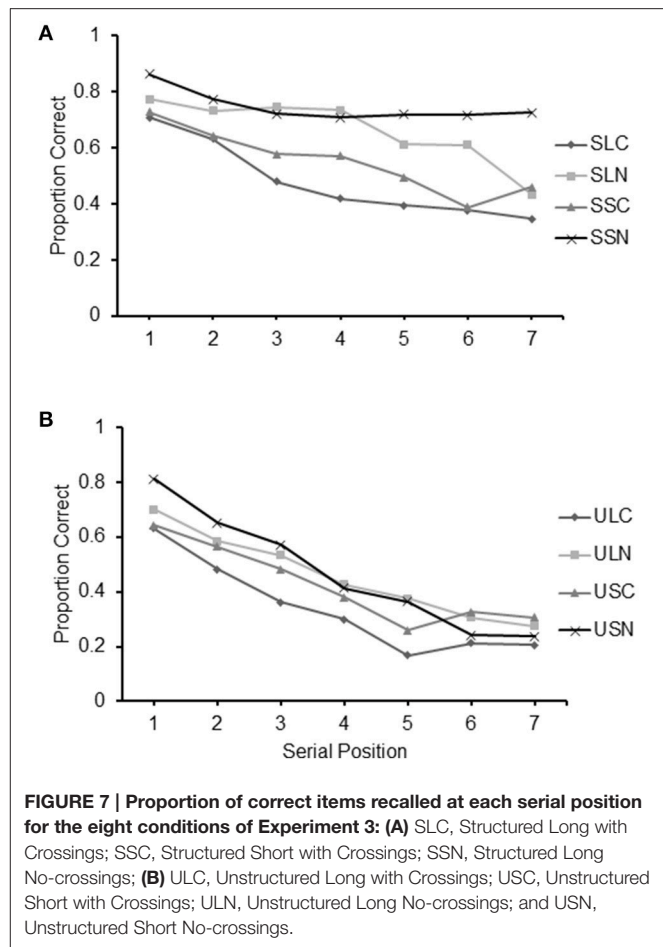
each of them (see Howell, 2002). As can be seen from the figure, path-crossings explained a larger portion of variance than path-length. Importantly, the largest portion of variance is explained by residual effects of structure that cannot be explained by either path-crossings or path-length.

Serial Position Analysis

Serial position curves for all conditions are reported in Figure 7. For clarity the curves are presented separately for conditions featuring structured sequences (Figure 7A) and those featuring unstructured sequences (Figure 7B). An ocular inspection of the figure indicates the presence of serial position effects in all conditions. Albeit present in all curves, a decreasing level of recall in relation to the serial position of the items is particularly evident in conditions featuring unstructured sequences.

A 7 (serial position: 1/2/3/4/5/6/7) \times 2 (structure: structured/unstructured) \times 2 (path-length: long/short) \times 2 (crossings: with/without crossings) repeated measures ANOVA confirmed the presence of a main effect of serial position [$F_{(6, 156)} = 86.81, p < 0.001, \eta_p^2 = 0.770$] in addition to the effects of structure [$F_{(1, 26)} = 193.33, p < 0.001, \eta_p^2 = 0.881$], path-length [$F_{(1, 26)} = 54.87, p < 0.001, \eta_p^2 = 0.678$], and crossings [$F_{(1, 26)} = 132.99, p < 0.001, \eta_p^2 = 0.836$]. The interactions Structure by Crossings [$F_{(1, 26)} = 18.87, p < 0.001, \eta_p^2 = 0.421$], Structure by Path-length by Crossings [$F_{(1, 26)} = 4.83, p < 0.05, \eta_p^2 = 0.157$], Structure by Serial position [$F_{(6, 156)} = 19.85, p < 0.05, \eta_p^2 = 0.443$], Path-length by Serial position [$F_{(6, 156)} = 3.37, p < 0.05, \eta_p^2 = 0.115$], Structure by Path-length by Serial position [$F_{(6, 156)} = 6.42, p < 0.001, \eta_p^2 = 0.198$], Crossings by Serial position [$F_{(6, 156)} = 6.34, p < 0.001, \eta_p^2 = 0.196$], Structure by Crossings by Serial position [$F_{(6, 156)} = 6.18, p < 0.001, \eta_p^2 = 0.192$] were all significant.

Although it is difficult to pinpoint the exact nature of these complex interactions, they are likely to be due to various small variations occurring at different points of the serial position curve of the different conditions that otherwise showed a relatively similar pattern. In fact, trend analyses carried out individually for the different conditions revealed that they all had a significant linear component [$9.95 < F_{(6, 156)} < 165.39$, all $ps < 0.001$, $2.77 < \eta_p^2 < 0.864$] that could be indicative of a primacy effect as performance deteriorated in line with the serial position of the items in the sequence. The quadratic component of the trend was



highly significant in conditions SLC, SLN, SSN, ULC [$16.28 < F_{(1, 26)} < 45.87, 0.374 < \eta_p^2 < 0.481$] and USC and significant in condition USN [$F_{(1, 26)} = 6.79, p < 0.05, \eta_p^2 = 0.205$]. It was not significant in conditions SSC and ULN.

The graphs suggest the lack of a recency effect, apart from increases in recall between items 6 and 7 in condition SSC, $t_{(26)} = 2.92, p < 0.01$ and between items 5 and 6 in conditions USN, $t_{(26)} = 4.43, p < 0.001$ and USC, $t_{(26)} = 2.64, p < 0.05$, that may be indicative of such an effect.

Discussion

In Experiment 3 we carried out a systematic manipulation of sequence structure as previously defined in the literature (Bor et al., 2003) and two path characteristics that have been shown to affect SSR (Kemps, 2001; Orsini et al., 2001, 2004; Parmentier et al., 2005; Parmentier and Andrés, 2006; Fagot and De Lillo, 2011). The effects of structure and path characteristics can be confounded (see Parmentier et al., 2005, 2006; De Lillo and Lesk, 2010). Thus, the aim of Experiment 3 was to evaluate whether or not the effects of path characteristics can explain in their entirety purported effects of structure. The results indicate that they cannot. The experiment confirmed the presence of the effects of path characteristics reported in the literature, namely, path-length and path-crossings (Parmentier

et al., 2005, 2006). However, the results suggest that beneficial effects related to structure played a role on top of effects related to path characteristics. Moreover, the effects of structure emerged as the one explaining most of the variance in the data, followed by effects of path-crossings and finally by path-length effects.

As outlined above, the literature on organizational factors in SSR, often refers to the possible role of perceptual grouping factors in relation to the beneficial effects of structure (Kemps, 1999; Bor et al., 2003; Rossi-Arnaud et al., 2005; Ridgeway, 2006; Bor, 2012; Hurlstone et al., 2014). Perceptual grouping effects occurring at the time of observing the to-be-reproduced sequence could possibly take place in SSR tasks affording an aerial view of items presented in a rapid sequence. In situations where participants are required to slowly navigate through the items to apprehend the sequence to be reproduced the occurrence of perceptual grouping is much less likely. In Experiment 4 we aimed to assess if effects of structure emerge in navigational tasks too.

EXPERIMENT 4

In Experiment 4, we used a clustered configuration similar to the one used in Experiments 1 and 2 and a 3×3 matrix of locations. This allowed a direct comparison of the effects of structure and path-length with the two configurations and different operational definitions of structure. Importantly, however, SSR was assessed as part of an immersive virtual reality navigational task that required changes of viewpoint and slow movements by the participant to apprehend consecutive items in the sequence (see Supplementary Video). Visual perceptual grouping normally occurs between elements that are presented as part of the same visual display viewed from a fixed viewpoint. Albeit recent

evidence suggest a different timescale for the processing of different stages of grouping, visual grouping processes seem to be completed within a time range spanning a few hundred milliseconds to a second (Kurylo, 1997; Han et al., 1999, 2002; Brick Larkin and Kurylo, 2013). Thus, we reasoned, if effects of structure in SSR emerge when the to-be-recalled sequence is presented as a part of a large scale navigation task, lasting over a minute and requiring continuous changes of viewpoint, then it is unlikely that they are caused by perceptual grouping processes.

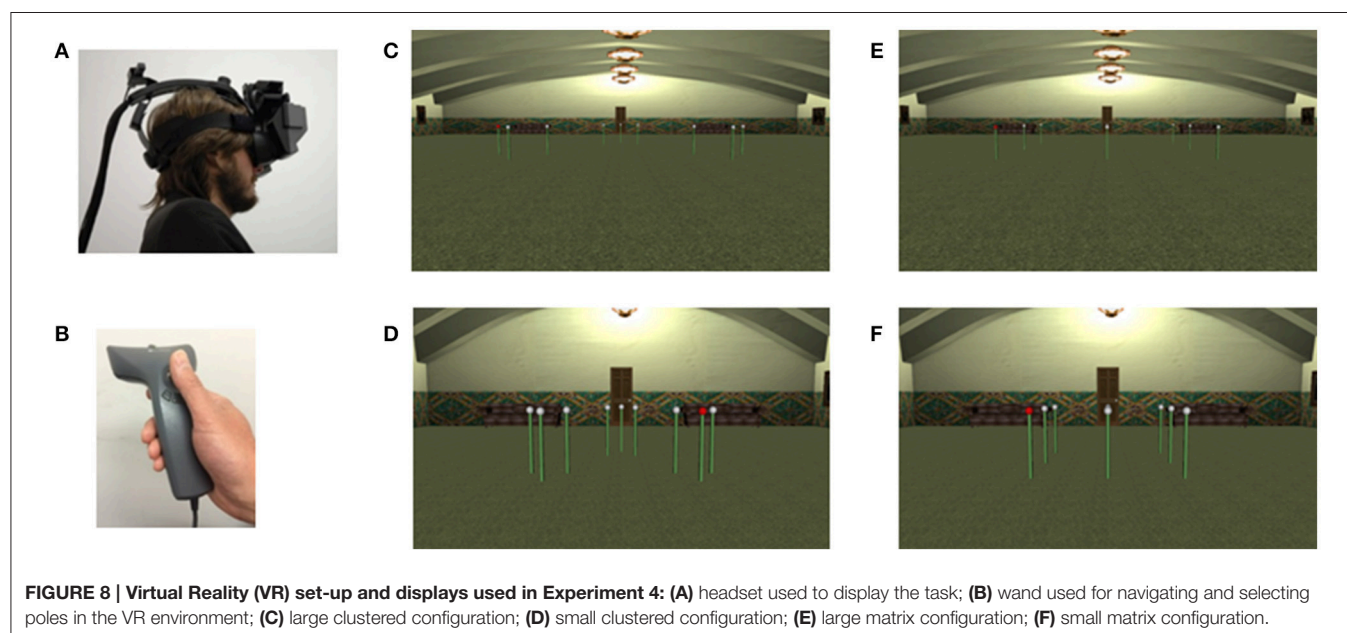
Methods

Participants

Twenty one participants (10 females and 11 males, age range 18–42) took part in the experiment. They were psychology undergraduates, who received course credits for participation, or members of a participant panel, comprising mainly of post-graduate students and staff, who were paid a small fee for taking part.

Materials, Design, and Procedure

The experiment took place in a Virtual Reality laboratory equipped with an NVIS nVisor stereoscopic head mounted display (see **Figure 8A**). An Inter-Sense position tracker determined the viewpoint depending on the head and the body movement of the participants who operated a hand held wand to navigate and produce responses. Traveling in the virtual environment was controlled by moving with the thumb a small joystick located on the wand (see **Figure 8B**). Traveling speed was set to 1 meter per second. The software was developed in house using Vizard 3.0. (WorldViz) and allowed the presentation of a virtual environment consisting of a set of 9 poles surmounted by a white sphere within a large virtual hall with richly textured surfaces and several landmarks (see **Figures 8C–F**).



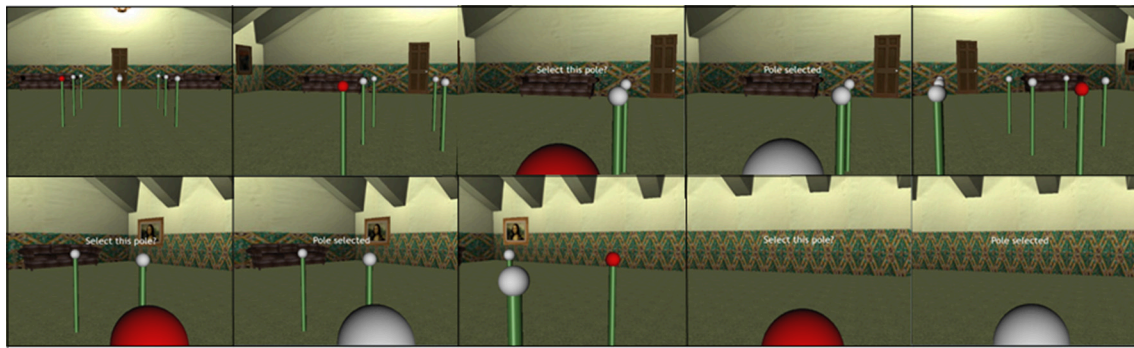


FIGURE 9 | Examples of viewpoints experienced during the presentation phase of a trial of Experiment 4. The figure shows that approaching the required pole, selecting it, it and then turning within the virtual reality environment to determine the location of the next pole in the to-be-recalled sequence required dramatic changes of view point. This made it impossible for visual perceptual grouping to contribute to effects of structure in this implementation of a serial recall task.

The general serial recall procedure was implemented as described below. Each trial featured a presentation phase and a recall phase. In the presentation phase, one of the white spheres surmounting the poles turned red until the participant traveled through the environment, approached and selected it by operating the wand. As soon as the pole was selected it turned white again and a second pole within the array turned red, and so on, until all the 9 poles had been selected by the participant. Thus, the selection of consecutive items during the presentation of the sequence required the participants to search and navigate toward different points of the environment. Since traveling time was set at 1 meter per second and it could take time for participants to rotate their head and identify the next item to reach, the presentation of the sequence was a lengthy process sometimes lasting one or more minutes to complete. Following the completion of the presentation phase, the recall phase immediately ensued. All poles turned white. The starting position was reinstated and the participant had to recall the sequence by navigating through the virtual environment and selecting the poles in the same order as in the presentation phase. The head movement of the participant was tracked by the inter-sense system and used to update the view-point producing a vivid immersive experience, and importantly determining that the viewpoint changed continuously (see **Figure 9**, and Supplementary Video, for an example of viewpoint experienced during the presentation of the sequence), depending on the position of the head and the body of the participant throughout the presentation of the sequence.

The design comprised a within subjects manipulation of the configuration of poles (clustered or matrix), the structure of the sequence (structured or unstructured), and the distance between the poles path-length (long and short) of the sequence to be recalled. The configurations and display sizes are shown in **Figures 8C–F**.

Configuration and path-length were manipulated by changing the spatial arrangement and the distance of the poles, respectively. Depending on the configuration condition the poles were arranged as a 3×3 square matrix or a clustered configuration. The path-length was manipulated by making the

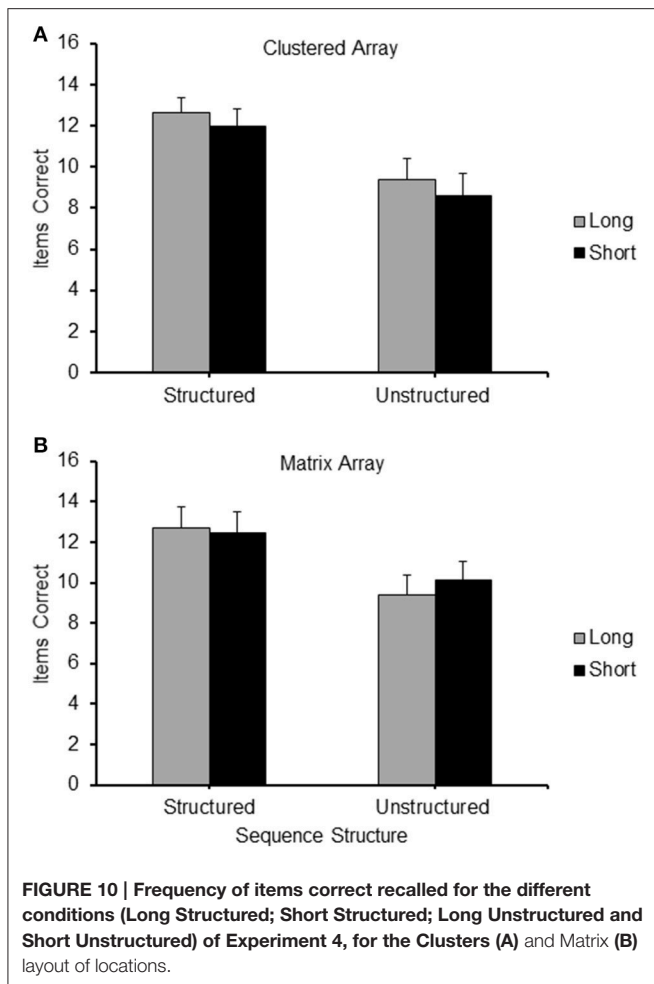
inter-pole distance in the long path condition 3 times longer than the distance in the short path condition. In particular, the minimum possible distance between the poles was 2.1 m in the long path condition and 0.7 m in the short path condition. The starting point from the center of the configuration of poles was 14.7 and 4.9 m in the short-path condition. For the clustered configuration, the structure of the sequence was manipulated as in Experiment 1. For the matrix configuration, it was manipulated as in Experiment 3. Participants received alternating trials of the short-path and the long path condition, with the starting position randomized across participants. Apart from this constraint, the conditions were randomized across trials. Each participant received two trials per condition for a total of 16 trials. Because of the weight of the headset, to ensure comfort, participants were given a short break every two trials, when required, and a 10 min break after eight trials.

Results

The mean frequency of items correctly recalled in the different conditions of Experiment 4 is depicted in **Figure 10A** (Clustered configuration) and **Figure 10B** (Matrix configuration).

A 2 (Configuration: clusters/matrix) $\times 2$ (Path-length: long/short) $\times 2$ (Structure: structured/unstructured) repeated measure ANOVA carried out on the frequency of items correctly recalled in the different conditions revealed a highly significant main effect for structure [$F_{(1, 20)} = 49.86, p < 0.001$] and a significant main effect of configuration [$F_{(1, 20)} = 6.94, p < 0.05$], where structured sequences were recalled significantly more accurately than unstructured sequences, and sequences in the matrix configuration were recalled slightly more accurately than the clustered configuration. By contrast, there was no significant difference in recall performance between long and short path-length [$F_{(1, 20)} = 0.021, p = \text{n.s.}$]. None of the interactions proved significant.

Similarly to the analysis of the results of Experiment 1, we carried out the critical comparison between the Long Structured and the Short Unstructured conditions in using paired sample *t*-tests on the mean frequency of correct items observed in these two conditions. The means for these conditions are



reported as part of **Figure 10**. The observed level of accuracy was higher in the long structured condition compared to the short unstructured condition in both the clustered [$t_{(20)} = 2.44, p < 0.05$] and the matrix configuration [$t_{(20)} = 4.04, p = 0.001$].

Because each condition only featured two trials in Experiment 4, the data set generated for this experiment was not suitable for serial position analyses.

Discussion

The results of Experiment 4 confirm the main findings of Experiment 1. Moreover, they show that participants can benefit from effects of structure even when they do not have a bird's eye view of the configuration of items. Importantly, the view point of the participants changed continuously during the task, as a function of their position within the virtual environment and mostly afforded seeing only a portion of the configuration of poles. Furthermore, each presentation phase lasted at least 1 min. Thus, the task characteristics, both in terms of viewpoint and timescale, make it unlikely that visual perceptual grouping mechanisms underpin the benefits of structure in this task. Additionally, the results of Experiment 4 confirmed that path-length has a marginal role in serial recall, even in conditions that should exacerbate its role, as in navigation. This pattern

of results seems to be very robust and confirms the results of other experiments carried out with a similar procedure but with a different and smaller sample of participants (De Lillo et al., 2014).

Experiment 4 did not address specifically the role of path-crossings in situations where participants do not have a bird's eye view of the configuration of test items. This was because Experiment 4 already featured several variables and because it is not possible to systematically vary path-crossings and distance at the same time in the clustered configuration. Therefore, the effects of this variable in relation to traveling distance was addressed in Experiment 5 by focusing exclusively on the matrix configuration and using in this VR navigational environment a design similar to that used in Experiment 3.

EXPERIMENT 5

Methods

Participants

Twenty participants (10 female and 10 male, with a mean age of 19.90, $SD = 2.56$) undergraduate psychology students from the University of Leicester took part in Experiment 5 and received course credits for their participation. All participants reported normal or corrected to normal vision.

Materials, Design, and Procedure

The apparatus used was the same as described for Experiment 4. The VR environment was similar to that used for Experiment 4 and the administration of the trials followed the same procedure. However, the configuration of poles in the VR environment and the design of the experiment were modeled on those of Experiment 3. Thus, Experiment 5 featured a 5×5 matrix of items and the same eight experimental conditions used for Experiment 3: Structured Long with Crossings (SLC); Structured Short with Crossings (SSC); Structured Long No-crossings (SLN); Structured Short No-crossings (SSN); Unstructured Long with Crossings (ULC); Unstructured Short with Crossings (USC); Unstructured Long No-crossings (ULN); and Unstructured Short No-crossings (USN). Examples of different types of sequences are presented in **Figure 5**. Because of the length of the testing session and the weight of the head mounted display, only two sequences for each of the conditions was presented to the participants, as in Experiment 4. The sequences were pseudo-randomly selected from the pool of sequences for each participant and each presentation with the constraint that the first eight trials had to feature one of each of the conditions. After the presentation of the first eight trials the participants were given the opportunity to take a short break before being presented with the second trial of all the conditions.

Results

As for Experiment 3, a 2 (Structure: structured/unstructured) \times 2 (Crossings: with/without crossings) \times 2 (Path-length: long/short) repeated measures ANOVA carried out on the frequency of items correctly recalled in the different condition revealed significant main effects for all the three factors: structure [$F_{(1, 19)} = 11.66, p < 0.01, \eta_p^2 = 0.380$], with structured sequences recalled with a higher level of accuracy than non-structured sequences (see

Figure 11A); crossings [$F_{(1, 19)} = 14.73, p = 0.001, \eta_p^2 = 0.437$] with sequences without crossings recalled at a higher level of accuracy than sequences with crossings (see **Figure 11B**); and path-length [$F_{(1, 19)} = 7.69, p < 0.05, \eta_p^2 = 0.288$], with longer sequences reported at a higher level of accuracy than shorter sequences (see **Figure 11C**). None of the interactions proved significant.

Eta Squared

As for Experiment 3, η^2 were calculated in order to evaluate the proportion of variance explained by the different factors. Albeit the values of η^2 for the main effects were smaller than those observed in Experiment 3, they followed the same pattern. Structure accounted for most of the total variance ($\eta^2 = 0.084$), followed by crossings ($\eta^2 = 0.066$), and finally path-length ($\eta^2 = 0.034$). The values of η^2 obtained for the error and the interactions were $\eta^2 = 0.760$ (error) and $\eta^2 = 0.055$ (interactions), respectively.

As only 2 trials were collected for each condition, it was not feasible to carry out a serial position analysis for this experiment.

Discussion

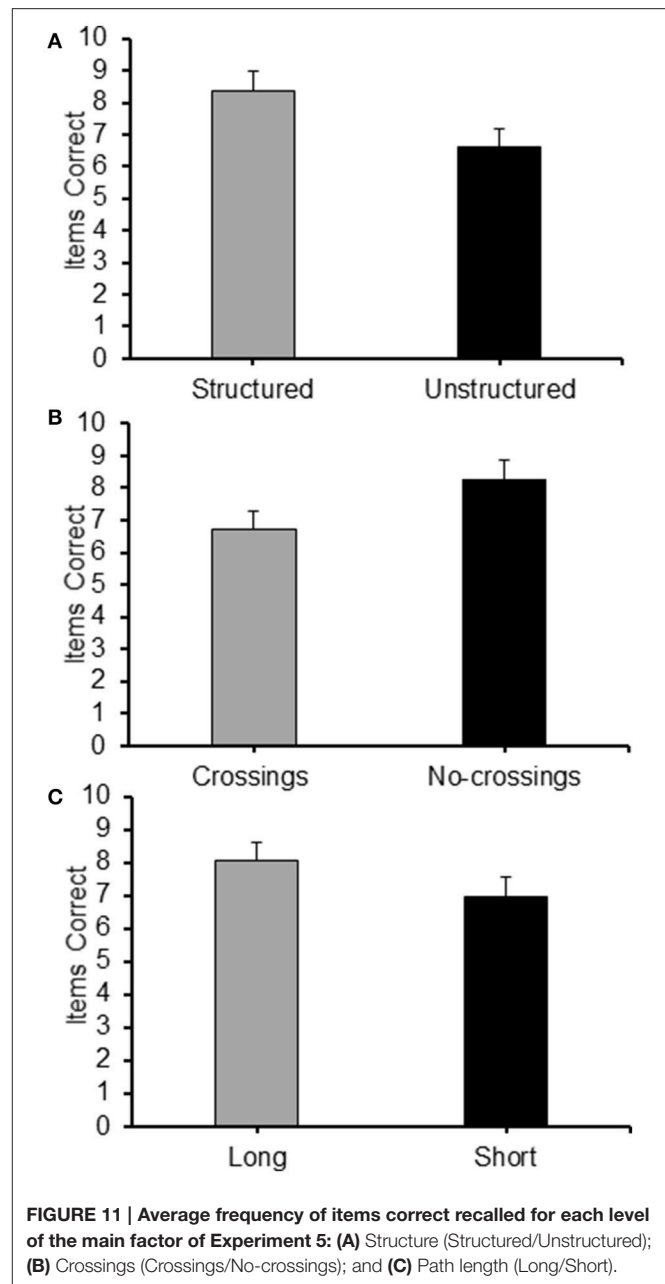
The results of Experiment 5 confirmed those obtained in Experiment 3. They indicate that the effects of structure and path characteristics are very robust and emerge in conditions where the viewpoint of the participants changed continuously during the presentation of the sequence and during recall. For these effects to emerge it is not necessary to have a bird's eye view of the display. This finding is particularly important in relation to the effects of structure. It confirmed that the effect of structure cannot be entirely explained by path characteristics alone and also that it can emerge when the operation of perceptual grouping principles is unlikely.

GENERAL DISCUSSION

We have reported five experiments exploring the contribution of different factors to immediate SSR. We assessed effects of path-length, path crossing and visual perceptual organization and determined the presence of residual effects of stimulus structure which are beneficial to recall but cannot be attributed to these factors.¹

Path-length and path-crossings have been previously shown to affect spatial span as measured by the standard Corsi test (Orsini et al., 2001, 2004) and the related dot task (Parmentier et al., 2005; Parmentier and Andrés, 2006; Guerard and Tremblay, 2012). Because these tests of immediate spatial recall use irregular arrangements of items it is difficult to determine with them which structural affordances of the display participants may detect and use to facilitate the recall of given sequences. Hence, the need to use spatial layouts of items specifically designed to allow, or

¹The set of experiments reported here focused on accuracy rather than response time (RT) and RT analyses were outside the scope of the present study. For detailed analyses of timing in relation to the composition of clustered sequences see De Lillo and Lesk (2010). For traveling time analyses in virtual reality tasks on related topics see Logie et al. (2011) and Trawley et al. (2011).



prevent, forms of spatiotemporal organization which are known a priori to afford memory coding of a particular kind.

For these reasons, our first two experiments used a Corsi-type task featuring a configuration of identical icons arranged as spatial clusters defined by the relative proximity of the items (De Lillo, 2004; Parmentier et al., 2006; De Lillo and Lesk, 2010). Within this type of clustered configuration, sequences that are segregated by clusters afford hierarchical memory coding (De Lillo, 2004; De Lillo and Lesk, 2010). However, effects of spatial clustering there can be confounded by path-length. In fact, it has been suggested that path-length alone could explain improved recall in such displays (Parmentier et al., 2006).

The results of Experiment 1 showed a dissociation between path-length and structure defined as segregation by spatial clusters. Structure proved an important determinant of spatial recall performance. By manipulating path-length in large and small arrays of spatial items (see also Smyth and Scholey, 1994) we showed that there are effects of segregation by clusters which are independent from, and stronger than, path-length. In fact, in Experiment 1 structured sequences were recalled at a higher level of accuracy than non-structured sequences irrespectively of path-length.

Although this finding is consistent with the notion that the segregation of sequences by spatial clusters affords hierarchical memory representations of the sequences that facilitates recall (De Lillo, 2004), in Experiment 1 it was not possible to test this possibility directly. Moreover, because most sequences segregated by clusters have fewer crossings than unstructured sequences, it was not possible to determine in Experiment 1 whether or not there are effects of segregation by clusters which can be attributed to hierarchical coding, in addition to possible beneficial effects deriving from the presence of fewer crossings in these sequences.

In Experiment 2 we controlled for the effects of crossings and tested the hierarchical coding hypothesis of the benefits of clustering in SSR, by exclusively using sequences segregated by clusters and manipulating the temporal phrasing of sequence presentation. Critically, Experiment 2 provided independent evidence that recall in sequences segregated by spatial clusters is supported by hierarchical coding. A lower level of recall was observed in conditions featuring the insertion of pauses within cluster boundaries and which were not compatible with a retrieval process that exploited such organization. In this respect, our results conform to those of studies that manipulated the presentation of pauses within and between hypothesized chunks to infer the presence of hierarchical coding in the spatial (Bor et al., 2003) and other domains (Farrell and Lelievre, 2012).

Thus, taken together the results of the first two experiments provide synergistic support for the notion that the beneficial effects of segregating sequences by spatial clusters derives from hierarchical coding in spatial working memory. The results are also consistent with evidence of hierarchical organization in SSR obtained in studies where the size and number of spatial clusters were manipulated in conditions otherwise similar to those of the present study (De Lillo and Lesk, 2010). There emerged that RT at cluster boundaries is proportional to the number of items in the cluster, reflecting the time taken to access the order of report of the items within that cluster (subordinate level of the hierarchy). By contrast, a component of the initial RT, which was deemed to be an expression of the time taken to retrieve the order of report of the clusters into which the sequence was segregated, was proportional to the number of clusters (superordinate level of the hierarchy). Albeit consistent with those of previous studies featuring clustered Corsi arrays (De Lillo, 2004; De Lillo and Lesk, 2010), the results of Experiment 1 and 2 provide crucial information concerning whether or not path-length could account on its own for the improved recall observed in sequences segregated by spatial clusters. The results show that this is not the case.

In clustered arrays such as those used in the first two experiments (as well as in irregular arrays, see Ridgeway, 2006) it is difficult to systematically manipulate the presence of path-crossings and path-length independently. Therefore, in our Experiment 3 we used a matrix of locations where it was possible to do so. We constructed sequences with nested characteristics so that a fully factorial design could be used to assess the portion of variance in the data explained by each of these factors. The results confirmed the presence of effects of all these factors with residual positive effects of structure. Structure was defined as in other studies that have used a square matrix of locations to administer a Corsi-type tasks (Bor et al., 2003). The presence vs. absence of structure in the sequences produced the largest portion of variance in recall. This suggests that principles other than path-length and crossings are most important in determining ease of recall in serial spatial memory.

Compared to spatially clustered arrays, matrices of locations make it more difficult to infer the reason for the benefits of structure defined as having consecutive serial spatial items within the same column, row or diagonal in a square matrix of locations. Bor et al. (2003) proposed that the benefits are derived from chunking and are based on gestalt principles. Similar claims have been made in other studies where gestalt principles such as continuity and symmetry were explicitly applied to the sequences (Kemps, 2001; Rossi-Arnaud et al., 2005). Reference to the formation of gestalten that facilitate SSR raises the question of whether perceptual grouping principles need to operate during the presentation of the sequences for this to occur.

The position of several authors converge on the notion that perceptual processes may account for the formation of chunks at encoding (Kemps, 2001; Bor et al., 2003; Parmentier et al., 2005; Avons, 2007; Bor, 2012). When reviewing these studies Bor (2012) queried whether the process of organizing the to-be-remembered material occurs “on-the-fly” that is, “spotting the pattern as a powerful new rule to apply on each trial” (Bor, 2012, p. 179). On the basis of the fact that an activation of the fusiform gyrus occurs in addition to the activation of the DLPFC in trials featuring structured sequences, Bor et al. (2003) suggested that chunking in this domain may be related to the object perception functions of this cortical region. A study based on self-reports of participants presented with variations of the Corsi test has attempted to dissociate what the author refers to as perceptual grouping and strategic grouping in the assessment of spatial span (Ridgeway, 2006). The study concluded that both perceptual and strategic grouping occur and give rise to enhanced memory performance in this task. This conclusion was based on the fact that some participants reported the use of strategies consisting of dividing the sequence to be reported in subgroups even in sequences where a clear-cut separation between some of its sub-sequences was not present. This is a potentially important distinction. Nevertheless, so far the relative contribution of each of these two types of grouping to serial recall accuracy has not been clarified.

On the basis of these considerations we deemed it important to determine whether or not perceptual grouping is essential for the benefits of structuring in spatial working memory to occur. We addressed this issue in our last two experiments by assessing

the effects of structure and path characteristics in conditions that made perceptual grouping unlikely to occur.

The immersive virtual reality serial recall task used in Experiments 4 and 5 was designed so that participants would navigate within the configuration of items forming the display. As a navigational version of the Corsi test, the VR task developed here is similar to walking Corsi tasks where participants observe a model walking along a route connecting locations in a real-life environment that they are required subsequently to reproduce (Piccardi et al., 2008, 2013; Nemmi et al., 2013). The walking Corsi has proved a very useful diagnostic tool for the detection of specific topographical memory deficits and sex differences in spatial memory (Piccardi et al., 2013) but has not been used to assess the role of organizational factors in WM. Our study differs from those featuring the walking Corsi as we used structured arrays of locations where different types of sequences were systematically manipulated rather than irregular arrays and sequences resembling those used in the traditional Corsi test.

Moreover, the virtual reality task used here differs in important aspects to the walking Corsi. During the observation of the sequences in the walking Corsi the observer stands at a distance from the array of locations traveled by the model. Thus, the entire sequence is seen from a single vantage point outside the configuration that the participant will be required to explore. Finally, the set of locations featured in the walking Corsi test are marked on the floor, making the test very similar to the standard Corsi and very useful for the comparison of spatial span in reaching and navigational space (Nemmi et al., 2013). In contrast with the walking Corsi, in our tasks during the presentation of the sequences participants were required to approach each item in the sequence and select it. Only then, a cue concerning the location of the next item was displayed in the navigational space. Thus, the entire sequence could never be observed from a single vantage point. The use of a head mounted display with head trackers ensured that the viewpoint of the participants changed at every movement the head performed to scan the array to search for the item to approach. This, together with the virtual movement throughout the array of items and the timing of the selection of each item, made it unlikely that participants could rely on visual grouping to form chunks. Yet, in Experiment 4 we still observed facilitating effects of structure in both the clustered and the matrix arrangement of items and in sequences characterized by either a short or long path. Moreover, even in a situation which made longer movements particularly costly in terms of time and distance traveled, path-length could not explain on its own the benefit of structure. In fact, as in Experiment 1, participants recalled more accurately structured long path sequences than unstructured short path sequences.

In Experiment 5 we used a design similar to that of Experiment 3 and simultaneously manipulated structure, path-length, and crossings in a VR environment. It emerged again that the effect of structure can be dissociated from the effect of the other two factors and explains a larger portion of variance than crossings and path-length. Path-length had a small effect and counterintuitively sequences with a longer path-length were recalled slightly more accurately than sequences with shorter paths. This indicates once again that beneficial effects of structure

cannot always be reduced to the shortening of path-length often associated with structured sequences. A possible reason for the small advantage for sequences with longer path-length in this experiment is that a longer distance, on average, between consecutive items makes it easier for the participants to encode items as pertaining to distinctive sub-regions of space in the environment.

Interestingly, it has been proposed that path-length effects in spatial recall tasks presented on computer monitors that afford an aerial view of the display could be related to perceptual grouping processes. Longer path-length would hinder grouping principles that otherwise strengthen the coding of the transition between successive items of sequences and result in less accurate recall (Guerard and Tremblay, 2012). The lack of a positive effect of shorter path-length in our navigational serial recall tasks used to prevent grouping could be consistent with this theoretical interpretation of path-length effects.

Most importantly, taken together, the results of Experiments 4 and 5 indicate that visual perceptual grouping processes do not need to occur for the benefit of structure in serial recall to emerge. Thus, the encoding of structure is likely to have occurred at a post perceptual stage of processing. Our results are neutral to the issues of whether these effects of structure occur during encoding, rehearsal or recall. Yet there is evidence for the fact that they are likely to occur at encoding. For example, in the study by Bor et al. (2003) the selective of the dorsolateral prefrontal cortex associated with effects of structure was observed during the encoding phase only. Also, effects of path-length (Guerard and Tremblay, 2012) and crossings have been shown to be unaffected by manipulations of the amount of rehearsal and by concurrent tapping performed during rehearsal (Parmentier and Andrés, 2006).

It has been proposed that effects of structure in SSR obtained by imposing symmetry, good continuation, and repetition of parts of the sequence in translated positions are to be attributed to the participation of long-term memory and that this could occur at recall (see Kemps, 2001; Rossi-Arnaud et al., 2005, for a discussion of this point). This could be possible. However, it has been shown (De Lillo and Lesk, 2010) that with serial recall tasks presented on touch-screens the effect of clustering and structure occur also in tasks which do not require the reproduction of the sequence at recall (i.e., participants are required to judge if two sequences are the same or not). Therefore, we consider it unlikely that the effects of structure are built at that time or that they are related to the motor plan of the sequence (see also Farrell and Lelievre, 2012, for similar conclusions in relation to temporal grouping).

One likely possibility is that during the encoding phase people start to build a mental image of the search space and the path followed and that the benefits of grouping occur at that stage and can persist during rehearsal. The possible role of mental images in spatial serial recall in the Corsi test makes intuitive sense and has been envisaged by several researchers (see Berch et al., 1998; Guerard and Tremblay, 2012). It is possible that the DLPFC activity that accompanies the benefits of structure in spatial recall (Bor et al., 2003) is related to the activation of mental images of visual patterns

conforming to familiar shapes or gestalten. This would be consistent with studies that indicate an involvement of the dorsal part of the prefrontal cortex in tasks requiring mentally imagining visual patterns (Ishai et al., 2000; Mechelli et al., 2004). It is known that grouping can occur in mental images. The extent to which the advantage for structure is mediated by static representations of mentally constructed visual or dynamic shifts of attention rehearsing the pattern remains to be determined.

In summary, there is some inconsistency in the literature in relation to the role of coding strategies, path characteristics and perceptual grouping in explaining beneficial effects of organization. Concepts related to one or the other of these factors are often used interchangeably or considered confounded. The current study clarifies that neither path-characteristics nor perceptual grouping principles can account for all effects of organization in SSR. Importantly, path-length throughout this series of experiments emerged to be the least important factor in explaining the beneficial effects of structure. This is of great interest because path-length is perhaps the path characteristic that is least likely to be related with the efficient encoding of data. Yet, it is the factor that most often is confounded with effects of organization. We have shown that even in navigational tasks where the cost of traveling is exacerbated and produces a large discrepancy in the time taken to observe the sequence at encoding, shorter paths are not always associated with better recall. Interestingly, comparative studies of humans and monkeys on serial recall tasks in small scale tests presented on touch-screens, where path-length and structure were dissociated showed that whereas humans are more sensitive to structure, monkeys are more sensitive to path-length (Fagot and De Lillo, 2011). This may be indicative of a peculiarity of human higher level cognition consisting in an enhanced ability to pick-up structure and use it to efficiently encode information, which could have implications for both the understanding of the evolutionary origins of human

cognition and for the refinement of primate models of human memory.

The short-term retention of spatiotemporal structures is ubiquitous in behavior. Yet the cognitive processes supporting it are still poorly understood and comparatively under-investigated in experimental psychology. We believe that disentangling the effects of the efficient use of data structure from effects of path-characteristics in serial recall should open-up new lines of essential research in working memory. These should be aimed at characterizing further the processes responsible for it in different task domains, the extent to which they dissociate in neuropsychological conditions, during lifespan and in animal models of human cognition.

AUTHOR CONTRIBUTIONS

CD conceived the study, analyzed the data, and wrote the article. MK collected and analyzed data for Experiment 4 and contributed to the write-up of the article. DP collected data of Experiment 3, contributed to the analyses and write-up of that experiment and provided comments on drafts of the article.

ACKNOWLEDGMENTS

We thank Tony Andrews for programming experiment 1-3, Kevin McCracken for programming Experiments 4-5, Josephine Marson for helping collecting and analyzing data for experiments 1 and 2, Chelsey Nightingale for helping collecting data for Experiment 5 and Robin Green for useful comments on a draft of this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.01686/full#supplementary-material>

REFERENCES

- Avons, S. E. (2007). Spatial span under translation: a study of reference frames. *Memory Cogn.* 35, 402–417. doi: 10.3758/BF03193281
- Avons, S. E., and Oswald, K. (2008). Using temporal order to identify spatial reference frames. *Percept. Psychophys.* 70, 1068–1080. doi: 10.3758/PP.70.6.1068
- Baddeley, A. (1992). Working memory. *Science* 255, 556–559.
- Baddeley, A. D. (2001). Is working memory still working? *Am. Psychol.* 56, 849–864. doi: 10.1037/0003-066X.56.11.851
- Berch, D. B., Krikorian, R., and Huha, E. M. (1998). The Corsi block-tapping task: methodological and theoretical considerations. *Brain Cogn.* 38, 317–338. doi: 10.1006/brcg.1998.1039
- Boduroglu, A., and Shah, P. (2014). Configural representations in spatial working memory. *Visual Cogn.* 22, 102–124. doi: 10.1080/13506285.2013.875499
- Bor, D. (2012). *The Ravenous Brain*. Philadelphia, PA: Perseus; Basic Books.
- Bor, D., Duncan, J., Wiseman, R. J., and Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron* 37, 361–367. doi: 10.1016/S0896-6273(02)01171-6
- Botvinick, M. M., Wang, J., Cowan, E., Roy, S., Bastianen, C., Mayo, J. P., et al. (2009). An analysis of immediate serial recall performance in a macaque. *Anim. Cogn.* 12, 671–678. doi: 10.1007/s10071-009-0226-z
- Brick Larkin, G., and Kurylo, D. D. (2013). Perceptual grouping and high-order cognitive ability. *J. Indiv. Dif.* 34, 153–158. doi: 10.1027/1614-0001/a000110
- Busch, R. M., Farrell, K., Lisdahl-Medina, K., and Krikorian, R. (2005). Corsi block-tapping task performance as a function of path configuration. *J. Clin. Exp. Neuropsychol.* 27, 127–134. doi: 10.1080/138033990513681
- Cambridge Cognition (2006). *CANTABclipse Test Administration Guide / Manual version 3.0.0*. Cambridge: Cambridge Cognition Limited.
- Cestari, V., Saraulli, D., Spataro, P., Lega, A., Sciarretta, A., Marques, V. R., et al. (2013). Memory for symmetry and perceptual binding in patients with schizophrenia. *Acta Psychol.* 144, 594–603. doi: 10.1016/j.actpsy.2013.09.016
- Cornoldi, C., and Vecchi, T. (2003). *Visuo-Spatial Working Memory and Individual Differences*. Hove: Psychology Press.
- Corsi, P. M. (1972). *Human Memory and the Medial Temporal Region of the Brain*. Dissertation, Abstracts International, 34(02), 891B (University microfilms No. AA105-77717).
- Crowder, R. G. (1969). Behavioral strategies in immediate memory. *J. Verbal Learn. Verbal Behav.* 8, 524–528. doi: 10.1016/S0022-5371(69)80098-8
- De Lillo, C. (2004). Imposing structure on a Corsi-type task: evidence for hierarchical organization based on spatial proximity in serial-spatial memory. *Brain Cogn.* 55, 415–426. doi: 10.1016/j.bandc.2004.02.071
- De Lillo, C. (2012). “The use of spatial structure in working memory: a comparative standpoint,” in *The Complex Mind: an Interdisciplinary Approach*,

- eds D. McFarland, K. Stenning, and M. McGonigle-Chalmers (Chippenhams: Eastbourne, Palgrave-MacMillan), 38–54.
- De Lillo, C., Kirby, M., and James, F. C. (2014). Spatial working memory in immersive virtual reality foraging: path organisation, travelling distance and search efficiency in humans (*Homo sapiens*). *Am. J. Primatol.* 76, 436–446. doi: 10.1002/ajp.22195
- De Lillo, C., and Lesk, V. (2010). Spatial clustering and hierarchical coding in immediate serial recall. *Eur. J. Cogn. Psychol.* 22, 216–246. doi: 10.1080/09541440902757918
- Fagot, J., and De Lillo, C. (2011). A comparative study of working memory: immediate serial recall in baboons (*Papio papio*) and humans. *Neuropsychologia* 49, 3870–3880. doi: 10.1016/j.neuropsychologia.2011.10.003
- Farrell, S., and Lelièvre, A. (2012). The dynamics of access to groups in working memory. *J. Exp. Psychol.* 38, 1659–1674. doi: 10.1037/a0028469
- Fischer, M. H. (2001). Probing spatial working memory with the Corsi blocks task. *Brain Cogn.* 45, 143–154. doi: 10.1006/brcg.2000.1221
- Guerard, K., and Tremblay, S. (2012). The effects of path length and display size on memory for spatial information. *Exp. Psychol.* 59, 147–152. doi: 10.1027/1618-3169/a000137
- Han, S., Ding, Y., and Song, Y. (2002). Neural mechanisms of perceptual grouping in humans as revealed by high density event related potentials. *Neurosci. Lett.* 319, 29–32. doi: 10.1016/S0304-3940(01)02502-2
- Han, S., Humphreys, G. W., and Chen, L. (1999). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Percept. Psychophys.* 61, 661–674. doi: 10.3758/BF03205537
- Howell, D. (2002). *Statistical Methods for Psychology, 5th Edn.* Duxbury, MA; Pacific Grove, CA.
- Hurlstone, M. J., Hitch, G. J., and Baddeley, A. D. (2014). Memory for serial order across domains: an overview of the literature and directions for future research. *Psychol. Bull.* 140, 339–373. doi: 10.1037/a0034221
- Imbo, I., Szmalec, A., and Vandierendonck, A. (2009). The role of structure in age-related increases in visuo-spatial working memory span. *Psychol. Belgica* 49, 275–290. doi: 10.5334/pb-49-4-275
- Ishai, A., Ungerleider, L. G., and Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990. doi: 10.1016/S0896-6273(00)00168-9
- Jarvis, H. L., and Gathercole, S. E. (2003). Verbal and non-verbal working memory and achievements on National Curriculum tests at 11 and 14 years of age. *Educ. Child Psychol.* 20, 123–140.
- Jones, D. M., Farrand, P., Stuart, G., and Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *J. Exp. Psychol.* 21, 103–115. doi: 10.1037/0278-7393.21.4.1008
- Kaplan, E., Fein, D., Morris, R., and Delis, D. C. (1991). *WAIS-R NI Manual*. San Antonio, TX: Psychological Corporation.
- Kemps, E. (1999). Effects of complexity on visual-spatial working memory. *Eur. J. Cogn. Psychol.* 11, 335–356. doi: 10.1080/713752320
- Kemps, E. (2001). Complexity effects in visuo-spatial working memory: Implications for the role of long-term memory. *Memory* 9, 13–27. doi: 10.1080/09658210042000012
- Kessels, R. P., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., and de Haan, E. H. F. (2000). The Corsi Block-Tapping Task: standardization and normative data. *Appl. Neuropsychol.* 7, 252–258. doi: 10.1207/S15324826AN0704_8
- Klahr, D., Chase, W. G., and Lovelace, E. A. (1983). Structure and process in alphabetic retrieval. *J. Exp. Psychol.* 9, 462–477. doi: 10.1037/0278-7393.9.3.462
- Kurylo, D. D. (1997). Time course of perceptual grouping. *Percept. Psychophys.* 59, 142–147. doi: 10.3758/BF03206856
- Lashley, K. S. (1951). “The problem of serial order in behavior,” in *Cerebral Mechanisms in Behavior*, ed L. A. Jeffries (New York, NY: Wiley), 112–147.
- Logie, R. H., Trawley, S., and Law, A. (2011). Multitasking: Multiple, domain-specific cognitive functions in a virtual environment. *Mem. Cognit.* 39, 1561–1574. doi: 10.3758/s13421-011-0120-1
- Mechelli, A., Price, C. J., Friston, K. J., and Ishai, A. (2004). Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cereb. Cortex* 14, 1256–1265. doi: 10.1093/cercor/bbh087
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97.
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *Cortex* 27, 272–277.
- Milton, K. (1993). Diet and primate evolution. *Sci. Am.* 269, 86–93. doi: 10.1038/scientificamerican0893-86
- Nemmi, F., Boccia, M., Piccardi, L., Galati, G., and Guariglia, C. (2013). Segregation of neural circuits involved in spatial learning in reaching and navigational space. *Neuropsychologia* 51, 1561–1570. doi: 10.1016/j.neuropsychologia.2013.03.031
- Orsini, A., Pasquandibiseglie, M., Picone, L., and Tortora, R. (2001). Factors which influence the difficulty of the spatial path in Corsi's block-tapping test. *Percept. Mot. Skills* 92, 732–738. doi: 10.2466/pms.2001.92.3.732
- Orsini, A., Simonetta, S., and Marmorato, M. S. (2004). Corsi's block-tapping test: some characteristics of the spatial path which influence memory. *Percept. Mot. Skills* 98, 382–388. doi: 10.2466/pms.98.2.382-388
- Parmentier, F. B. R., and Andrés, P. (2006). The impact of path crossing on visuo-spatial serial memory: encoding or rehearsal effect? *Q. J. Exp. Psychol.* 59, 1867–1874. doi: 10.1080/17470210600872154
- Parmentier, F. B., Andres, P., Elford, G., and Jones, D. M. (2006). Hierarchical organization in visuo-spatial serial memory: interaction of temporal order with spatial and temporal grouping. *Psychol. Res.* 70, 200–217. doi: 10.1007/s00426-004-0212-7
- Parmentier, F. B., Elford, G., and Maybery, M. T. (2005). Transitional information in spatial serial memory: path characteristics affect recall performance. *J. Exp. Psychol.* 31, 412–427. doi: 10.1037/0278-7393.31.3.412
- Piccardi, L., Bianchini, F., Argento, O., De Nigris, A., Maialetti, A., Palermo, L., et al. (2013). The walking Corsi test (WalCT): standardization of topographical memory in an Italian population. *Neurol. Sci.* 34, 971–978. doi: 10.1007/s10072-012-1175-x
- Piccardi, L., Iaria, G., Ricci, M., Bianchini, F., Zompanti, L., and Guariglia, C. (2008). Walking in the Corsi test: which type of memory do you need? *Neurosci. Lett.* 432, 127–131. doi: 10.1016/j.neulet.2007.12.044
- Ridgeway, D. (2006). Strategic grouping in the spatial span memory task. *Memory* 14, 990–1000. doi: 10.1080/09658210601010797
- Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., and van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Hum. Mov. Sci.* 26, 525–554. doi: 10.1016/j.humov.2007.04.001
- Rossi-Arnaud, C., Pieroni, L., and Baddeley, A. (2005). Symmetry and binding in visuo-spatial working memory. *Neuroscience* 139, 393–400. doi: 10.1016/j.neuroscience.2005.10.048
- Rossi-Arnaud, C., Spataro, P., and Longobardi, E. (2012). Effects of pointing on the recall of simultaneous and sequential visuospatial arrays: a role for retrieval strategies? *Psychol. Res.* 76, 699–712. doi: 10.1007/s00426-011-0394-8
- Smirni, P., Villardita, C., and Zappalá, G. (1983). Influence of different paths on spatial memory performance in the Block-Tapping Test. *J. Clin. Neuropsychol.* 5, 355–359. doi: 10.1080/01688638308401184
- Smyth, M. M., and Scholey, K. A. (1992). Determining spatial span: the role of movement time and articulation rate. *Q. J. Exp. Psychol.* 45A, 479–501.
- Smyth, M. M., and Scholey, K. A. (1994). Characteristics of spatial memory span: is there an analogy to the word length effect, based on movement time? *Q. J. Exp. Psychol.* 47, 91–117. doi: 10.1080/14640749408401145
- St Clair-Thompson, H. L. (2007). The effects of cognitive demand upon relationships between working memory and cognitive skills. *Q. J. Exp. Psychol.* 60, 1378–1388. doi: 10.1080/17470210601025505
- Trawley, S. L., Law, A. S., and Logie, R. H. (2011). Event-based prospective remembering in a virtual world. *Q. J. Exp. Psychol.* 64, 2181–2193. doi: 10.1080/17470218.2011.584976
- Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale—Third edition: Administration and Scoring Manual*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997b). *Wechsler Memory Scale – Third edition: Administration and Scoring Manual*. San Antonio, TX: The Psychological Corporation.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 De Lillo, Kirby and Poole. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Working Memory for Sequences of Temporal Durations Reveals a Volatile Single-Item Store

Sanjay G. Manohar* and Masud Husain

Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Hermann Josef Mueller,
Ludwig Maximilian University
of Munich, Germany
Mark A. Elliott,
National University of Ireland Galway,
Ireland
Stephen Emrich,
Brock University, Canada

*Correspondence:

Sanjay G. Manohar
sanjay.manohar@psy.ox.ac.uk

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 30 June 2016

Accepted: 10 October 2016

Published: 26 October 2016

Citation:

Manohar SG and Husain M (2016)
Working Memory for Sequences of
Temporal Durations Reveals a Volatile
Single-Item Store.
Front. Psychol. 7:1655.
doi: 10.3389/fpsyg.2016.01655

When a sequence is held in working memory, different items are retained with differing fidelity. Here we ask whether a sequence of brief time intervals that must be remembered show recency effects, similar to those observed in verbal and visuospatial working memory. It has been suggested that prioritizing some items over others can be accounted for by a “focus of attention,” maintaining some items in a privileged state. We therefore also investigated whether such benefits are vulnerable to disruption by attention or expectation. Participants listened to sequences of one to five tones, of varying durations (200 ms to 2 s). Subsequently, the length of one of the tones in the sequence had to be reproduced by holding a key. The discrepancy between the reproduced and actual durations quantified the fidelity of memory for auditory durations. Recall precision decreased with the number of items that had to be remembered, and was better for the first and last items of sequences, in line with set-size and serial position effects seen in other modalities. To test whether *attentional filtering* demands might impair performance, an irrelevant variation in pitch was introduced in some blocks of trials. In those blocks, memory precision was worse for sequences that consisted of only one item, i.e., the smallest memory set-size. Thus, when irrelevant information was present, the benefit of having only one item in memory is attenuated. Finally we examined whether *expectation* could interfere with memory. On half the trials, the number of items in the upcoming sequence was cued. When the number of items was known in advance, performance was paradoxically worse when the sequence consisted of only one item. Thus the benefit of having only one item to remember is stronger when it is unexpectedly the only item. Our results suggest that similar mechanisms are used to hold auditory time durations in working memory, as for visual or verbal stimuli. Further, solitary items were remembered better when more items were expected, but worse when irrelevant features were present. This suggests that the “privileged” state of one item in memory is particularly volatile and susceptible to interference.

Keywords: working memory, attention, duration, serial position effect, focus of attention

INTRODUCTION

When a series of items is held in working memory, not all items are held with equal fidelity. Items early in the sequence may be forgotten, whereas items at the very start of a sequence may be easier to find. The *final* item in a sequence may also be held in a more “active,” privileged or prioritized state (Allen et al., 2014; Hu et al., 2016). This is known as the “recency effect,” and has been shown

to be volatile, susceptible to a number of attentional manipulations (Davelaar et al., 2005; Hu et al., 2014). It decays quickly (Postman and Phillips, 1965; LaRocque et al., 2014), may be selectively impaired by TMS or lesions to modality-specific cortex (Vallar and Papagno, 1986; Zokaei et al., 2014a), and may relate to earlier items being forgotten through retroactive interference (Kool et al., 2014). For these reasons, it has been postulated that the benefits enjoyed by the final item in a sequence arise because it remains in the *focus of attention*.

Recent studies of working memory have begun to use continuous recall measures, which allow the precision or fidelity with which items are stored to be quantified. Most of these studies have used visual working memory, measuring the precision of storing spatial locations, colors or orientations (Bays and Husain, 2008; Zhang and Luck, 2008). These paradigms require participants to reproduce their memory of a continuously variable feature, for example by adjusting a dial. The reported feature can then be compared to the veridical feature, providing a trial-wise, quantitative precision measure. Recently, these precision paradigms have been extended to auditory and vibrotactile frequencies, and similar effects have been demonstrated, indicating that features in various modalities may all be encoded in a similar way (Kumar et al., 2013; Joseph et al., 2015). In neural models of working memory, the ability to hold several continuous features in memory has been taken to suggest that the feature dimensions are encoded in a set of independent feature-tuned channels, which are activated upon encoding each feature for each item in memory (Compte et al., 2000; Wimmer et al., 2014).

Could a similar storage method be used to hold temporal durations in memory? Periods of time are abstract: durations do *not* traditionally form a parameterised space represented by cells in sensory transduction. Intervals of time might need to be explicitly extracted or inferred from other kinds of representation (Matthews and Meck, 2016). Durations are also unusual things to hold in short-term memory. Despite this, it appears that we do in fact possess working memory for durations (Teki and Griffiths, 2014). Indeed we are able to repeat rhythms that we hear, for example in music, poetry or speech (McAuley, 2010). But it is not clear that the same mechanisms would be involved, as those that subserve visual or verbal working memory. The presence of set-size, serial position, and attentional effects could provide evidence for commonality of mechanisms.

Human time perception has been most commonly studied with simple interval estimation, reproduction and comparison tasks (Grondin, 2010). A number of factors increase or decrease the perceived duration of an interval. Practice can lengthen perceived durations (Eisler, 1976), as can arousal (Wittmann, 2013), whereas aging shortens them (Baudouin et al., 2006b). *Attention* and *expectation* play particularly important roles in interval timing. Attentional loads shorten perceived durations while they are experienced (Brown, 1985, 1997; Block et al., 2010) but lengthen the reproduction of durations (Fortin and Breton, 1995; Baudouin et al., 2006a,b). Evidence from patients also implicates attention in timing,

with patients reporting shorter and less accurate estimates of durations (Danckert et al., 2007). We therefore studied whether attentional demands might alter retention of durations in working memory, by introducing variation of an irrelevant feature.

Importantly, expectation also impacts on timing. The presence of distractors during a time judgment task can lengthen the subjective duration of a stimulus, but this effect only arises when the distractors are unexpected (Penney et al., 2014). Similarly, producing an interval that is interrupted by a pause late in the interval leads to overestimation; this effect persisted on trials when a break did not actually occur, but was expected to occur (Fortin and Massé, 2000). These results suggest that expectation of an upcoming event shortens perceived durations. In the present study we investigate whether simply expecting an event could enhance memory retention for durations.

We set out to test a direct analog of visual working memory experiments, in the time domain. In particular we asked, does memory for durations show similar set-size and serial position effects as visual working memory? Further, we enquired whether set-size and serial position effects are susceptible to manipulation of attention and expectation. We asked whether the need to filter irrelevant information, and the expectation of the end of a sequence, altered the recency effect. We hypothesized that any attentional benefits would be attenuated if irrelevant features were being ignored. Regarding temporal expectation, we predicted that the unexpected end of a sequence can confer a recency benefit, whereas if the ends of sequences were expected, this advantage would be lost.

GENERAL METHODS

Participants were instructed to listen to each sequence of tones, and remember the time each one lasted for. They were told that after a delay, they would see a signal indicating which of the items in the sequence they had to recall (probed by serial order), and that they had to press and hold a key to try and match that duration as precisely as they could (**Figure 1A**).

Participants sat in a dimly lit room viewing a CRT monitor at a distance of 40 cm from a chinrest. Tones were presented through a pair of stereo speakers, situated either side of the computer screen 50 cm in front of the subjects, at shoulder height. Tones comprised a sine wave at 440 Hz (Experiments 1 and 3). Each tone was modulated to taper linearly over the first and last 10 ms, to minimize transients. The durations to be remembered were selected from a uniform distribution between 200 and 2000 ms. Sequences of 1 to 5 durations (Experiment 1) or 1 to 4 durations (Experiments 2 and 3) were chosen, with proportionally more trials for higher set-sizes. This permitted each serial position in each set-size to be probed equally frequently. The tones were separated by a fixed 500 ms inter-stimulus interval. After the end of the final tone, there was a 1000 ms silent retention interval.

At the end of the retention interval, the computer screen displayed a cue indicating which item was to be recalled. This was done graphically by displaying a row of squares, each

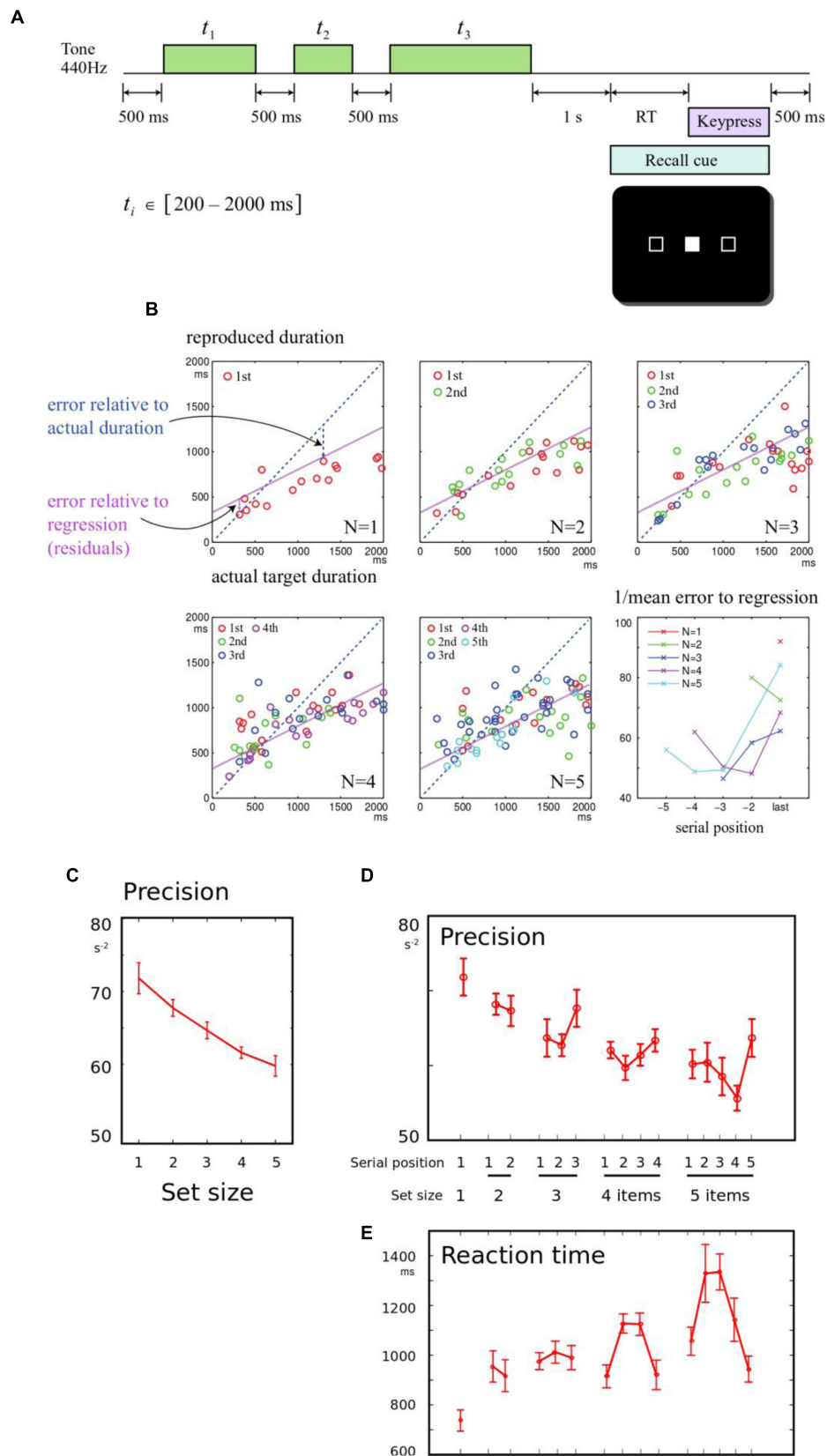


FIGURE 1 | Continued

FIGURE 1 | (A) Memory for durations task: In order to study how a series of durations are held in working memory, participants were asked to listen to a sequence of one to five tones. After a 1 s delay, they were cued to one of the tones by its serial position. They had to reproduce the duration of the cued tone by holding down a response key for a matching duration. The durations to be remembered were drawn from a uniform distribution between 200 and 2000 ms. **(B)** Example of results from a single participant: Panels correspond to sequence lengths 1–5 items. In each panel, the response durations of all trials are plotted, as a function of the corresponding target duration. The blue diagonal dotted line indicates perfect performance, where responses would be identical to the heard durations. The purple line indicates a linear regression fit to all the subject's responses. The slope is flat and intercept is positive, indicating that short durations are overestimated and long durations are underestimated. Errors were calculated relative to the regression line. The final panel shows the precision (reciprocal of the root mean square error) calculated for each set-size and serial position, for this subject. Colors indicate different set-sizes, and the final item for each sequence length is aligned to the right. **(C)** Precision falls with longer sequences: When more durations needed to be remembered, the overall precision of reported durations was reduced. Data were collapsed across the serial positions, and the inverse mean error for each sequence length is shown. Error bars indicate within-subject error for the effect of set-size, across all participants. **(D)** Precision shows primacy and recency effects: The mean error was broken down by serial position, demonstrating an overall benefit for the last item in a sequence (recency effect). **(E)** Response times mirror memory precision: Responses were faster when fewer items had to be remembered. Serial effects were also observed, with faster responses for the first and last items in 4- or 5-item sequences (primacy and recency effects), as predicted by an information-accumulation model of response time.

representing one of the items heard on the current trial, in sequential order from left to right. One square was filled in, indicating the item that had to be recalled. For example, if four tones were heard, there would be four boxes, and to indicate that the first tone should be recalled, the left-most square was filled in white, whereas the other three were hollow frames.

Participants then pressed and released the key, to indicate their memory of the duration of the indicated tone. After the response, an inter-trial interval of 500 ms followed, and the next trial began. No feedback was provided.

In all experiments, 10 practice trials were performed, and participants were debriefed to check they understood the task, before the experiment began.

EXPERIMENT 1: WORKING MEMORY FOR DURATIONS

Methods

Experiment 1 required participants to remember 1- to 5-item sequences, and each serial position in the sequence was probed equally often. This gave 15 trial types, with more 5-item trials than 1-item trials. There were 60 trials per block, in 4 blocks, separated by a 2-min break. 15 participants performed this experiment.

Participants were recruited from the UCL Psychology subject pool, and were aged 18–36 years (mean 26.5 years). All subjects gave informed written consent as approved by the UCL Research Ethics Committee.

Results

Our primary measure was recall error. As expected, there was an overall linear relationship between the recalled duration and the corresponding presented duration, and this relationship showed systematic overestimation of shorter intervals, and underestimation of longer intervals (**Figure 1B**). This demonstrates a well-studied bias in interval reproduction, and in line with other studies we used a linear fit to model this bias (Jazayeri and Shadlen, 2010, 2015). This permitted us to measure error relative to each individual's linear fit, as the residuals of the regression. Thus on each trial, the discrepancy relative to this fit could be calculated, that indicated the fidelity of memory recall. Memory fidelity, or *precision*, was quantified as the reciprocal of

the root mean squared error, calculated for each condition for each subject.

First, the effect of set-size was examined, collapsing across serial positions. Increasing set-size strongly reduced precision [**Figure 1C**, $F(4,56) = 9.53$, $p < 0.001$]. To establish how set-size and serial position influenced recall, a one-way repeated measures ANOVA was first performed across all set-size and serial position conditions. For five set-sizes, this gave 15 conditions. The conditions differed significantly [**Figure 1D**, $F(14,193) = 3.46$, $p < 0.001$]. The primacy effect was not significant [ANOVA of first and second item in sequences length 2 to 5, $F(1,98) = 0.52$, $p > 0.05$], but there was a significant recency effect [last vs. penultimate item in sequences length 2 to 5, $F(1,98) = 6.64$, $p = 0.012$].

A further linear regression within each condition produced similar results (Supplementary Figure S1), confirming that set-size and serial position effects were truly due to precision, rather than systematic bias. All results were also robust to normalizing by logarithmic transformations of the times (Supplementary Materials).

Reaction times (RT) were measured from the probe onset (appearance of filled box) until the button was initially depressed (initiation of the production interval). This interval therefore represents the time taken to identify the probed item, bring its duration to mind, and prepare a response. RT tended to be greater whenever precision was lower, and exhibited significant set-size effects (**Figure 1E**). RT was also faster for both the first and last items of a sequence, exhibiting both recency and primacy [$F(1,98) = 7.22$, $p = 0.009$ and $F(1,98) = 7.60$, $p = 0.007$]. The findings are in keeping with information-accumulation models of retrieval from memory that have been proposed in visual working memory (Pearson et al., 2014; Schneegans and Bays, 2016).

EXPERIMENT 2: EFFECT OF VARIATION ON AN IRRELEVANT FEATURE DIMENSION

Methods

We next asked whether the presence of an irrelevant feature would alter memory fidelity, as a function of set-size or serial

position. Variation in this additional feature might invoke attentional filtering, and thus impair memory performance specifically for items that rely on attention.

In experiment 2, the pitch of the tone was randomly chosen between 440 and 880 Hz. It was emphasized to participants that the pitch was irrelevant, and that only duration had to be remembered. In some blocks, the pitch of *each tone* in a trial was varied randomly. In other blocks, the pitch of tones within a trial was kept constant, but randomly selected for each trial (**Figure 2A**). There were 1 to 4 items in each sequence, and different serial positions were probed on each trial. There were thus 10 combinations of set-size and probe position, with 9 repeats in each block giving 90 trials in each of four blocks. The two block-wise conditions, variable vs. constant pitches, were counterbalanced in order across subjects, such that eight participants performed blocks in the order “ABBA,” and eight in the order “BAAB.” For one participant, who did the constant block first, two blocks of data were lost, so their data were discarded, giving a total of 15 participants.

Results

Precision was compared with a 2-way ANOVA, with factor 1 distinguishing the 10 possible combinations of set-size and probe, and factor 2 indicating the block type, i.e., the presence or absence of variation in the irrelevant feature. An interaction was observed between the probed item and presence of variation [$F(9,266) = 2.16, p = 0.025$], in addition to a main effect of item [$F(9,266) = 2.56, p = 0.008$], with no main effect of variability [$F(1,266) < 0.1$]. This interaction suggests that attentional filtering had selective effects on some memory conditions (**Figures 2B,C**). *Post hoc* tests revealed that the interaction was driven by variation impairing recall specifically in the 1-item condition [$t(14) = 2.83, p = 0.013$], but no effects of variation were observed for any serial position for the other set-sizes. Therefore, only when a single duration had to be remembered, was there an effect of expecting variability in the current block. The same results were obtained when using separate regressions for each condition, set-size, and serial position. This indicates that the filtering effect was not due to a change in bias (Supplementary Materials). The pairwise tests were robust to normalization by log transform and non-parametric *U*-test.

There was no effect of variability on the primacy or recency effect, as quantified by interactions with the difference between the first two or last two items in 2- to 4- item sequences (both $F < 1.26, p > 0.05$).

Reaction times showed strong set-size effects as before (**Figure 2D**). However, there was no main effect of variability, no interaction with primacy [$F(1,154) = 0.22$], and a trend for variability to reduce the RT recency effect [$F(1,154) = 3.27, p = 0.073$].

Incidentally we noted that higher pitched tones were perceived as 1% longer in keeping with greater subjective intensity (Goldstone and Lhamon, 1974), but this effect did not interact with variability in our study (Supplementary Materials).

EXPERIMENT 3: EFFECT OF EXPECTING A SEQUENCE'S LENGTH

Methods

In experiment 3, half the trials began with a cue screen lasting 500 ms, and the other half of trials began with a cross at the screen center. The cue screen consisted of a horizontal set of empty boxes, with the number of boxes indicating the number of tones that would be presented on the upcoming trial (**Figure 3A**). After the cue, the tones were presented and probed as in Experiments 1 and 2. There were 90 trials in four blocks, with all conditions interleaved. Fourteen participants performed the experiment, but one did not complete the task, leaving 13 datasets.

Results

Precision was compared using a 2-way ANOVA, with factor 1 distinguishing the 10 possible set-size/probe conditions, and factor 2 indicating whether the set-size cue was present or absent. There was a main effect of item probed [**Figure 3B**, $F(9,228) = 2.08, p = 0.032$], no main effect of cue presence [$F(1,228) = 1.54, p > 0.05$], but a significant interaction between item and cue [$F(9,228) = 2.00, p = 0.040$]. This interaction was driven by a significant cue effect only for the 1-item condition [**Figure 3C**, *post hoc t*-test, $t(12) = 4.04; p = 0.009$], with no significant differences for any serial position in any other set-size. There was no effect of cue upon primacy or recency (both $F < 0.23$). The same effects were found when condition-wise regression was used to calculate the precision (Supplementary Materials), and the effect was robust to log-transform normalization and non-parametric *U*-tests.

Reaction times was significantly greater when no cue was presented [**Figure 3D**, $F(1,228) = 36, p < 0.001$], with no interaction with item [$F(9,228) = 1.02, p > 0.05$]. This indicated that expecting the end of the sequence improved the speed of responding generally. This confirms that cueing did indeed have the anticipated effect of improving expectation of when the probe would occur – but this was in sharp contrast to the above findings that precision was unchanged or *worse* with the cue.

DISCUSSION

This study asked whether memory for a sequence short auditory durations follows well-known laws associated with working memory in other modalities. The results confirm the existence of set-size and serial position effects that are in line with other modalities (**Figure 1D**). We then asked whether attention and expectation could modulate memory for durations. When a variable irrelevant feature was introduced into the sequences, memory for single items was worse, suggesting that the high performance normally observed for single items may be susceptible to attentional disruption (**Figure 2C**). When the number of items was expected, we found a similar disruption of the ability to recall a single item in memory. Thus single items are best remembered when they are unexpectedly the only item (**Figure 3C**). The benefit of having to remember one item, rather than a sequence is thus disrupted by the presence of irrelevant

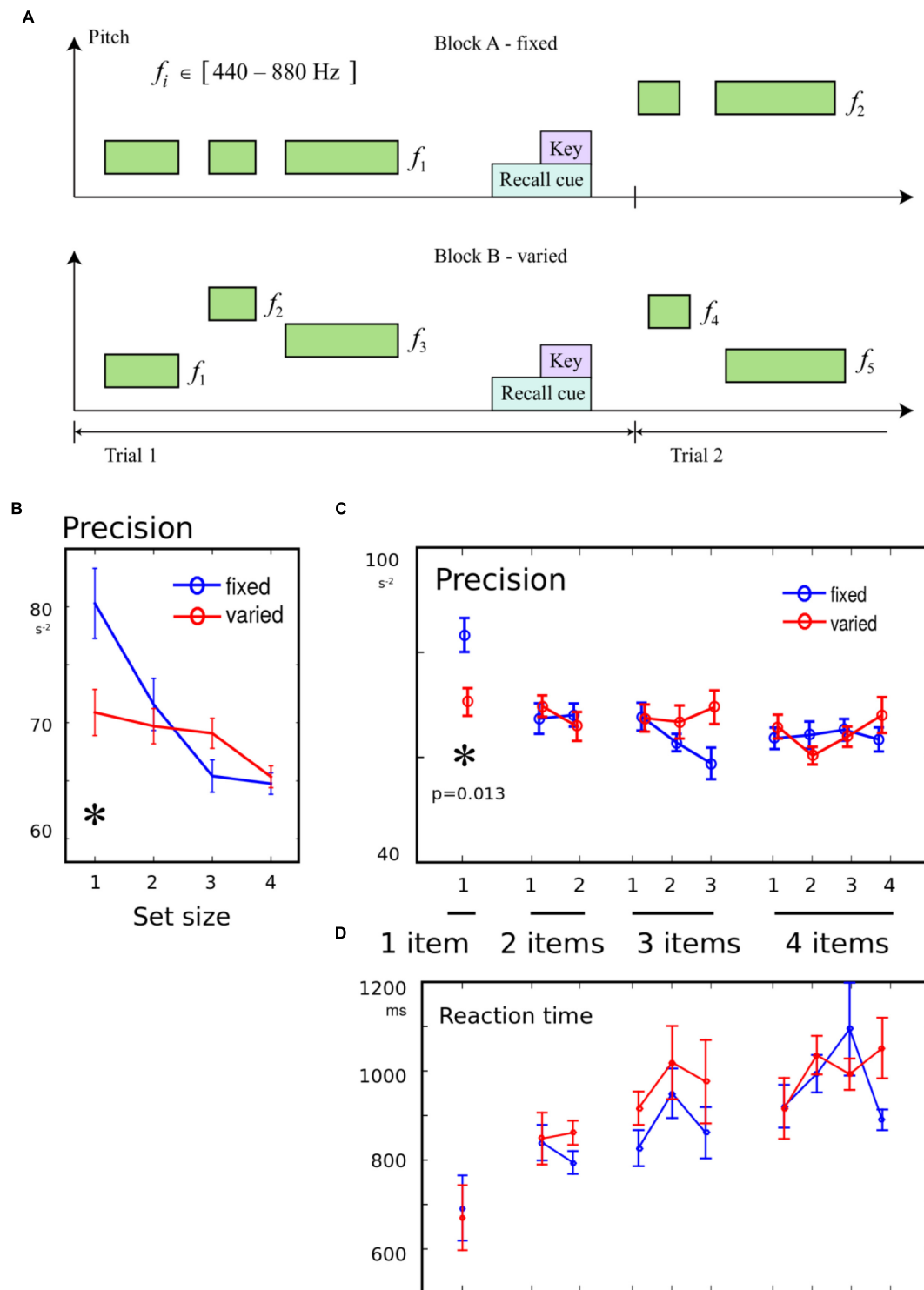


FIGURE 2 | (A) Variation in pitch was an irrelevant feature: To examine whether the introduction of an irrelevant variation in pitch would worsen recall, two of the four blocks had the same pitch for all tones in each trial (“fixed” condition). In the remaining blocks, each tone on every trial had a randomly chosen pitch (“variable” condition). Between 1 and 4 tones were presented on each trial. **(B)** Precision for a single tone was worse in variable block: When only one item had to be remembered, memory precision was worse in blocks where pitches varied from trial to trial, compared to blocks when they were constant. There was no effect of variability when multiple items had to be remembered. **(C)** No effect of variability on serial position effects: The trials were broken down according to serial position in the sequence. Although there was an interaction of variability with memory condition, the only effect of variability was seen for the one-tone condition. **(D)** No effect of variability on reaction time (RT): There were no differences in the time to initiate a response, between variable blocks and fixed blocks.

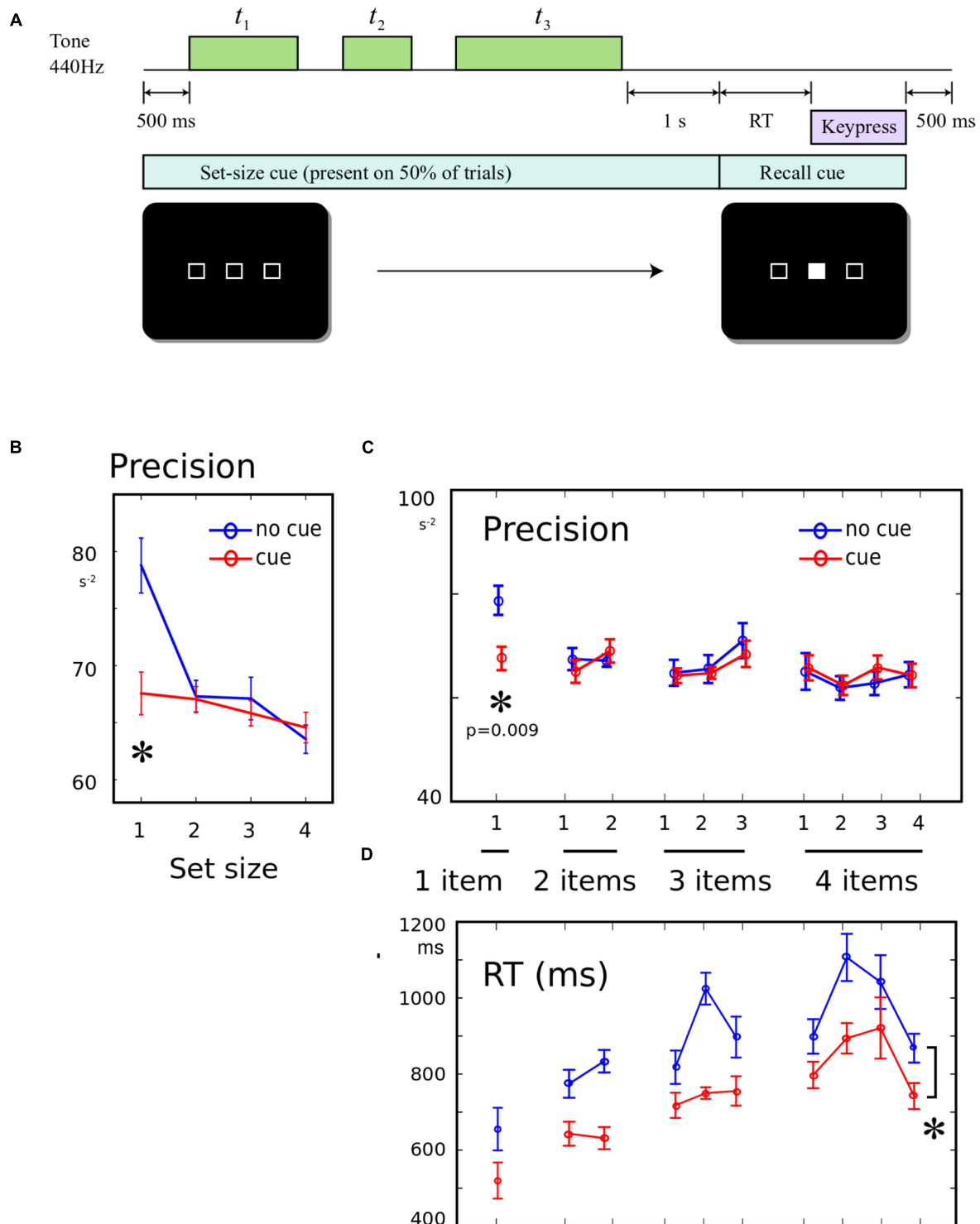


FIGURE 3 | (A) Informing participants about the number of upcoming tones: 500 ms prior to the start of each trial, a screen was presented. On half the trials, this screen comprised a row of empty boxes, with the number of boxes corresponding to the number of tones that will be presented on this trial. The cue remained on-screen until the recall cue. On the remainder of trials, a cross was displayed instead, giving no information about the number of tones. **(B)** Precision for a single memory item was worse when number of items is known: The recall precision was significantly lower when a single tone was presented, if it was *expected* that the tone would be single, compared to when it was unexpectedly single. **(C)** No effect of pre-cueing set-size on other conditions: There was no worsening or improvement in recall when the number of items was known in advance, and there was no change in the shape of the serial position curve. **(D)** Response times faster when set-size was known in advance: On cued trials, RT was significantly shorter (main effect of cue). This did not interact with serial position.

information, but enhanced when more items in the sequence are expected.

One possible interpretation of these findings is that both filtering and expectation interfere with a common aspect of the maintenance of singular items. Could the benefits for single items be mediated by an attentional focus, conferring more mnemonic resources on an item that is isolated? In attentional focus models, some items in memory are held in a “privileged” state (Oberauer, 2002; Cowan, 2011). If multiple items are held in working memory, not all items are recalled with the same accuracy, and some of the differences between items may arise because of a privileged or attended state conferred to one item (Lewis-Peacock et al., 2012; Postle, 2015). The benefit can be transferred among items (Zokaei et al., 2014b), and may explain the susceptibility of recency effects to attentional manipulations (Öztekin et al., 2010; Morrison et al., 2014; Souza et al., 2016). Our findings add weight to suggestions that working memory may contain one high-resolution but volatile representation. However, it is notable that we did *not* find attentional disruptions for the last item of longer sequences. This suggests that the recency effect might not always be susceptible to attentional load or expectation.

There may be other, more complex reasons for the disruption by expectation and filtering. Eye movements may distort time perception (Morrone et al., 2005; Burr et al., 2010) and thus the cue preceding the stimulus might alter time perception. Another possibility is that expectation of an event (e.g., further items that might be presented) can *increase* perceived durations (Fortin and Massé, 2000; Penney et al., 2014). Alternatively, a dual-task effect might occur in the filtering condition, which is known to *shorten* perceived durations of stimuli (Block et al., 2010). In Experiment 3 pitch differences between notes may also increase the perceived duration of gaps between tones (Crowder and Neath, 1995; Lake et al., 2014).

However, all these effects would be expected to lead to distortions of perception and thus *systematic biases* to over- or under-estimate the duration. In contrast, our results suggested no bias but an increase in error, i.e., the *variability* of responses around the same fixed duration. This might suggest that the measured effects occur at encoding or storage, rather than being perceptual biases. Could focusing attention explain the results of experiments 2 and 3?

An important property of attention is its refractoriness, as characterized by attentional blink or inhibition of return. These phenomena impose temporal capacity limits on deployment of attention to sequentially presented items. Could the set-size cue itself capture attention in experiment 3, and impair encoding of the subsequent tone? We think this is a less likely explanation, because the attentional blink tends to arise between 150 and 450 ms after a stimulus (Shapiro et al., 1997), whereas the gaps in our task were 500 ms. Further, this might also be expected to slow down RTs, whereas we in fact observed faster RTs. We suggest an alternative explanation: when the recall cue is expected after the item, the preparation of the response begins as soon as the tone ends. Note that such immediate response preparation cannot occur in any other type of trial, because either the item-to-be-probed is unknown, or the end of the sequence is not

expected. Early response preparation could be the factor that leads to disruption of duration memory.

The present study is one of few that examine working memory for durations that are not intrinsically rhythmic. Most studies that investigate memory for multiple durations test our ability to discriminate rhythms, i.e., sequences of durations that are integer multiples of a discrete, quantised beat (Penhune et al., 1998; Teki et al., 2011; Grahn, 2012b). Such studies do demonstrate limitations in the number of durations that can be remembered, but give no indication of the precision with which each duration is remembered. Rhythm discrimination may in fact predispose subjects to use discrete categorical strategies for representing time, whereas for non-rhythmic time sequences, different neural mechanisms are thought to be recruited (Grahn and Brett, 2007; Grube et al., 2010; Joseph et al., 2016).

It is possible that in our task, durations could either be encoded individually, as absolute time intervals, or as relative times approximating a rhythmic structure. The present study is not able to distinguish these two possibilities. Rhythm perception involuntarily leads to complex changes in perceived intensity and timing, which vary according to expertise (Povel and Essens, 1985). Perceiving rhythm also leads to phase-dependent facilitation for many aspects of auditory perception and cognition (McAuley, 2010; Grahn, 2012a). Expecting further items in a sequence (Experiment 3) could potentially promote adoption of a rhythmic strategy, and this strategic change might drive the improvements observed with expectation. Rhythm-perceptual effects may be overlaid upon working memory effects, and could lead to more efficient storage of intervals at the expense of precision (Jones and Ralston, 1991; Large, 2002), similar to “lossy compression” or configural effects observed in visual memory (Alvarez, 2011). Further study would be required to directly measure the effect of rhythm-based encoding on duration memory.

We cannot exclude that our findings might be specific to the auditory modality, rather than representing a general effect in temporal cognition. The duration of auditory stimuli are generally reproduced more precisely than visual stimuli, with sounds being perceived as lasting around 20% longer than lights of matched duration (Goldstone and Lhamon, 1974; Wearden et al., 2006; McAuley and Henry, 2010). As we were primarily interested in the precision of memory, we used auditory stimuli for this experiment. Moreover, we tested temporal memory using “filled durations” – i.e., a tone lasting for the desired duration, as opposed to a gap in a tone, or the interval between a pair of clicks delineating an “unfilled” interval. The use of tones minimizes bias caused by start and end markers themselves (Rammsayer and Leutner, 1996), and produces more precise interval reproduction than the duration of gaps, whose durations tend to be systematically underestimated (Wearden et al., 2007).

How might neurones encode time durations in memory? Single time intervals could be reproduced by gradually varying neural activity during the encoding period which, upon termination of the interval, determines the subsequent rate-of-rise of an accumulator (Jazayeri and Shadlen, 2015) – somewhat like a pendulum that swings back to the height it was released from. But in order to use such an arrangement for sequences

of several durations, an elaborate orchestration of segregated neuronal populations would be required (Kleinman et al., 2016). One way of achieving this might be to harness existing domain-general working memory processes.

Our results do suggest a conserved pattern of storage for remembering many different kinds of information, even including short durations. But for this to occur, durations should be encoded by a similar mechanism that has been proposed for other sensory modalities – i.e., a channel-based place code. What is the evidence that time intervals might be encoded by duration-selective channels? First, adaptation effects can be observed when we repeatedly hear a fixed duration, and these cross-modal adaptation effects are highly redolent of those observed in visual orientation and spatial frequency channels (Aaen-Stockdale et al., 2010; Heron et al., 2012). Second, neurophysiology provides evidence for duration-selective channels, for example with channels of the order of 30 ms in the brain stem (Brand et al., 2000; Aubie et al., 2009), around 100 ms in primary auditory cortex (He et al., 1997), up to 400 ms in V1 there, and in prefrontal cortex units have been found that are selective for durations up to 4 s (Yumoto et al., 2011). These neural representations could provide a substrate for storing duration information in working memory. Further, if duration-selective channels of this kind operate similarly to classical visuospatial or auditory feature domains studied in working memory, then similar capacity limits should be evident. In line with this, holding more than one duration in memory reduces the precision with which they can be remembered, and pre-cueing one of several durations can selectively improve memory (Teki and Griffiths, 2014).

Alternative classes of neural model have been proposed to explain how a single interval might be reproduced. First, pacemaker-accumulator models postulate a signal occurring at a fixed average rate that is integrated by a counter, and then compared to some threshold (Treisman, 1963; Gibbon et al., 1984). Second, population clock models posit that neural ensembles transition through a sequence of states in a probabilistic manner to produce accurate timing (Buonomano and Laje, 2010). Third, coincidences of noisy cortical oscillations may be detected by striatal neurons, rendering them sensitive to

“beats” that occur after a learned interval (Oprisan and Buhusi, 2014). However, none of these proposals can straightforwardly account for the ability to hold multiple durations in mind, as observed in the current task. Functional imaging findings suggest that sensorimotor thalamocortical-basal-ganglia pathways may subserve the more complex aspects of temporal cognition (Schubotz et al., 2000; Schubotz and von Cramon, 2001; Grahn, 2012a). Indeed working memory may itself be central in producing an interval, because some form of counter needs to be maintained online during the interval (Brown, 1997; Gu et al., 2015). Individuating items in working memory and interval timing might utilize the same temporal context cues, an idea supported by correlations between memory performance and temporal discrimination performance (Unsworth and Engle, 2005; Broadway and Engle, 2011). Interval timing and working memory might thus be two modes of operation of the same neural system (Gu et al., 2015).

In summary, we show that several temporal durations can be held in working memory at once, and they are subject to standard sequential working memory limits. We demonstrated that the memory of single auditory durations in memory is especially susceptible to manipulations of attention and expectation.

AUTHOR CONTRIBUTIONS

SM conceived, conducted, and analyzed data from the experiments. SM and MH wrote the manuscript.

FUNDING

This work was funded by a Wellcome Trust Principal Fellowship to MH 098282.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.01655/full#supplementary-material>

REFERENCES

- Aaen-Stockdale, C., Hotchkiss, J., Heron, J., and Whitaker, D. (2010). Time channels: adaptation of auditory and visual duration. *Perception* 39:66.
- Allen, R. J., Baddeley, A. D., and Hitch, G. J. (2014). Evidence for two attentional components in visual working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 1499–1509. doi: 10.1037/xlm0000002
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends Cogn. Sci.* 15, 122–131. doi: 10.1016/j.tics.2011.01.003
- Aubie, B., Becker, S., and Faure, P. A. (2009). Computational models of millisecond level duration tuning in neural circuits. *J. Neurosci.* 29, 9255–9270. doi: 10.1523/JNEUROSCI.1085-09.2009
- Baudouin, A., Vanneste, S., Isingrini, M., and Pouthas, V. (2006a). Differential involvement of internal clock and working memory in the production and reproduction of duration: a study on older adults. *Acta Psychol. (Amst.)* 121, 285–296. doi: 10.1016/j.actpsy.2005.07.004
- Baudouin, A., Vanneste, S., Pouthas, V., and Isingrini, M. (2006b). Age-related changes in duration reproduction: involvement of working memory processes. *Brain Cogn.* 62, 17–23. doi: 10.1016/j.bandc.2006.03.003
- Bays, P. M., and Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science* 321, 851–854. doi: 10.1126/science.1158023
- Block, R. A., Hancock, P. A., and Zakay, D. (2010). How cognitive load affects duration judgments: a meta-analytic review. *Acta Psychol. (Amst.)* 134, 330–343. doi: 10.1016/j.actpsy.2010.03.006
- Brand, A., Urban, R., and Grothe, B. (2000). Duration tuning in the mouse auditory midbrain. *J. Neurophysiol.* 84, 1790–1799.
- Broadway, J. M., and Engle, R. W. (2011). Individual differences in working memory capacity and temporal discrimination. *PLoS ONE* 6:e25422. doi: 10.1371/journal.pone.0025422
- Brown, S. W. (1985). Time perception and attention: the effects of prospective versus retrospective paradigms and task demands on perceived duration. *Percept. Psychophys.* 38, 115–124. doi: 10.3758/BF03198848

- Brown, S. W. (1997). Attentional resources in timing: interference effects in concurrent temporal and nontemporal working memory tasks. *Atten. Percept. Psychophys.* 59, 1118–1140. doi: 10.3758/BF03205526
- Buonomano, D. V., and Laje, R. (2010). Population clocks: motor timing with neural dynamics. *Trends Cogn. Sci.* 14, 520–527. doi: 10.1016/j.tics.2010.09.002
- Burr, D. C., Ross, J., Binda, P., and Morrone, M. C. (2010). Saccades compress space, time and number. *Trends Cogn. Sci.* 14, 528–533. doi: 10.1016/j.tics.2010.09.005
- Compte, A., Brunel, N., Goldman-Rakic, P. S., and Wang, X.-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* 10, 910–923. doi: 10.1093/cercor/10.9.910
- Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: making sense of competing claims. *Neuropsychologia* 49, 1401–1406. doi: 10.1016/j.neuropsychologia.2011.01.035
- Crowder, R. G., and Neath, I. (1995). The Influence of pitch on time perception in short melodies. *Music Percept. Interdiscip. J.* 12, 379–386. doi: 10.2307/40285672
- Danckert, J., Ferber, S., Pun, C., Broderick, C., Striemer, C., Rock, S., et al. (2007). Neglected time: impaired temporal perception of multisecond intervals in unilateral neglect. *J. Cogn. Neurosci.* 19, 1706–1720. doi: 10.1162/jocn.2007.19.10.1706
- Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., and Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychol. Rev.* 112, 3–42. doi: 10.1037/0033-295X.112.1.3
- Eisler, H. (1976). Experiments on subjective duration 1868–1975: a collection of power function exponents. *Psychol. Bull.* 83, 1154–1171. doi: 10.1037/0033-2909.83.6.1154
- Fortin, C., and Breton, R. (1995). Temporal interval production and processing in working memory. *Atten. Percept. Psychophys.* 57, 203–215. doi: 10.3758/BF03206507
- Fortin, C., and Massé, N. (2000). Expecting a break in time estimation: attentional time-sharing without concurrent processing. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1788–1796. doi: 10.1037/0096-1523.26.6.1788
- Gibbon, J., Church, R. M., and Meck, W. H. (1984). Scalar timing in memory. *Ann. N. Y. Acad. Sci.* 423, 52–77. doi: 10.1111/j.1749-6632.1984.tb23417.x
- Goldstone, S., and Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment: I. Sounds are judged longer than lights. *Percept. Mot. Skills* 39, 63–82. doi: 10.2466/pms.1974.39.1.63
- Grahn, J. A. (2012a). Neural mechanisms of rhythm perception: current findings and future perspectives. *Top. Cogn. Sci.* 4, 585–606. doi: 10.1111/j.1756-8765.2012.01213.x
- Grahn, J. A. (2012b). See what I hear? Beat perception in auditory and visual rhythms. *Exp. Brain Res.* 220, 51–61. doi: 10.1007/s00221-012-3114-8
- Grahn, J. A., and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* 19, 893–906. doi: 10.1162/jocn.2007.19.5.893
- Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* 72, 561–582. doi: 10.3758/APP.72.3.561
- Grube, M., Cooper, F. E., Chinnery, P. F., and Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11597–11601. doi: 10.1073/pnas.0910473107
- Gu, B.-M., van Rijn, H., and Meck, W. H. (2015). Oscillatory multiplexing of neural population codes for interval timing and working memory. *Neurosci. Biobehav. Rev.* 48, 160–185. doi: 10.1016/j.neubiorev.2014.10.008
- He, J., Hashikawa, T., Ojima, H., and Kinouchi, Y. (1997). Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *J. Neurosci.* 17, 2615–2625.
- Heron, J., Aaen-Stockdale, C., Hotchkiss, J., Roach, N. W., McGraw, P. V., and Whitaker, D. (2012). Duration channels mediate human time perception. *Proc. R. Soc. B Biol. Sci.* 279, 690–698. doi: 10.1098/rspb.2011.1131
- Hu, Y., Allen, R. J., Baddeley, A. D., and Hitch, G. J. (2016). Executive control of stimulus-driven and goal-directed attention in visual working memory. *Atten. Percept. Psychophys.* 78, 2164–2175. doi: 10.3758/s13414-016-1106-7
- Hu, Y., Hitch, G. J., Baddeley, A. D., Zhang, M., and Allen, R. J. (2014). Executive and perceptual attention play different roles in visual working memory: evidence from suffix and strategy effects. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1665–1678. doi: 10.1037/a0037163
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590
- Jazayeri, M., and Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Curr. Biol.* 25, 2599–2609. doi: 10.1016/j.cub.2015.08.038
- Jones, M. R., and Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Mem. Cognit.* 19, 8–20. doi: 10.3758/BF03198492
- Joseph, S., Iverson, P., Manohar, S., Fox, Z., Scott, S. K., and Husain, M. (2015). Precision of working memory for speech sounds. *Q. J. Exp. Psychol.* 68, 2022–2040. doi: 10.1080/17470218.2014.1002799
- Joseph, S., Teki, S., Kumar, S., Husain, M., and Griffiths, T. D. (2016). Resource allocation models of auditory working memory. *Brain Res.* 1640(Part B), 183–192. doi: 10.1016/j.brainres.2016.01.044
- Kleinman, M. R., Sohn, H., and Lee, D. (2016). A two-stage model of concurrent interval timing in monkeys. *J. Neurophysiol.* 116, 1068–1081. doi: 10.1152/jn.00375.2016
- Kool, W., Conway, A. R. A., and Turk-Browne, N. B. (2014). Sequential dynamics in visual short-term memory. *Atten. Percept. Psychophys.* 76, 1885–1901. doi: 10.3758/s13414-014-0755-7
- Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., et al. (2013). Resource allocation and prioritization in auditory working memory. *Cogn. Neurosci.* 4, 12–20. doi: 10.1080/17588928.2012.716416
- Lake, J. I., LaBar, K. S., and Meck, W. H. (2014). Hear it playing low and slow: how pitch level differentially influences time perception. *Acta Psychol. (Amst.)* 149, 169–177. doi: 10.1016/j.actpsy.2014.03.010
- Large, E. (2002). Perceiving temporal regularity in music. *Cogn. Sci.* 26, 1–37. doi: 10.1016/S0364-0213(01)00057-X
- LaRocque, J. J., Eichenbaum, A. S., Starrett, M. J., Rose, N. S., Emrich, S. M., and Postle, B. R. (2014). The short- and long-term fates of memory items retained outside the focus of attention. *Mem. Cognit.* 43, 453–468. doi: 10.3758/s13421-014-0486-y
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., and Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24, 61–79. doi: 10.1162/jocn_a_00140
- Matthews, W. J., and Meck, W. H. (2016). Temporal cognition: connecting subjective time to perception, attention, and memory. *Psychol. Bull.* 142, 865–907. doi: 10.1037/bul0000045
- McAuley, J. D. (2010). “Tempo and rhythm,” in *Music Perception*, eds M. Riess Jones, R. R. Fay, and A. N. Popper (New York, NY: Springer), 165–199.
- McAuley, J. D., and Henry, M. J. (2010). Modality effects in rhythm processing: auditory encoding of visual rhythms is neither obligatory nor automatic. *Atten. Percept. Psychophys.* 72, 1377–1389. doi: 10.3758/APP.72.5.1377
- Morrison, A. B., Conway, A. R. A., and Chein, J. M. (2014). Primacy and recency effects as indices of the focus of attention. *Front. Hum. Neurosci.* 8:6. doi: 10.3389/fnhum.2014.00006
- Morrone, M. C., Ross, J., and Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nat. Neurosci.* 8, 950–954. doi: 10.1038/nn1488
- Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 411–421. doi: 10.1037/0278-7393.28.3.411
- Oprisan, S. A., and Buhusi, C. V. (2014). What is all the noise about in interval timing? *Philos. Trans. R. Soc. B Biol. Sci.* 369:20120459. doi: 10.1098/rstb.2012.0459
- Öztekin, I., Davachi, L., and McElree, B. (2010). Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single store. *Psychol. Sci.* 21, 1123–1133. doi: 10.1177/0956797610376651
- Pearson, B., Raškevičius, J., Bays, P. M., Pertzov, Y., and Husain, M. (2014). Working memory retrieval as a decision process. *J. Vis.* 14, 2. doi: 10.1167/14.2.2
- Penhune, V. B., Zatorre, R. J., and Evans, A. C. (1998). Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J. Cogn. Neurosci.* 10, 752–765. doi: 10.1162/089892998563149
- Penney, T. B., Yim, E. N. K., and Ng, K. K. (2014). Distractor expectancy effects on interval timing. *Timing Time Percept.* 2, 1–19. doi: 10.1163/22134468-00002010

- Postle, B. R. (2015). The cognitive neuroscience of visual short-term memory. *Curr. Opin. Behav. Sci.* 1, 40–46. doi: 10.1016/j.cobeha.2014.08.004
- Postman, L., and Phillips, L. W. (1965). Short-term temporal changes in free recall. *Q. J. Exp. Psychol.* 17, 132–138. doi: 10.1080/17470216508416422
- Povel, D.-J., and Essens, P. (1985). Perception of Temporal Patterns. *Music Percept. Interdiscip. J.* 2, 411–440. doi: 10.2307/40285311
- Rammesayer, T. H., and Leutner, D. (1996). Temporal discrimination as a function of marker duration. *Percept. Psychophys.* 58, 1213–1223. doi: 10.3758/BF03207554
- Schneegans, S., and Bays, P. M. (2016). No fixed item limit in visuospatial working memory. *Cortex J. Devoted Study Nerv. Syst. Behav.* 83, 181–193. doi: 10.1016/j.cortex.2016.07.021
- Schubotz, R. I., Friederici, A. D., and Yves von Cramon, D. (2000). Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *Neuroimage* 11, 1–12. doi: 10.1006/nimg.1999.0514
- Schubotz, R. I., and von Cramon, D. Y. (2001). Interval and ordinal properties of sequences are associated with distinct premotor areas. *Cereb. Cortex* 11, 210–222. doi: 10.1093/cercor/11.3.210
- Shapiro, K. L., Raymond, J. E., and Arnell, K. M. (1997). The attentional blink. *Trends Cogn. Sci.* 1, 291–296. doi: 10.1016/S1364-6613(97)01094-2
- Souza, A. S., Rerko, L., and Oberauer, K. (2016). Getting more from visual working memory: retro-cues enhance retrieval and protect from visual interference. *J. Exp. Psychol. Hum. Percept. Perform.* 42, 890–910. doi: 10.1037/xhp0000192
- Teki, S., and Griffiths, T. D. (2014). Working memory for time intervals in auditory rhythmic sequences. *Audit. Cogn. Neurosci.* 5:1329. doi: 10.3389/fpsyg.2014.01329
- Teki, S., Grube, M., Kumar, S., and Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci.* 31, 3805–3812. doi: 10.1523/JNEUROSCI.5561-10.2011
- Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the “internal clock.” *Psychol. Monogr.* 77, 1–31. doi: 10.1037/h0093864
- Unsworth, N., and Engle, R. W. (2005). Individual differences in working memory capacity and learning: evidence from the serial reaction time task. *Mem. Cognit.* 33, 213–220. doi: 10.3758/BF03195310
- Vallar, G., and Papagno, C. (1986). Phonological short-term store and the nature of the recency effect: evidence from neuropsychology. *Brain Cogn.* 5, 428–442. doi: 10.1016/0278-2626(86)90044-8
- Wearden, J. H., Norton, R., Martin, S., and Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 716–729. doi: 10.1037/0096-1523.33.3.716
- Wearden, J. H., Todd, N. P. M., and Jones, L. A. (2006). When do auditory/visual differences in duration judgements occur? *Q. J. Exp. Psychol.* 59, 1709–1724. doi: 10.1080/17470210500314729
- Wimmer, K., Nykamp, D. Q., Constantinidis, C., and Compte, A. (2014). Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* 17, 431–439. doi: 10.1038/nn.3645
- Wittmann, M. (2013). The inner sense of time: how the brain creates a representation of duration. *Nat. Rev. Neurosci.* 14, 217–223. doi: 10.1038/nrn3452
- Yumoto, N., Lu, X., Henry, T. R., Miyachi, S., Nambu, A., Fukai, T., et al. (2011). A neural correlate of the processing of multi-second time intervals in primate prefrontal cortex. *PLoS ONE* 6:e19168. doi: 10.1371/journal.pone.0019168
- Zhang, W., and Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature* 453, 233–235. doi: 10.1038/nature06860
- Zokaei, N., Manohar, S., Husain, M., and Feredoes, E. (2014a). Causal evidence for a privileged working memory state in early visual cortex. *J. Neurosci.* 34, 158–162. doi: 10.1523/JNEUROSCI.2899-13.2014
- Zokaei, N., Ning, S., Manohar, S., Feredoes, E., and Husain, M. (2014b). Flexibility of representational states in working memory. *Front. Hum. Neurosci.* 8:853. doi: 10.3389/fnhum.2014.00853

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Manohar and Husain. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Isochronous Sequential Presentation Helps Children Orient Their Attention in Time

Katherine A. Johnson^{1*}, Marita Bryan¹, Kira Polonowita¹, Delia Decroupet¹ and Jennifer T. Coull²

¹ School of Psychological Sciences, University of Melbourne, Parkville, VIC, Australia, ² Laboratoire des Neurosciences Cognitives, Aix-Marseille Université, CNRS, Marseille, France

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Zhicheng Lin,
Ohio State University, USA
Steve Majerus,
University of Liège, Belgium

*Correspondence:

Katherine A. Johnson
kajo@unimelb.edu.au

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 20 June 2016

Accepted: 05 September 2016

Published: 22 September 2016

Citation:

Johnson KA, Bryan M, Polonowita K,
Decroupet D and Coull JT (2016)
Isochronous Sequential Presentation
Helps Children Orient Their Attention
in Time. *Front. Psychol.* 7:1417.
doi: 10.3389/fpsyg.2016.01417

Knowing when an event is likely to occur allows attentional resources to be oriented toward that moment in time, enhancing processing of the event. We previously found that children (mean age 11 years) are unable to use endogenous temporal cues to orient attention in time, despite being able to use endogenous spatial cues (arrows) to orient attention in space. Arrow cues, however, may have proved beneficial by engaging exogenous (automatic), as well as endogenous (voluntary), orienting mechanisms. We therefore conducted two studies in which the exogenous properties of visual temporal cues were increased, to examine whether this helped children orient their attention in time. In the first study, the location of an imperative target was predicted by the direction of a left or right spatial arrow cue while its onset was predicted by the relative duration of a short or long temporal cue. To minimize the influence of rhythmic entrainment in the temporal condition, the foreperiod (500 ms/1100 ms) was deliberately chosen so as not to precisely match the duration of the temporal cue (100 ms/400 ms). Targets appeared either at cued locations/onset times (valid trials) or at unexpected locations/onset times (invalid trials). Adults' response times were significantly slower for invalid versus valid trials, in both spatial and temporal domains. Despite being slowed by invalid spatial cues, children (mean age 10.7 years) were unperturbed by invalid temporal cues, suggesting that these duration-based temporal cues did not help them orient attention in time. In the second study, we enhanced the exogenous properties of temporal cues further, by presenting multiple temporal cues in an isochronous (rhythmic) sequence. Again, to minimize automatic entrainment, target onset did not match the isochronous interval. Children (mean age 11.4 years), as well as adults, were now significantly slowed by invalid cues in both the temporal and spatial dimension. The sequential, as opposed to single, presentation of temporal cues therefore helped children to orient their attention in time. We suggest that the exogenous properties of sequential presentation provide a temporal scaffold that supports the additional attentional and mnemonic requirements of temporal, as compared to spatial, processing.

Keywords: temporal attention, spatial attention, rhythm, temporal prediction, temporal expectation, exogenous attention

INTRODUCTION

Our attentional system allows us to filter distracting stimuli to efficiently process relevant information. Moreover, we can flexibly direct (or “orient”) attentional resources to a stimulus appearing at a specific location in space (Posner, 1980) or moment in time (Coull and Nobre, 1998), allowing individuals to better process that stimulus and so respond to the environment in an appropriate manner (Coull, 2009; Carrasco, 2011). The ability to orient attention in space has a long and considerable research history, with the research of Wundt, James, and Helmholtz in the 19th century (Carrasco, 2011) providing a platform for more recent experimental research (Posner, 1980; Enns and Brodeur, 1989; Corbetta and Shulman, 2002; Posner, 2016). The ability to orient attention in time using similar experimental paradigms is a newer research area, with the vast majority of research being conducted with adult participants (Coull and Nobre, 1998; Correa et al., 2004; Nobre et al., 2007). Nevertheless, a previous developmental study of our own, which examined attention in both time and space with valid and invalid cues, has shown that children have difficulties using a symbolic cue to voluntarily orient their attention in time, but can use an arrow cue to orient their attention in space (Johnson et al., 2015). The spatial properties of the arrow space cue, however, may have induced automatic, in addition to voluntary, attentional mechanisms (Ristic and Kingstone, 2009), making the time and space cues unbalanced. Two studies are presented in which we manipulated the temporal properties of the time cues to test whether this would help children to orient their attention in time.

In the spatial domain, two partially segregated attentional orienting systems in the human brain – the exogenous and endogenous systems – have been extensively described (Corbetta et al., 2002). The exogenous spatial orienting system directs attention automatically to salient stimuli in the environment. In a typical exogenous spatial orienting task, a cue is presented in a peripheral location. A target then appears, either in the same location as the cue (valid trial) or in the non-cued location (invalid trial), and the participant must respond as quickly as possible to the appearance of the target (Posner, 1980). Responses are quicker to targets appearing in the cued than the invalidly cued spatial location. This “validity effect” is observed even when the location of the cue does not necessarily predict where the target will eventually appear (Ristic et al., 2002; Tipples, 2002), attesting to the automatic nature of the spatial orienting mechanism. The endogenous orienting system, on the other hand, directs attention through goal-directed voluntary mechanisms, based on knowledge and expectation (Corbetta et al., 2002). In a typical endogenous spatial orienting paradigm, attention is voluntarily directed to one location or another in response to symbolic or abstract centrally presented cues (e.g., arrows) that provide information about where the upcoming target is likely to appear (Posner, 1980). Again, RTs are faster to targets appearing in validly cued, rather than invalidly cued, locations. While the exogenous orienting system operates from infancy (Harmon et al., 1994; Richards, 2000), the use of endogenous cues to voluntarily orient attention in space develops later, around 6–8 years (Pearson and Lane, 1990;

Rueda et al., 2004; Wainwright and Bryson, 2005; Iarocci et al., 2009).

The endogenous spatial orienting paradigm has now been adapted for the temporal domain, with symbolic temporal cues predicting when, rather than where, the target would appear (Coull and Nobre, 1998). Temporally predictive cues allowed attention to be endogenously oriented toward the predicted moment in time, with RTs being faster for validly, rather than invalidly, cued targets (Coull et al., 2000; Correa et al., 2004, 2010). Despite the growing number of studies in adults, only a handful of studies have investigated whether children can use endogenous temporal cues to orient their attention in time. Two very recent studies by Mento and Tarantino (2015) and Mento and Vallesi (2016) have shown that children (aged 6–12) can use symbolic temporal cues to endogenously orient their attention in time. When we compared spatial and temporal orienting directly in our own study, however, children (average age 11 years) had difficulties using a symbolic cue to orient their attention in time, even though they could use an arrow cue to orient their attention in space (Johnson et al., 2015). We concluded that the development of endogenous temporal orienting lags behind that of endogenous spatial orienting. The symbolic spatial cues used in the Johnson et al. (2015) study were left or right-facing arrows. Arrows, though symbolic and therefore assumed to orient spatial attention endogenously, have actually been shown to also induce exogenous orienting of attention (Eimer, 1997; Tipples, 2002; Ristic and Kingstone, 2009; Olk, 2014). The relative benefits of spatial, over temporal, cues in Johnson et al.’s (2015) study might have been due, therefore, to the additional exogenous orienting mechanisms induced by the arrow stimuli. By contrast, Johnson et al.’s (2015) temporal cues were more purely endogenous in nature: short and long lines represented short or long temporal intervals. There is evidence that children as young as 4 or 5 represent time in spatial terms, with the spatial length of a stimulus biasing temporal estimates of its duration (Casasanto et al., 2010; Bottini and Casasanto, 2013). It is therefore possible that short/long lines could induce exogenous, as well as endogenous, orienting of attention to short/long temporal intervals. This hypothesis has not yet, to our knowledge, been formally tested. In the absence of evidence, the possibility remains that the developmental lag for temporal, versus spatial, orienting reported by Johnson et al. (2015) was not due to the spatial versus temporal nature of the cues, but rather to their differential capacity for inducing exogenous and endogenous attentional mechanisms.

We therefore decided to compare temporal and spatial orienting in children using symbolic cues that were hypothesized to induce exogenous, as well as endogenous, mechanisms in both temporal and spatial domains. Just as the physical location of a stimulus can act as an exogenous spatial cue, the physical duration of a stimulus can act as an exogenous temporal cue. Indeed the constant duration of the intervals delineating an isochronous (rhythmic) stimulus sequence guides temporal attention to moments in time that are in phase (on beat) with the temporal structure of the sequence, without the need for attentional instruction (Klein and Jones, 1996; Large and Jones, 1999; Jones et al., 2002; Rohenkohl et al., 2011; Sanabria et al.,

2011; Triviño et al., 2011; Bolger et al., 2014). The temporal predictability of isochronous sequences improves sensorimotor processing of events occurring in phase with the rhythm, both enhancing perceptual sensitivity (Barnes and Jones, 2000; Jones et al., 2002; Lange et al., 2003; Morillon et al., 2014) and speeding visual target detection (Coull and Nobre, 2008; Bolger et al., 2014). As such, the temporal properties of stimulus presentation may be considered as an exogenous temporal cue. Like exogenous spatial cues, exogenous temporal cues can orient attention automatically. For example, Sanabria et al. (2011) showed that RTs to targets presented in phase with the temporal rhythm were faster even though the target was equally likely to appear out of phase with the rhythm (Sanabria et al., 2011). Similarly, responses were faster to in-phase targets even when participants were not required to attend to the rhythmic sequence (Rohenkohl et al., 2011; Breska and Deouell, 2014) or when participants had to simultaneously perform a demanding secondary working memory task (de la Rosa et al., 2012; Cutanda et al., 2015).

Like exogenous spatial orienting, exogenous temporal orienting also appears to operate from infancy. Infants detect unexpected temporal patterns (Brannon et al., 2004, 2007; vanMarle and Wynn, 2006), and can use rhythm to create temporal expectancies (Haith et al., 1988; Colombo and Richman, 2002; Philips-Silver and Trainor, 2005; Bergeson and Trehub, 2006; Werner et al., 2009; Winkler et al., 2009; Brandon and Saffran, 2011). These studies suggest that infants have temporally predictive information processing capabilities (Trainor, 2012). From around the age of 4 years, children can tap in time with isochronous, rhythmic, and musical sequences, and can discriminate between the tempos of two drum sequences (Drake et al., 2000). Moreover, older children (around 11 years) can use temporal patterns to predict in time when a target will appear, allowing them to respond more quickly to that target (Durstun et al., 2007). In sum, children are able to build up a temporal expectancy based on the temporal properties of sensory input, which directs their attention exogenously in time.

The two new studies reported here were designed to test the prediction that children would be able to benefit from a temporal cue when the cue conveyed temporal information in a more exogenous manner. In the first study, temporal information was conveyed by the actual presentation duration of the temporal cue: the cue was presented for either 100 or 400 ms to indicate that the target would appear after a short (500 ms) or a long (1100 ms) delay. In the second study, temporal information was again conveyed by the duration of the cue, though this time the cue was presented five times in a row to reinforce the representation of cue duration.

STUDY 1: TEMPORAL INFORMATION CONVEYED BY CUE DURATION

In the Johnson et al. (2015) study both the space and time cues were presented for 100 ms, a presentation time typical of these types of cognitive studies (Coull and Nobre, 1998). The presentation duration of the cue, however, can itself convey temporal information about when the target is expected to

appear in a bottom-up, or exogenous, manner. Presenting the temporal cue for either a long (e.g., 400 ms) or a short (e.g., 100 ms) duration to reflect the upcoming period between the cue and the target's appearance (the foreperiod, "FP") may help the participant to extract temporally pertinent information from the cue's physical appearance. While the temporal cue used in this study retained the visual qualities of our previous paradigm (Johnson et al., 2015) – it was an abstract symbol comprised of thick or thin lines that predicted a long (1100 ms) or short (500 ms) FP, respectively – we additionally incorporated temporal information in a more exogenous manner, by manipulating the duration of cue presentation. It is important to note that the presentation duration of the temporal cue was not equal to the upcoming FP; participants had to estimate the *relative* duration of the cue (short/long) and apply this information to a new set of timing parameters, in order to predict when (soon/late) the target would occur.

Children in the age range tested in this study are able to estimate the duration of stimuli as accurately as adults (Droit-Volet and Wearden, 2001; McCormack et al., 2005; Droit-Volet and Coull, 2015; Droit-Volet, 2016). For instance, 10-year-old children performed similarly to adults on two variants of the temporal generalization task, in which participants were presented with a pair of stimuli and asked to judge whether they were of the same duration (McCormack et al., 2005). In the temporal bisection task participants are trained to recognize two stimulus durations as either "short" or "long," and are then tested on a range of probe durations and asked to decide whether a probe is either short or long. Children as young as 3 years of age are able to complete this task with orderly data, demonstrating an ability to process temporal information albeit less accurately than 5- or 8-year-olds (Droit-Volet and Wearden, 2001), while 10-year-olds perform as well as adults (Droit-Volet and Coull, 2015). We were therefore confident that the 10- to 12-year-olds in our study would be able to accurately time the duration of the time cue.

The aim of study 1 was to investigate if children would be able to use the physical duration of the time cue to predict when the target would appear so as to speed response times (RTs). The hypothesis was that children and adults would show the validity effect for both the time and space cues, with faster responses to validly versus invalidly cued trials.

Materials and Methods

Participants

Thirty-three typically developing children (20 female) and 30 adults (20 female) participated in the study. Thirteen children (eight female) were excluded as they made over 50 omission errors on the task, suggesting task disengagement. Please note that the outcome of the results remained the same when these participants were included in the sample. The final sample consisted of 20 children (12 female) and 30 adults (20 female). The children ranged in age from 10 to 12 years (mean 10.7, *SD* 0.8); the adults ranged in age from 18 to 23 years (mean 19.2, *SD* 1.3). The children were recruited from two primary schools in Melbourne, Victoria. The adults were recruited from

the University of Melbourne first year cohort of Psychology students via a Research Experience Program, for which they received course credit. Children's estimated full-scale intelligence quotient (IQ) was calculated using the WISC-IV (Wechsler, 2004); 16 children completed the four subtest assessment using Block Design, Similarities, Digit Span and Coding, whilst four children completed the two subtest assessment using Block Design and Vocabulary. The children's estimated full scale IQs were calculated using Sattler's method (Sattler and Dumont, 2004), and all scored above 70 (mean 105, *SD* 10, range 81–121).

The University of Melbourne Human Research Ethics Committee and the Catholic Education Office in the Archdiocese of Melbourne approved the study, in accordance with the 1964 Declaration of Helsinki. Parents and children provided written informed consent prior to each child's participation in the study. Adult participants provided written informed consent prior to the study.

Experimental Task

All participants completed a modified version of the spatial and temporal orienting task (Coull and Nobre, 1998), which was presented using E-prime Software (Psychology Software Tools) on a 15-inch laptop computer. The modification comprised the use of a duration-based temporal cue. Participants were presented with a central stimulus display containing a central diamond and two peripheral boxes (**Figure 1**). The participants were asked to maintain their gaze on the central stimulus and use the information presented there to help predict the appearance of the upcoming target, an 'x', in one of the two peripheral boxes. The aim of the participants was to respond as quickly as possible to the appearance of the target, by pressing the down arrow on the computer keyboard. The participants simply needed to detect the appearance of the target.

Three Cue conditions were presented in separate blocks, the order of which was counterbalanced across participants (**Figure 1**). Within a trial, participants were initially exposed to the background stimulus display for a 600, 700, 800, 900, or 1000 ms inter-trial interval, randomized across trials. In the space condition, the line comprising the left or right side of the central stimulus thickened slightly for 250 ms, indicating the likely appearance of the target in the left or right peripheral box, respectively. In the time condition, the outline of the central stimulus thickened either very slightly for 100 ms or to a much greater extent for 400 ms, indicating that the target was likely to appear soon (500 ms FP) or later (1100 ms FP), respectively. In the neutral condition, the outline of the central stimulus thickened slightly for 250 ms but did not provide any specific information about the likely location or FP of the target and so simply alerted the participant to the upcoming target. For all Cue conditions the background stimulus display then remained unchanged for a FP of either 500 or 1100 ms. The timing of the FP started with the offset of the cue. The target then appeared in either the left or right peripheral box for 100 ms. Following target presentation, the background stimulus display was shown for 1500 ms, to allow for participant's responses, before the next trial commenced.

For the space and time conditions, 32 valid, 8 invalid, and 4 catch trials (44 trials in total) were presented in each of three consecutive blocks (132 trials per condition). For the neutral condition, 16 trials were presented in each of three consecutive blocks (48 trials). Prior to each block, participants were informed of the nature of the cue in the upcoming block. Each block lasted for between 2 to 3 min and participants were able to take rest breaks between blocks. The whole task, with breaks, took approximately 20–25 min.

Participants were provided with a training set of 32 valid trials for the space and time conditions, and 16 trials for the neutral condition, prior to the experimental session. This was to ensure they understood the instructions and, for the time condition, to learn the association between the duration cues and the short and long FPs. The spatial and temporal cues were trained for the same number of trials so that they were subject to the same degree of learning-induced transfer from endogenous to exogenous attention control (Lin et al., 2016). The participants were asked to identify the meaning of each of the cues to ensure understanding of the cues. They were reminded to respond to target appearance as quickly as possible.

Procedure

The children were tested in a quiet setting at their schools. The adults were tested in a quiet testing room in the School of Psychological Sciences at the University of Melbourne.

Data Analysis

RTs of less than 100 ms (errors of omission, extremely fast RTs) were excluded from the RT analyses (see **Table 1** for a count of omission errors). Any RTs to the catch trials were also excluded. For each participant, the mean RT was calculated per trial type, and group means (*M*) and standard deviations (*SDs*) were then calculated. The data were normally distributed.

Statistics

Statistical analysis was carried out using IBM SPSS software version 23. The validity effect was investigated with a three-way mixed factorial ANOVA with Group (adults, children) as a between-subjects factor and Cue type (space, time) and Validity (valid, invalid) as within-subjects factors. The validity effect was calculated using data from the 500 ms FP trials only to avoid confounding the temporal validity effect with Variable FP effects. In cued RT paradigms, the probability of target appearance increases with the length of the FP – the “Hazard Function” (Luce, 1986) – which leads to faster RTs at longer FPs (Woodrow, 1914) – the “Variable FP effect” (Niemi and Näätänen, 1981). In temporal orienting paradigms, the RT benefits of the Hazard Function render the RT benefits of the temporally valid cue negligible at long FPs (Coull and Nobre, 1998; Coull et al., 2000; Vallesi et al., 2013). To obtain a clean measure of temporal orienting effects, we therefore constrained our analysis to the short (500 ms) data-point (Rohenkohl et al., 2011, for a similar approach). Refer to **Table 2** for data for the 1100 ms condition.

Data from the neutral condition was used to calculate the Variable FP effect and the sequential effect. The sequential effect reflects the fact that a participant's RT to the upcoming

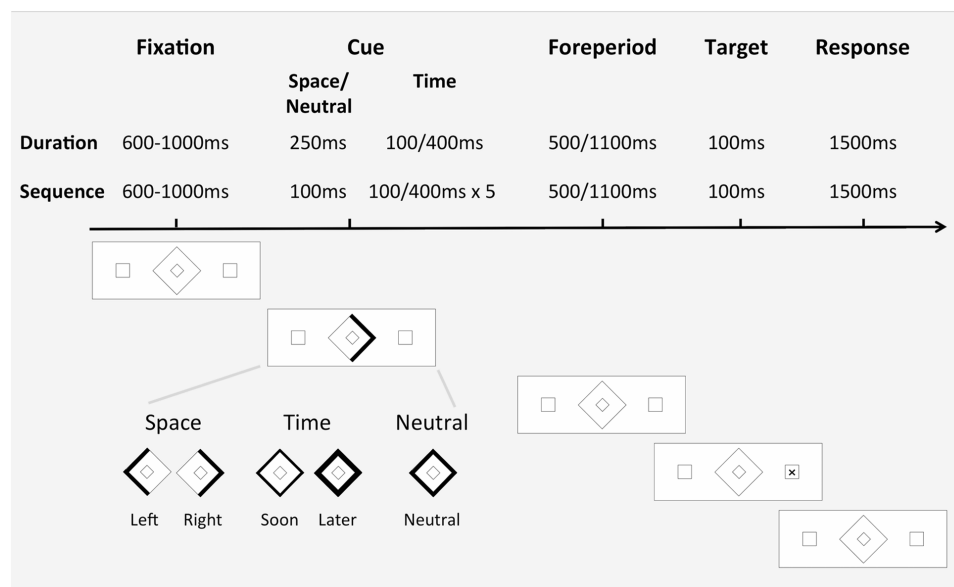


FIGURE 1 | Sequence of events in one trial (right spatial, valid) and examples of the Cue stimuli used in the three orienting conditions. In the Space trials, either the left- or right-hand side of the stimuli thickened slightly, indicating the likely appearance of the target in the left or right peripheral box, respectively. In the neutral condition, the outline of the central stimulus thickened slightly but did not provide any specific information about the likely location or FP of the target and so simply alerted the participant to the upcoming target. In Study 1 in the duration time condition, the outline of the central stimulus thickened either very slightly for 100 ms or to a much greater extent for 400 ms, indicating that the target was likely to appear soon (500 ms FP) or later (1100 ms FP), respectively. In Study 2 in the sequential time condition, for the short FP the background stimulus display was shown for 100 ms and then the outline of the central stimulus thickened very slightly for 100 ms. This off/on cycle occurred five times in a row, to indicate that the target was likely to appear soon. For the long FP, the background stimulus display appeared for 100 ms, then a thick outline of the central stimulus appeared for 400 ms. This off/on cycle occurred five times in a row, signaling that the target was going to appear later.

target depends on the duration of the FP of the previous trial (Woodrow, 1914; Karlin, 1959; Baumeister and Joubert, 1969). Responses are slower on short FP trials when the previous trial had a long FP; in contrast responses for long FP trials are not influenced by the previous trial's FP (Los and van den Heuvel, 2001; Los and Heslenfeld, 2005). The sequential effect is thought to be an automatic form of temporal prediction (Los and van den Heuvel, 2001; Triviño et al., 2011; Vallesi et al., 2013, 2014). The variable FP and sequential effects were investigated with a three-way mixed factorial ANOVA involving Group (adults, children), FP of the current trial, i.e., $FP(n)$ (500 ms, 1100 ms), and FP of the previous trial, i.e., $FP(n-1)$ (500 ms, 1100 ms). The FP and sequential effects were investigated using the neutral trials only, to avoid confounds from any effects associated with the space and time cues. The alpha level was set at 0.05 and Bonferroni-adjustments were made for pair-wise comparisons.

Results

Spatial and Temporal Validity effects

Significant Group $F(1,48) = 8.648$, $p = 0.005$, $\eta_p^2 = 0.153$, Cue $F(1,48) = 13.354$, $p = 0.001$, $\eta_p^2 = 0.218$, and Validity $F(1,48) = 57.741$, $p < 0.001$, $\eta_p^2 = 0.546$ main effects were further explained by a significant Group by Cue by Validity interaction, $F(1,48) = 5.828$, $p = 0.020$, $\eta_p^2 = 0.108$ (Figure 2; Table 2). This was broken down by Group. For the adults,

there was a significant Cue main effect, $F(1,29) = 15.137$, $p = 0.001$, $\eta_p^2 = 0.343$, such that adults responded significantly more quickly to the time than space cues. There was also a significant Validity main effect, $F(1,29) = 57.627$, $p < 0.001$, $\eta_p^2 = 0.665$, whereby adults responded significantly more quickly to the valid than invalid trials. There was no significant Cue by Validity interaction, $F(1,29) = 0.431$, $p = 0.517$, $\eta_p^2 = 0.015$, suggesting valid cues were equally beneficial in space and time. For the children, on the other hand, there was a significant Cue by Validity interaction, $F(1,19) = 15.379$, $p = 0.001$, $\eta_p^2 = 0.447$. On the space trials, children responded significantly more quickly to the valid than invalid trials, $p < 0.001$. On the time trials, however, there was no significant difference in MRT between the valid and invalid trials, $p = 0.988$. On valid trials there was no significant difference in MRT between the space and time cues, $p = 0.519$. On invalid trials, the children responded significantly more slowly to the space cues than the time cues, $p = 0.014$.

Sequential and Variable Foreperiod Effects

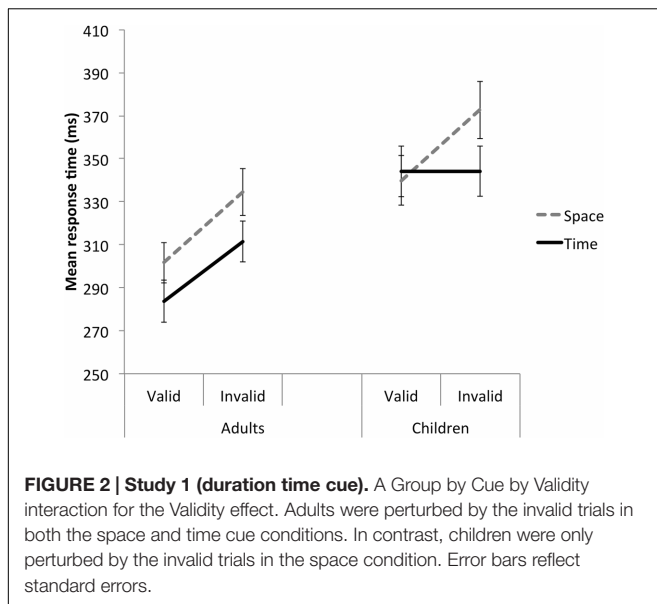
Significant FP, $F(1,48) = 5.711$, $p = 0.021$, $\eta_p^2 = 0.106$, and $FP(n-1)$, $F(1,48) = 21.774$, $p < 0.001$, $\eta_p^2 = 0.312$, were further explained by a significant FP by $FP(n-1)$ interaction, $F(1,48) = 19.422$, $p < 0.001$, $\eta_p^2 = 0.288$ (Figure 3; Table 3). At the 500 ms FP, participants were significantly faster to respond to targets when the previous trial's FP was also 500 ms rather than

TABLE 1 | The median and interquartile range (in parentheses) measures of the number of omission errors, for the Adult and Child groups for Study 1 (duration time cue) and Study 2 (sequence time cue) on the various levels of the Cue (space, time), Validity (valid, invalid), and Foreperiod (500 ms, 1100 ms) independent variables of the spatial and duration temporal orienting task.

Study	Group	Space Valid 500 ms	Space Invalid 500 ms	Space Valid 1100 ms	Space Invalid 1100 ms	Time Valid 500 ms	Time Invalid 500 ms	Time Valid 1100 ms	Time Invalid 1100 ms	Neutral 500 ms	Neutral 1100 ms
Duration	Adults	0.5 (2)	0.0 (0)	1.0 (3)	0.0 (1)	1.0 (2)	0.0 (1)	1.0 (2)	0.0 (0)	0.0 (1)	0.0 (2)
Duration	Children	3.5 (3)	1.0 (2)	4.0 (7)	1.0 (2)	2.0 (4)	0.0 (1)	2.5 (4)	1.5 (3)	1.0 (2)	1.0 (5)
Sequence	Adults	0.0 (2)	0.0 (0)	1.0 (1)	0.0 (1)	0.0 (1)	0.0 (0)	0.0 (1)	0.0 (1)	0.0 (0)	0.0 (1)
Sequence	Children	2.0 (2)	1.0 (2)	2.0 (4)	0.0 (1)	3.0 (4)	0.0 (1)	1.0 (3)	0.0 (1)	0.0 (1)	1.0 (4)

TABLE 2 | Mean and standard deviation (in parentheses) measures of response time, in milliseconds, for the Adult and Child groups for Study 1 (duration time cue) and Study 2 (sequence time cue) on the various levels of the Cue (space, time), Validity (valid, invalid), and Foreperiod (500 ms, 1100 ms) independent variables of the spatial and duration temporal orienting task.

Study	Group	Space Valid 500 ms	Space Invalid 500 ms	Space Valid 1100 ms	Space Invalid 1100 ms	Validity effect	Time Valid 500 ms	Time Invalid 500 ms	Validity effect	Time Valid 1100 ms	Time Invalid 1100 ms	Validity effect
Duration	Adults	302 (48)	334 (53)	291 (38)	315 (42)	32	284 (54)	312 (51)	28	296 (42)	304 (48)	8
Duration	Children	340 (57)	373 (69)	340 (51)	355 (59)	33	344 (52)	344 (54)	0	337 (46)	338 (51)	1
Sequence	Adults	304 (44)	336 (55)	296 (35)	318 (44)	32	288 (51)	333 (51)	35	284 (35)	313 (37)	29
Sequence	Children	352 (65)	407 (80)	333 (58)	356 (60)	54	335 (77)	416 (82)	81	343 (80)	357 (79)	14



1100 ms, $p < 0.001$, reflecting the sequential effect. At 1100 ms FP, there was no significant difference in RT between the FP($n - 1$) 500 ms and FP($n - 1$) 1100 ms trials, $p = 0.274$, reflecting the asymmetrical nature of the sequential effect. When the previous trial's FP was short, participants responded with similar RTs between the 500 and 1100 ms FP trials, $p = 0.749$. When the previous trial's FP was long, participants were significantly slower to respond to the target at 500 ms FP compared with the 1100 ms FP trials, $p < 0.001$.

The adults performed the task with significantly faster MRT than the children, $F(1,48) = 11.313$, $p = 0.002$, $\eta_p^2 = 0.191$. There were no other significant interactions.

Discussion

As expected, adults' attention was guided by both the time and space cues: participants responded with a significantly slower MRT to invalidly cued targets in both dimensions. The results, however, failed to support our hypothesis that manipulating the duration of the temporal cue would enable temporal orienting in children. We found that children were not significantly perturbed by the invalid time cues, suggesting that they were not using the duration cue to anticipate when the target would appear. In contrast, presenting the children with the invalid space cue did result in a significantly longer RT compared to the valid space cue, suggesting the children were using the space cues to anticipate *where* the target would appear. Moreover, the children demonstrated the sequential and variable FP effects, supporting previous results (Vallesi and Shallice, 2007; Johnson et al., 2015; Mento and Tarantino, 2015), indicating that children implicitly processed the temporal information available in the trial structure. Their responses were faster to the targets in the long FP trials, reflecting the variable FP effect, and their responses were slower when the preceding trial's cue-target interval was longer than that of the current trial, reflecting the sequential effect.

Indeed, our data further suggest that these effects were even stronger in children compared with the adults (Johnson et al., 2015).

Overall, our results suggest that the temporal information conveyed by the duration-based time cue was not enough to help children anticipate when the target would appear. Their performance was very similar to that of the children in the Johnson et al. (2015) study. Although children in this age range can estimate stimulus duration as well as adults (McCormack et al., 2005; Droit-Volet and Coull, 2015; Droit-Volet, 2016), it appears that they are not yet able to use this information to make temporal predictions in order to optimize behavior. Yet previous studies have shown that even young infants derive temporal expectations from isochronous or rhythmic sequences of stimuli (Haith et al., 1988; Colombo and Richman, 2002; Philips-Silver and Trainor, 2005; Werner et al., 2009; Winkler et al., 2009; Brandon and Saffran, 2011). In Study 2 therefore, we further increased the exogenous temporal information conveyed by the temporal cue by presenting cues in an isochronous sequence. We aimed to test whether this would help children of 10–12 years to orient attention in time in order to speed responding to a temporally predictable event.

STUDY 2: TEMPORAL INFORMATION CONVEYED BY AN ISOCHRONOUS SEQUENCE

In this study, we presented the duration-based temporal cue several times in an isochronous sequence, in order to enhance the temporal properties of the cue. Many previous studies in adults have shown that the variability of duration estimates decreases with the number of stimulus repetitions (Keele et al., 1989; Schulze, 1989; Drake and Botte, 1993; Ivry and Hazeltine, 1995; Grondin et al., 2001; Merchant et al., 2008; Grondin, 2012). We therefore hypothesized that presenting our duration-based time cue multiple times would help children form a more robust temporal memory of cue duration, which would then help them form temporal predictions concerning target onset time. In the current study, five repetitions of the temporal cue were presented in quick succession, with either a 100 ms short on and 100 ms off cycle, or a 400 ms long on and 100 ms off cycle, indicating that the upcoming FP would be short (500 ms) or long (1100 ms), respectively. Importantly, the time of target onset was not in phase with the preceding rhythm: instead, as in Study 1, participants had to extract the relative duration of the cue and extrapolate it to a new set of timing parameters, in order to predict when the target would occur. In this way, the isochronous time cue was not simply entraining temporal attention in a purely exogenous manner (Klein and Jones, 1996; Large and Jones, 1999; Jones et al., 2002). Sanabria et al. (2011, Experiment 3) have already shown that adult participants extrapolate the temporal information provided by non-predictive rhythmic sequences to non-matching FPs in order to speed responding (Sanabria et al., 2011). The aim of study 2 was to investigate whether children could similarly use the temporal information provided by an isochronous sequence to anticipate

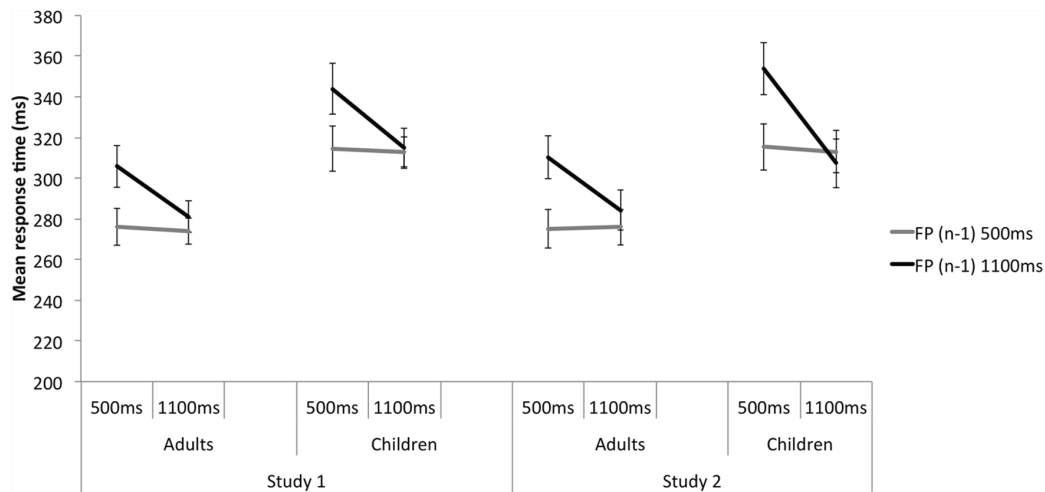


FIGURE 3 | A significant $FP(n) \times FP(n-1)$ interaction in the Neutral trials for Study 1 on the left and Study 2 on the right. This indicates the presence of the sequential effect, in both adults and children equally. Participants responded more slowly when the current trial's FP was short and was preceded by a long FP trial, compared with a preceding short FP trial. Responses did not vary significantly when the current long FP trial was preceded by a long or short FP. Results were very similar across studies 1 and 2. Error bars reflect standard errors.

TABLE 3 | Mean and standard deviation (in parentheses) measures of response time, in milliseconds, for the Adult and Child groups of Study 1 (duration time cue) and Study 2 (sequence time cue), on the Neutral trials, on the various levels of the previous trial foreperiod $FP(n-1)$ (500 ms, 1100 ms) and the current trial $FP(n)$ (500 ms, 1100 ms) of the spatial and duration temporal orienting task.

Study	Group	FP(n-1) 500 ms	FP(n-1) 1100 ms	FP(n-1) 500 ms	FP(n-1) 1100 ms
		FP(n) 500 ms	FP(n) 500 ms	FP(n) 1100 ms	FP(n) 1100 ms
Duration	Adults	276 (43)	306 (57)	274 (37)	281 (39)
Duration	Children	315 (59)	344 (55)	313 (33)	315 (47)
Sequence	Adults	275 (40)	310 (51)	276 (37)	284 (49)
Sequence	Children	315 (63)	354 (65)	313 (57)	307 (61)

when the target would appear. The space cue was the same as in Study 1 (arrows) and is consistent with our previous research (Johnson et al., 2015). Our hypothesis was that children and adults would show the validity effect for both the time and space cues.

Materials and Methods

Participants

Twenty-four typically developing children (15 female) and 31 adults (26 female) participated in the study. Four children and one adult were excluded as they made over 50 omission errors on the task, suggesting task disengagement. One adult was excluded, as her RTs were greater than 2.5 SD above the adult group. Please note that the outcome of the results remained the same when these participants were included in the sample. The final sample consisted of 20 children (13 female) and 29 adults (24 female). The children ranged in age from 10 to 12 years (mean 11.4, *SD* 0.6); the adults ranged in age from 18 to 32 years (mean 20.4, *SD* 3.3). The children were recruited from a primary school in Melbourne, Victoria. The adults were recruited from the University of Melbourne first year cohort of Psychology students via a Research Experience Program, for which they received

course credit. To ensure that the children could understand the task instructions, their estimated full-scale IQ was calculated using the WISC-IV (Wechsler, 2004); 18 children completed the four subtest assessment using Block Design, Similarities, Digit Span and Coding, whilst two children completed the two subtest assessment using Block Design and Vocabulary. The children's estimated full scale IQs were calculated using Sattler's method (Sattler and Dumont, 2004), and all scored above 70 (mean 107, *SD* 9, range 83–119).

The University of Melbourne Human Research Ethics Committee and the Catholic Education Office in the Archdiocese of Melbourne approved the study, in accordance with the 1964 Declaration of Helsinki. Parents and children provided written informed consent prior to each child's participation in the study. Adult participants provided written informed consent prior to the study.

Experimental Task

The modified version of the spatial and temporal orienting task (Coull and Nobre, 1998) used in Study 1 was further modified for Study 2. The space and neutral conditions were the same as per Study 1. In the time condition, the central stimulus was presented in an isochronous sequence (Figure 1), indicating

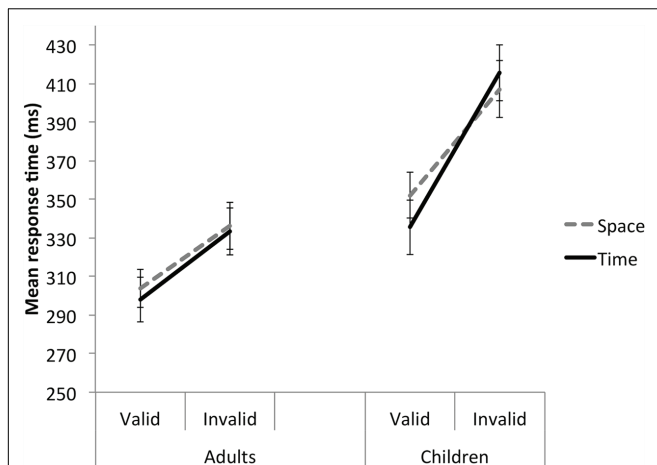


FIGURE 4 | Study 2 (sequence time cue). Both the adult and child groups were perturbed by the invalid cues, in both the temporal and spatial dimension, with a greater effect shown by the child group. Error bars reflect standard errors.

that the target was likely to appear soon (500 ms FP) or later (1100 ms FP). For the short FP, the background stimulus display was presented for 100 ms, and then the outline of the central stimulus thickened very slightly for 100 ms. This off/on cycle occurred five times in a row, to indicate that the target was likely to appear soon. For the long FP, the background stimulus display appeared for 100 ms, then a thick outline of the central stimulus appeared for 400 ms. This off/on cycle occurred five times in a row, signaling that the target was going to appear later. In the neutral condition, a slightly thickened outline of the central stimulus appeared once for 100 ms and did not provide any specific information about the likely location or FP of the target. The timing of the FP started with the offset of the cue and, in the case of the time cue, with the last stimulus of the isochronous sequence.

Within a trial, participants were initially exposed to the background stimulus display for 600, 700, 800, 900, or 1000 ms inter-trial interval, which was randomized across trials. During the spatial and neutral conditions, the cue was then presented to participants for 100 ms. During the time condition, the cue was presented in an isochronous sequence for 1 s in total for the short FP trials and 2.5 s in total for the long FP trials. For all Cue conditions the background stimulus then remained unchanged for a delay of either 500 or 1100 ms, after which the target appeared in either the left or right peripheral box for 100 ms. Following target presentation in each trial, the background stimulus display was shown for 1500 ms before the next trial commenced.

For the space and time conditions, 32 valid, 8 invalid, and 4 catch trials (44 trials in total) were presented in each of three consecutive blocks (132 trials per condition). For the neutral condition 16 trials were presented in each of three blocks (48 trials altogether). Prior to each block commencing, participants were informed of the nature of the cue in the

upcoming block. Each block lasted for between 2 to 3 min and participants were able to take rest breaks in between blocks. The whole task, with breaks, took approximately 25–30 min.

Participants were provided with a training set of 32 valid trials for the space and time conditions, and 16 trials for the neutral condition, prior to the experimental session. This was to ensure they understood the instructions and to learn the association between the isochronous sequence and the short and long FPs in the time condition. The participants were asked to identify the meaning of each of the cues to ensure understanding of the cues. They were reminded to respond to target detection as quickly as possible.

Procedure

The children were tested in a quiet setting at their school. The adults were tested in a quiet testing room in the School of Psychological Sciences at the University of Melbourne.

Data Analysis

RTs of less than 100 ms (errors of omission, extremely fast RTs) were excluded from analyses. Any RTs to the catch trials were also excluded. For each participant, the mean RT was calculated per trial type and group means (M) and standard deviations (SDs) for each trial type were calculated. The data were normally distributed.

Statistics

Statistical analysis was carried out using IBM SPSS software version 23. The validity effect was investigated with a three-way mixed factorial ANOVA with Group (adults, children) as a between-subjects factor, and Cue type (Space, Time) and Validity (valid, invalid) as within-subjects factors. As in Study 1, only the 500 ms FP trials were analysed. Please refer to **Table 2** for the 1100 ms data. The FP and sequential effects were investigated with a three-way mixed factorial ANOVA involving Group (adults, children), FP of the current trial, i.e., $FP(n)$ (500, 1100 ms), and FP of the previous trial, i.e., $FP(n - 1)$ (500 and, 1100 ms), on the neutral trials only. The alpha level was set at 0.05 and Bonferroni-adjustments were made for pair-wise comparisons.

Results

Spatial and Temporal Validity Effects

Importantly, there were no significant main effects or interactions involving Cue (**Figure 4**; **Table 2**). A Validity main effect, $F(1,47) = 103.65$, $p < 0.001$, $\eta_p^2 = 0.688$, and a Group main effect, $F(1,47) = 13.392$, $p = 0.001$, $\eta_p^2 = 0.222$, were further explained by a Validity by Group interaction, $F(1,47) = 11.341$, $p = 0.002$, $\eta_p^2 = 0.194$. Although both adults and children had significantly slower RTs to targets in the invalid than valid trials, both $p < 0.001$, Cohen's d adults 0.72, children 0.95, the interaction was most likely driven by the particularly slow responses to invalid trials made by the children, as seen in **Figure 4**.

There were no other significant main or interaction effects.

Variable Foreperiod and Sequential Effects

Significant $FP(n)$, $F(1,47) = 23.3$, $p < 0.001$, $\eta_p^2 = 0.32$, and $FP(n-1)$, $F(1,47) = 32.1$, $p < 0.001$, $\eta_p^2 = 0.41$, main effects were further explained by a significant $FP(n)$ by $FP(n-1)$ interaction, $F(1,47) = 32.17$, $p < 0.001$, $\eta_p^2 = 0.41$ (Figure 3; Table 3). RTs were significantly slower when the current short $FP(n)$ trial was preceded by a long $FP(n-1)$ trial than by a short $FP(n-1)$ trial ($p < 0.001$) – the sequential effect. In contrast, RTs did not vary significantly when a current long $FP(n)$ trial was preceded by a long or short $FP(n-1)$ trial, $p = 0.79$ – the asymmetric nature of the sequential effect. When the preceding trial was long, participants responded to the target with significantly faster MRTs at the current trial long FP compared with the short, $p < 0.001$. When the preceding trial was short, there was no significant difference in MRT between the current trial short and long FPs, $p = 0.884$.

There was a significant Group main effect, $F(1,47) = 6.7$, $p = 0.013$, $\eta_p^2 = 0.13$, with the adults responding significantly more quickly than the children. There were no interactions involving Group.

A Direct Comparison of the Validity Effect for the Temporal Cue from Studies 1 and 2

The Validity effect was compared across Studies 1 and 2, to directly test whether the provision of a duration versus a sequential cue was more beneficial in helping children (and adults) to orient attention in time. In terms of age of the two samples, there was no significant main effect of Study, $F(1,95) = 1.427$, $p = 0.235$, $\eta_p^2 = 0.015$, and no significant interaction between Group (adult, child) and Study (duration, sequence), $F(1,95) = 0.343$, $p = 0.559$, $\eta_p^2 = 0.004$. By design, there was a significant main effect of Group, $F(1,95) = 453.080$, $p < 0.001$, $\eta_p^2 = 0.827$.

A Group (adult, child) by Study (duration, sequence) by Validity (valid, invalid) three-way repeated measures ANOVA was conducted on the Temporal, 500 ms FP data. Significant Validity, $F(1,95) = 91.538$, $p < 0.001$, $\eta_p^2 = 0.491$, Study, $F(1,95) = 4.699$, $p = 0.033$, $\eta_p^2 = 0.047$, and Group, $F(1,95) = 21.863$, $p < 0.001$, $\eta_p^2 = 0.187$ main effects were further explained by a significant Validity by Study by Group interaction, $F(1,95) = 23.520$, $p < 0.001$, $\eta_p^2 = 0.198$ (Figure 5). This was broken down by Group. For the adults, there was a significant Validity main effect, with significantly faster responses to the valid than invalid trials, $p < 0.001$. There was no Study main effect and no significant Study by Validity interaction. For the children, there was a significant Study by Validity interaction, $F(1,38) = 35.037$, $p < 0.001$, $\eta_p^2 = 0.480$. For the duration study, there was no significant difference between the valid and invalid temporal trials, $p = 0.992$. For the sequence study, children were significantly faster to respond to the valid compared with invalid trials, $p < 0.001$. For valid trials, there was no significant difference in MRT between the duration and sequence studies, $p = 0.680$. For the invalid trials, in contrast, children in the sequence study were significantly slower to respond to the target than in the duration study, $p = 0.002$.

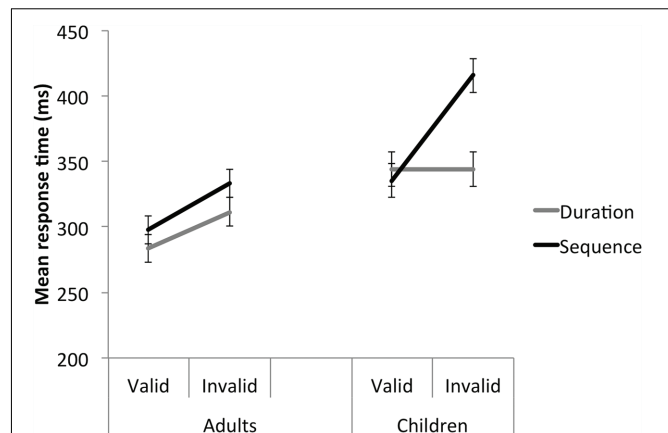


FIGURE 5 | Studies 1 and 2 – A direct comparison of the time cue trials at 500 ms FP between Studies 1 (duration time cue) and 2 (sequence time cue). For children, the absence of a response time perturbation by invalid time cues in the Duration study contrasts with a large response time perturbation in the Sequence study. Error bars reflect standard errors. Please note that the time cue data are also represented in Figures 2 and 4.

Discussion

Supporting our hypothesis, children were able to use the isochronous, visual sequence to orient their attention in time – both children and adults showed the validity effect for spatial and temporal orienting, and indeed the children showed a significantly larger temporal validity effect than the adults. There were no significant effects of Cue, suggesting that the isochronous sequence was as useful as the arrow cue in guiding the orienting of attention in time versus space, both for children and adults. Both children and adults showed the variable FP and asymmetric sequential effects, again supporting previous research. A follow-up analysis of the temporal cue trials only, directly comparing the validity effect of the two studies, confirmed the findings that the provision of the sequential cue aided children (and adults) to orient their attention in time, whereas the duration cue did not offer the same support for the children.

Isochronous sequences have attention-capturing properties (Jones et al., 1982; Breska and Deouell, 2014), orienting attention exogenously to moments in time that are in phase with the entraining rhythm. Notably, however, the rhythm of our sequence did not directly match the timing of the upcoming FP, suggesting that it did not merely rhythmically entrain participants' attention. The participants had to extract temporal information from the sequence and then extrapolate this information to the forthcoming FP, which was longer (500 ms) than the individual components (100 ms) of the sequence. It is possible, however, that the individual components of the isochronous sequences were combined by participants to produce an interval that was a harmonic of the upcoming FP. For instance, the five 100 ms cues may have been added to the five 100 ms intervals, summing to 1,000 ms, which is double the 500 ms FP. This harmonic may feasibly have helped entrain attention to

the upcoming FP. Further research is required to probe this question.

These data replicate previous findings in adults that fast and slow rhythms generate temporal expectancies that a target will appear after a short versus long delay, respectively, e.g., (Barnes and Jones, 2000; Rohenkohl et al., 2011; Sanabria et al., 2011; Breska and Deouell, 2014; Morillon et al., 2014). In this study, we have extended these findings to children aged around 11.

GENERAL DISCUSSION

To select and process important information in the environment requires a flexible cognitive system that can orient attention in both time and space. The exogenous characteristics of the time cue were emphasized firstly by using a duration-based cue, and secondly by using a sequential isochronous cue. The duration cue of Study 1 did not help children to predict target onset in order to speed responding. In contrast, the sequence cue of Study 2 did help children to orient their attention in time and they were strongly perturbed by the invalid time cues. A direct comparison of RTs in Studies 1 and 2 confirmed that sequential, rather than single, presentation of a duration cue was significantly more beneficial for performance. We found that children were able to use the physical temporal properties of sequential presentation to help estimate when the target would appear in the near future. Exaggerating the exogenous nature of the temporal cue therefore helped children orient their attention in time. To our knowledge, this is the first study showing that children can make use of isochronous rhythms to predict the time of target onset in order to speed responses to that target.

Although many prior studies have shown that children (Drake et al., 2000), and even infants (Haith et al., 1988; Colombo and Richman, 2002; Philips-Silver and Trainor, 2005; Werner et al., 2009; Winkler et al., 2009; Brandon and Saffran, 2011), *process* the temporal information inherent in isochronous rhythms, we demonstrate that children can then *use* this temporal information to guide and optimize their responses to temporally predictable events. In adults, the temporal predictability of isochronous or rhythmic sequences orient attention to moments in time that are in phase with the entraining rhythm, thereby optimizing processing of stimuli appearing at those precise moments (Doherty et al., 2005; Martin et al., 2005; Correa and Nobre, 2008; Ellis and Jones, 2010; Rohenkohl et al., 2011; Sanabria et al., 2011; de la Rosa et al., 2012; Miller et al., 2013; Cutanda et al., 2015). In our study, however, the interval before target appearance (500 ms) was not identical to that used in the isochronous sequence (100 ms). Participants had to use a *relatively* short (or long) cue to predict target onset after the relatively shorter (or longer) FP. Sanabria et al. (2011, Experiment 3) have already shown that adult participants can extrapolate the temporal information provided by rhythmic sequences to non-matching FPs (Sanabria et al., 2011). We now confirm this result in children. We suggest that our results do not simply reflect the entraining effects of an isochronous sequence, although it is possible that participants combined

the five sequential presentations of the 100 ms stimulus to cue the 500 ms FP. Further research in children, comparing the benefits of isochronous sequences on targets appearing at harmonic versus non-harmonic FPs, will be required to address this possibility.

Many previous studies have shown that multiple, as opposed to single, presentations of stimulus duration sharpen timing in motor and perceptual duration estimation tasks (Schulze, 1989; Drake and Botte, 1993; Merchant et al., 2008; Grondin, 2012). In these studies, the timing benefits of sequential presentation were measured explicitly by the accuracy and variability of duration judgments. We extend these findings in two ways: first by showing that sequential presentation improves timing as measured more implicitly by the speed of RT and, second, by demonstrating this effect in children as well as in adults. Importantly, in these previous studies, timing was improved by sequential presentation whether the test interval was contiguous (in-phase) with the reference rhythm or not (Keele et al., 1989; Ivry and Hazeltine, 1995; Grondin et al., 2001). The temporal benefits of sequential presentation were therefore unlikely to be due simply to entrainment, but rather were interpreted to reflect the construction of a more robust, accurate and less variable temporal template against which the test interval could then be compared. It is therefore possible that in our study, sequential presentation afforded a more temporally robust representation of the short or long cue in memory, which helped children to then orient their attention toward the moment in time at which the (non-contiguous) target was predicted to appear. This point is important because, as opposed to spatial processing, temporal processing depends upon a number of accessory cognitive processes, such as sustained attention and working memory (Michon, 1985; Zakay and Block, 1996; Fortin and Rousseau, 1998). To predict the moment at which the target is expected to appear, the moment of interval onset must be held in working memory and continuously compared to the currently elapsing time until the critical (predicted) time is reached. Children's timing abilities are known to correlate strongly with mnemonic and attentional capacity (Zélandi and Droit-Volet, 2011, 2012; Droit-Volet, 2013, 2016; Droit-Volet and Zélandi, 2013a,b; Droit-Volet and Coull, 2016) and, as compared to adults, their temporal sensitivity is disproportionately perturbed when their memory of the reference duration is deliberately degraded (Delgado and Droit-Volet, 2007). Therefore, the repeated, sequential presentation of the temporal cue in our study may have provided a robust temporal scaffold to counteract the additional cognitive demands of the temporal task.

In fact, this hypothesis may explain the discrepancy between the results of our own study (Johnson et al., 2015), and that of Mento and Tarantino (Mento and Tarantino, 2015). While we found that children aged 11 could not use an abstract symbolic temporal cue to orient their attention in time (Johnson et al., 2015), Mento and Tarantino (2015) found that children as young as 6 years did benefit from a symbolic temporal cue. One of the main differences between their paradigm and our own was the nature of cue presentation. In the Mento and

Tarantino study, the cue remained on the screen until target onset (600 ms short/1400 ms long), whereas in our study it was presented very briefly (100 ms) prior to an empty FP. It may be that the long presentation time of the cue in the Mento and Tarantino (2015) study provided a similarly robust temporal scaffold for interpretation of the cue.

CONCLUSION

By combining both exogenous and endogenous stimulus characteristics in the sequential time cue, similar to the combination of exogenous and endogenous features of an arrow space cue (Eimer, 1997; Tipples, 2002; Ristic and Kingstone, 2009; Olk, 2014), we found that children could successfully orient their attention in both time and space. Future research comparing spatial and temporal orienting directly should try to balance the endogenous and/or

exogenous characteristics of the temporal cue with that of the spatial cue.

AUTHOR CONTRIBUTIONS

Conception and design of the work: KJ and JC. Acquisition, analysis, and interpretation of data: KJ, MB, KP, DD, and JC. Drafting and revision of work: KJ, MB, KP, DD, and JC. Agreement to be accountable for all aspects of the work: KJ, MB, KP, DD, and JC.

ACKNOWLEDGMENTS

We would like to thank the children and teachers at several primary schools in Melbourne for their generous participation in this study.

REFERENCES

- Barnes, R., and Jones, M. R. (2000). Expectancy, attention, and time. *Cogn. Psychol.* 41, 254–311. doi: 10.1006/cogp.2000.0738
- Baumeister, A. A., and Joubert, C. E. (1969). Interactive effects on reaction time of preparatory interval length and preparatory interval frequency. *J. Exp. Psychol.* 82, 393–395. doi: 10.1037/h0028119
- Bergeson, T. R., and Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Percept.* 23, 345–360. doi: 10.1525/mp.2006.23.4.345
- Bolger, D., Coull, J. T., and Schön, D. (2014). Metrical rhythm implicitly orients attention in time, as indexed by improved target detection and left inferior parietal activation. *J. Cogn. Neurosci.* 26, 593–605. doi: 10.1162/jocn_a_00511
- Bottini, R., and Casasanto, D. (2013). Space and time in the child's mind: metaphoric or ATOMIC? *Front. Psychol.* 4:803. doi: 10.3389/fpsyg.2013.00803
- Brandon, M., and Saffran, J. R. (2011). Apparent motion enhances visual rhythm discrimination in infancy. *Atten. Percept. Psychophys.* 73, 1016–1020. doi: 10.3758/s13414-011-0106-x
- Brannon, E. M., Roussel, L. W., Meck, W., and Woldorff, M. (2004). Timing in the baby brain. *Cogn. Brain Res.* 21, 227–233. doi: 10.1016/j.cogbrainres.2004.04.007
- Brannon, E. M., Suanda, S., and Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Dev. Sci.* 10, 770–777. doi: 10.1111/j.1467-7687.2007.00635.x
- Breska, A., and Deouell, L. Y. (2014). Automatic bias of temporal expectations following temporally regular input independently of high-level temporal expectation. *J. Cogn. Neurosci.* 26, 1555–1571. doi: 10.1162/jocn_a_00564
- Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res.* 51, 1484–1525. doi: 10.1016/j.visres.2011.04.012
- Casasanto, D., Fotakopoulou, O., and Boroditsky, L. (2010). Space and time in the child's mind: evidence for a cross-dimensional asymmetry. *Cogn. Sci.* 34, 387–405. doi: 10.1111/j.1551-6709.2010.01094.x
- Colombo, J., and Richman, W. A. (2002). Infant timekeeping: attention and temporal estimation in 4-month-olds. *Psychol. Sci.* 13, 475–479. doi: 10.1111/1467-9280.00484
- Corbetta, M., Kincade, J. M., and Schulman, G. L. (2002). Neural systems for visual orienting and their relationship to spatial working memory. *J. Cogn. Neurosci.* 14, 508–523. doi: 10.1162/089892902317362029
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Correa, A., Lupiáñez, J., Milliken, B., and Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Percept. Psychophys.* 66, 264–278. doi: 10.3758/BF03194878
- Correa, A., and Nobre, A. C. (2008). Neural modulation by regularity and passage of time. *J. Neurophysiol.* 100, 1649–1655. doi: 10.1152/jn.90656.2008
- Correa, A., Triviño, M., Pérez-Dueñas, C., Acosta, A., and Lupiáñez, J. (2010). Temporal preparation, response inhibition and impulsivity. *Brain Cogn.* 73, 222–228. doi: 10.1016/j.bandc.2010.05.006
- Coull, J. T. (2009). Neural substrates of mounting temporal expectation. *PLoS Biol.* 7:e1000166. doi: 10.1371/journal.pbio.1000166
- Coull, J. T., Frith, C. D., Buchel, C., and Nobre, A. C. (2000). Orienting attention in time: behavioural and neuroanatomical distinction between exogenous and endogenous shifts. *Neuropsychologia* 38, 808–819. doi: 10.1016/S0028-3932(99)00132-3
- Coull, J. T., and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* 18, 7426–7435.
- Coull, J. T., and Nobre, A. C. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr. Opin. Neurobiol.* 18, 137–144. doi: 10.1016/j.conb.2008.07.011
- Cutanda, D., Correa, A., and Sanabria, D. (2015). Auditory temporal preparation induced by rhythmic cues during concurrent auditory working memory tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 790–797. doi: 10.1037/a0039167
- de la Rosa, M. D., Sanabria, D., Capizzi, M., and Correa, A. (2012). Temporal preparation driven by rhythms is resistant to working memory interference. *Front. Psychol.* 3:308. doi: 10.3389/fpsyg.2012.00308
- Delgado, M. D. L., and Droit-Volet, S. (2007). Testing the representation of time in reference memory in the bisection and the generalization task: the utility of a developmental approach. *Q. J. Exp. Psychol.* 60, 820–836. doi: 10.1080/17470210600790471
- Doherty, J. R., Rao, A. K., Mesulam, M. M., and Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* 25, 8259–8266. doi: 10.1523/JNEUROSCI.1821-05.2005
- Drake, C., and Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: evidence for a multiple-look model. *Percept. Psychophys.* 54, 277–286. doi: 10.3758/BF03205262
- Drake, C., Jones, M. R., and Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition* 77, 251–288. doi: 10.1016/S0010-0277(00)00106-2
- Droit-Volet, S. (2013). Time perception in children: a neurodevelopmental approach. *Neuropsychologia* 51, 220–234. doi: 10.1016/j.neuropsychologia.2012.09.023
- Droit-Volet, S. (2016). Development of time. *Curr. Opin. Behav. Sci.* 8, 102–109. doi: 10.1016/j.cobeha.2016.02.003
- Droit-Volet, S., and Coull, J. T. (2015). The developmental emergence of the mental time-line: spatial and numerical distortion of time judgement. *PLoS ONE* 10:e0130465. doi: 10.1371/journal.pone.0130465

- Droit-Volet, S., and Coull, J. T. (2016). Distinct developmental trajectories for explicit and implicit timing. *J. Exp. Child Psychol.* 150, 141–154. doi: 10.1016/j.jecp.2016.05.010
- Droit-Volet, S., and Wearden, J. H. (2001). Temporal bisection in children. *J. Exp. Child Psychol.* 80, 142–159. doi: 10.1006/jecp.2001.2631
- Droit-Volet, S., and Zélandi, P. S. (2013a). Development of time sensitivity and information processing speed. *PLoS ONE* 8:e71424. doi: 10.1371/journal.pone.0071424
- Droit-Volet, S., and Zélandi, P. S. (2013b). Development of time sensitivity: duration ratios in time bisection. *Q. J. Exp. Psychol.* 66, 671–686. doi: 10.1080/17470218.2012.712148
- Durston, S., Davidson, M. C., Mulder, M. J., Spicer, J. A., Galvan, A., Tottenham, N., et al. (2007). Neural and behavioral correlates of expectancy violations in attention-deficit hyperactivity disorder. *J. Child Psychol. Psychiatry* 48, 881–889. doi: 10.1111/j.1469-7610.2007.01754.x
- Eimer, M. (1997). Uninformative symbolic cues may bias visual-spatial attention: behavioral and electrophysiological evidence. *Biol. Psychol.* 46, 67–71. doi: 10.1016/S0301-0511(97)05254-X
- Ellis, R. J., and Jones, M. R. (2010). Rhythmic context modulates foreperiod effects. *Atten. Percept. Psychophys.* 72, 2274–2288. doi: 10.3758/APP.72.8.2274
- Enns, J. T., and Brodeur, D. A. (1989). A developmental study of covert orienting to peripheral visual cues. *J. Exp. Child Psychol.* 48, 171–189. doi: 10.1016/0022-0965(89)90001-5
- Fortin, C., and Rousseau, R. (1998). Interference from short term memory processing on encoding and reproducing brief durations. *Psychol. Res.* 61, 269–279. doi: 10.1007/s004260050031
- Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: evidence from interval discrimination, reproduction, and categorization. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 880–890. doi: 10.1037/a0027188
- Grondin, S., Ouellet, B., and Roussel, M.-E. (2001). About optimal timing and stability of Weber fraction for duration discrimination. *Acoust. Sci. Technol.* 22, 370–372. doi: 10.1250/ast.22.370
- Haith, M. M., Hazan, C., and Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Dev.* 59, 467–479. doi: 10.2307/1130325
- Harmon, C., Posner, M. I., Rothbart, M. K., and Thomas-Thrapp, L. (1994). Development of orienting to locations and objects in human infants. *Can. J. Exp. Psychol.* 48, 301–318. doi: 10.1037/1196-1961.48.2.301
- Iarocci, G., Enns, J. T., Randolph, B., and Burack, J. A. (2009). The modulation of visual orienting reflexes across the lifespan. *Dev. Sci.* 12, 715–724. doi: 10.1111/j.1467-7687.2009.00810.x
- Ivry, R. B., and Hazeltine, E. R. (1995). Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 3–18.
- Johnson, K. A., Burrows, E., and Coull, J. T. (2015). Children can implicitly, but not voluntarily, direct attention in time. *PLoS ONE* 10:e0123625. doi: 10.1371/journal.pone.0123625
- Jones, M. R., Boltz, M., and Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Percept. Psychophys.* 32, 211–218. doi: 10.3758/BF03206225
- Jones, M. R., Moynihan, H., Mackenzie, N., and Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol. Sci.* 13, 313–319. doi: 10.1111/1467-9280.00458
- Karlin, L. (1959). Reaction time as a function of foreperiod duration and variability. *J. Exp. Psychol.* 58, 185–191. doi: 10.1037/h0049152
- Keele, S. W., Nicoletti, R., Ivry, R. L., and Pokorny, R. A. (1989). Mechanisms of perceptual timing: beat-based or interval-based judgements? *Psychol. Res.* 50, 251–256. doi: 10.1007/BF00309261
- Klein, J. M., and Jones, M. R. (1996). Effects of attentional set and rhythmic complexity on attending. *Percept. Psychophys.* 58, 34–46. doi: 10.3758/BF03205473
- Lange, K., Rösler, F., and Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: an event-related potential study. *Psychophysiology* 40, 806–817. doi: 10.1111/1469-8986.00081
- Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychol. Bull.* 106, 119–159.
- Lin, Z., Lu, Z.-L., and He, S. (2016). Decomposing experience-driven attention: opposite attentional effects of previously predictive cues. *Atten. Percept. Psychophys.* doi: 10.3758/s13414-016-1101-z [Epub ahead of print].
- Los, S. A., and Heslenfeld, D. J. (2005). Intentional and unintentional contributions to nonspecific preparation: electrophysiological evidence. *J. Exp. Psychol. Gen.* 134, 52–72. doi: 10.1037/0096-3445.134.1.52
- Los, S. A., and van den Heuvel, C. E. (2001). Intentional and unintentional contributions to nonspecific preparation during reaction time foreperiods. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 370–386.
- Luce, R. D. (1986). *Response Times. Their Role in Inferring Elementary Mental Organization*. New York, NY: Oxford University Press.
- Martin, T., Egly, R., Houck, J. M., Pearson Bish, J., Barrera, B. D., Lee, D. C., et al. (2005). Chronometric evidence for entrained attention. *Percept. Psychophys.* 67, 168–184. doi: 10.3758/BF03195020
- McCormack, T., Wearden, J. H., Smith, M. C., and Brown, G. D. A. (2005). Episodic temporal generalization: a developmental study. *Q. J. Exp. Psychol. A* 58, 693–704. doi: 10.1080/02724980443000250
- Mento, G., and Tarantino, V. (2015). Developmental trajectories of internally and externally driven temporal prediction. *PLoS ONE* 10:e0135098. doi: 10.1371/journal.pone.0135098
- Mento, G., and Vallesi, A. (2016). Spatiotemporally dissociable neural signatures for generating and updating expectation over time in children: A High Density-ERP study. *Dev. Cogn. Neurosci.* 19, 98–106. doi: 10.1016/j.dcn.2016.02.008
- Merchant, H., Zarco, W., and Prado, L. (2008). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *J. Neurophysiol.* 99, 939–949.
- Michon, J. A. (1985). “The complete time experiencer,” in *Time, Mind and Behavior*, eds J. A. Michon and J. L. J. Jackson (Berlin: Springer).
- Miller, J. E., Carlson, L. A., and McAuley, J. D. (2013). When what you hear influences what you see: listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychol. Sci.* 24, 11–18. doi: 10.1177/0956797612446707
- Morillon, B., Schroeder, C. E., and Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nat. Commun.* 5, 1–9. doi: 10.1038/ncomms6255
- Niemi, P., and Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychol. Bull.* 89, 133–162. doi: 10.1037/0033-2909.89.1.133
- Nobre, A. C., Correa, A., and Coull, J. T. (2007). The hazards of time. *Curr. Opin. Neurobiol.* 17, 465–470. doi: 10.1016/j.conb.2007.07.006
- Olk, B. (2014). Effects of spatial, temporal and spatiotemporal cueing are alike when attention is directed voluntarily. *Exp. Brain Res.* 232, 3623–3633. doi: 10.1007/s00221-014-4033-7
- Pearson, D. A., and Lane, D. M. (1990). Visual attention movements: a developmental study. *Child Dev.* 61, 1779–1795. doi: 10.2307/1130837
- Philips-Silver, J., and Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science* 308, 1430. doi: 10.1126/science.1110922
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231
- Posner, M. I. (2016). Orienting of attention: then and now. *Q. J. Exp. Psychol. (Hove)* 69, 1864–1875. doi: 10.1080/17470218.2014.937446
- Richards, J. E. (2000). Localizing the development of covert attention in infants with scalp event-related potentials. *Dev. Psychol.* 36, 91–108. doi: 10.1037/0012-1649.36.1.91
- Ristic, J., Friesen, C. K., and Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychon. Bull. Rev.* 9, 507–513. doi: 10.3758/BF03196306
- Ristic, J., and Kingstone, A. (2009). Rethinking attentional development: reflexive and volitional orienting in children and adults. *Dev. Sci.* 12, 289–296. doi: 10.1111/j.1467-7687.2008.00756.x
- Rohenkohl, G., Coull, J. T., and Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE* 6:e14620. doi: 10.1371/journal.pone.0014620
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Pappert, Lercart, L., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012
- Sanabria, D., Capizzi, M., and Correa, A. (2011). Rhythms that speed you up. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 236–244. doi: 10.1037/a0019956
- Sattler, J. M., and Dumont, R. (2004). *Assessment of Children: WISC-IV and WPPSI-III Supplement*. San Diego, CA: Jerome M. Sattler Inc.

- Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Percept. Psychophys.* 45, 291–296.
- Tipples, J. (2002). Eye gaze is not unique: automatic orienting in response to uninformative arrows. *Psychon. Bull. Rev.* 9, 314–318. doi: 10.3758/BF03196287
- Trainor, L. J. (2012). Predictive information processing is a fundamental learning mechanism present in early development: evidence from infants. *Int. J. Psychophysiol.* 83, 256–258. doi: 10.1016/j.ijpsycho.2011.12.008
- Triviño, M., Arnedo, M., Lupiáñez, J., Chirivella, J., and Correa, A. (2011). Rhythms can overcome temporal orienting deficit after right frontal damage. *Neuropsychologia* 49, 3917–3930. doi: 10.1016/j.neuropsychologia.2011.10.009
- Vallesi, A., Arbula, S., and Bernardis, P. (2014). Functional dissociations in temporal preparation: evidence from dual-task performance. *Cognition* 130, 141–151. doi: 10.1016/j.cognition.2013.10.006
- Vallesi, A., Lozano, V. N., and Correa, A. (2013). Dissociating temporal preparation processes as a function of the inter-trial interval duration. *Cognition* 127, 22–30. doi: 10.1016/j.cognition.2012.11.011
- Vallesi, A., and Shallice, T. (2007). Developmental dissociations of preparation over time: deconstructing the variable foreperiod phenomena. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1377–1388.
- vanMarle, K., and Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Dev. Sci.* 9, F41–F49. doi: 10.1111/j.1467-7687.2006.00508.x
- Wainwright, A., and Bryson, S. E. (2005). The development of endogenous orienting: control over the scope of attention and lateral asymmetries. *Dev. Neuropsychol.* 27, 237–255. doi: 10.1207/s15326942dn2702_3
- Wechsler, D. (2004). *Wechsler Intelligence Scale for Children – IV*. London: Harcourt Brace & Company.
- Werner, L. A., Parrish, H. K., and Holmer, N. M. (2009). Effects of temporal uncertainty and temporal expectancy on infants' auditory sensitivity. *J. Acoust. Soc. Am.* 125, 1040–1049. doi: 10.1121/1.3050254
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., and Honing, H. (2009). Newborn infants detect the beat in music. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2468–2471. doi: 10.1073/pnas.0809035106
- Woodrow, H. (1914). The measurement of attention. *Psychol. Monogr.* 17, 1–158. doi: 10.1037/h0093087
- Zakay, D., and Block, R. A. (1996). “The role of attention in time estimation processes,” in *Time, Internal Clocks and Movement*, ed. M. A. Pastor (New York, NY: Elsevier Sciences), 143–164.
- Zélanti, P. S., and Droit-Volet, S. (2011). Cognitive abilities explaining age-related changes in time perception of short and long durations. *J. Exp. Child Psychol.* 109, 143–157. doi: 10.1016/j.jecp.2011.01.003
- Zélanti, P. S., and Droit-Volet, S. (2012). Auditory and visual differences in time perception? An investigation from a developmental perspective with neuropsychological tests. *J. Exp. Child Psychol.* 112, 296–311. doi: 10.1016/j.jecp.2012.01.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Johnson, Bryan, Polonowita, Decroupet and Coull. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Neurocognitive Performance of Visuospatial Attention in Children with Obesity

Chia-Liang Tsai^{1*}, Fu-Chen Chen², Chien-Yu Pan³ and Yu-Ting Tseng^{1,4}

¹ Lab of Cognitive Neurophysiology, Institute of Physical Education, Health and Leisure Studies, National Cheng Kung University, Tainan, Taiwan, ² Department of Recreational Sport and Health Promotion, National Pingtung University of Science and Technology, Tainan, Taiwan, ³ Department of Physical Education, National Kaohsiung Normal University, Kaohsiung, Taiwan, ⁴ School of Kinesiology, University of Minnesota, Minneapolis, MN, USA

The present study investigates the behavioral performance and event-related potentials (ERPs) in children with obesity and healthy weight children when performing a visuospatial attention task. Twenty-six children with obesity (obese group) and 26 healthy weight children (control group) were recruited. Their behavioral performance during a variant of the Posner paradigm was measured, and brain ERPs were recorded concurrently. The behavioral data revealed that the obese group responded more slowly, especially in the invalid condition, and exhibited a deficit in attentional inhibition capacity as compared to the control group. In terms of cognitive electrophysiological performance, although the obese group did not show significant differences on P3 latency elicited by the target stimuli when compared to the control group, they exhibited smaller P3 amplitudes when performing the visuospatial attention task. These results broaden previous findings, and indicate that childhood obesity is associated with a reduced ability to modulate the executive function network which supports visuospatial attention.

Keywords: obesity, attention, visuospatial ability, event-related potential, cognition

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Cihad Dundar,
Ondokuz Mayıs University, Turkey
Hong-Yan Bi,
Chinese Academy of Sciences, China

*Correspondence:

Chia-Liang Tsai
andytsai@mail.ncku.edu.tw

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 18 April 2016

Accepted: 24 June 2016

Published: 06 July 2016

Citation:

Tsai C-L, Chen F-C, Pan C-Y and
Tseng Y-T (2016) The Neurocognitive
Performance of Visuospatial Attention
in Children with Obesity.
Front. Psychol. 7:1033.
doi: 10.3389/fpsyg.2016.01033

INTRODUCTION

The prevalence of childhood overweightness and obesity has risen dramatically worldwide, and reached epidemic proportions over the past few decades (Wang and Lobstein, 2006; Liang et al., 2014). As a chronic condition, childhood obesity could be a contributor to the global burden of chronic disease, since children who are overweight or obese face a greater risk of developing various health problems, such as hypertension, cardiovascular disease, and type 2 diabetes mellitus (Reilly et al., 2003; Hannon et al., 2005). Children who are overweight are also prone to remain overweight as adults (Whitaker et al., 1997). Childhood obesity has thus become a major topic in discussion with regard to public health worldwide, due to the rise in premature obesity-related morbidities and related healthcare costs in recent decades. Moreover, some recent studies suggest that childhood obesity could not only be associated with adverse health sequelae, but also with cognitive problems (Davis and Cooper, 2011; Liang et al., 2014).

Executive functions, also called cognitive control, are a family of top-down mental processes, and this includes cognitive flexibility (i.e., response mapping, attention, or the ability to switch perspectives spatially), inhibition (e.g., response inhibitory and interference control), and working memory (Diamond, 2013). These aspects of cognition denote an ability to sustain or flexibly

redirect attention, as well as inhibit inappropriate behavioral responses (Robbins, 1998). Executive functions are thus important in individuals with obesity, as they can help them to orchestrate and maintain healthy goal-directed behaviors related to food intake and physical activity (Riggs et al., 2010). Moreover, numerous experimental studies demonstrate that problems with executive functions exist in children with obesity (Cserjesi et al., 2007; Li et al., 2008; Fergenbaum et al., 2009; Lokken et al., 2009; Kamijo et al., 2012a,b, 2014; Skoranski et al., 2013; Wirt et al., 2014).

Childhood obesity has negative associations with various aspects of neurocognitive functioning, such as attention and visuospatial performance (Liang et al., 2014). Individuals with obesity show diminished functional connectivity of the middle frontal gyrus and the lateral occipital cortex with the entire brain network, the regions involved in several brain circuits signaling perceptual processes, attention, executive, and motor functions (García-García et al., 2015). Compared to healthy weight children, those with obesity show reduced attentional focus (Cserjesi et al., 2007), and even during adulthood individuals with obesity still exhibit problems with regard to attention and cognitive flexibility (Cserjesi et al., 2009). In addition, more obese children/adolescents show lapses of attention (Pauli-Pott et al., 2010), and attention shifting/focus performance is associated with children's body weight (Wirt et al., 2015). In terms of visuospatial ability, previous studies demonstrated that children with higher body mass index (BMI) (i.e., over BMI 24) exhibited greater problems in visuospatial organization (Li et al., 2008) and mental rotation tasks (Jansen et al., 2011) compared to healthy weight children. Similarly, Martin et al. (2016) also found that boys with obesity showed significantly lower performance in visuospatial skills compared with those with a healthy weight, and obesity at age 3 years continued to predict decreased visuospatial skill at age 5 years (Martin et al., 2016).

Inhibitory control includes self-control, which is related to resisting temptation and not acting impulsively, as well as interference control, which involves cognitive inhibition (i.e., abilities related to suppressing prepotent mental representations) and selective attention/attentional inhibition (i.e., inhibitory control of attention which is related to the level of perception) (Diamond, 2013). There have been a few neuroimaging studies which find that individuals with obesity have significantly less gray matter volume in the prefrontal cortex, a brain region important in response inhibition (Raji et al., 2010; Horstmann et al., 2011; Maayan et al., 2011). Many studies have thus examined the relationship between childhood obesity and inhibitory control using the Go/No Go test, and found that pediatric obesity is linked to poorer inhibitory control abilities, and the degree of inhibitory control can help to predict bodyweight in children (Pauli-Pott et al., 2010; Kamijo et al., 2012a,b; Wirt et al., 2014). These previous studies demonstrated that children with obesity showed deficits in cognitive inhibition. However, although reduced attentional ability (Cserjesi et al., 2007; Davis and Cooper, 2011; Maayan et al., 2011) and poorer attention shifting performance (Pauli-Pott et al., 2010; Wirt et al., 2015) are associated with childhood obesity, no research has yet been conducted on the inhibitory control of attention, which

involves one of the interference control abilities at the level of perception (Diamond, 2013).

In the present study, a computerized serial reaction time (SRT) task (Robertson, 2007), the Posner paradigm, was used to assess the attention-related neurocognitive performances in the children with obesity. The task involves presenting a short sequence in each trial comprising a location cue followed by a target. In some trials the cue is valid, and in some trials it is not valid. The Posner paradigm requires a covert orienting of a visuospatial attention task, and is regarded as a valid and reliable measure for evaluating an individual's capacity for different attentional control modes (e.g., alerting, orienting, and shifting) (Posner et al., 2007). Given that the Posner paradigm task allows for the manipulation of executive function demands not only on attention and visuospatial processing abilities, but also on inhibitory control of attention [attentional inhibition, as measured by the inhibitory response effect (subtracting the mean RT between valid and invalid trials)] based on the response requirements (Posner and DiGirolamo, 1998; Theeuwes, 2010), we used it to elucidate the association between childhood obesity and neurocognitive performance in the current study. The visuospatial attention task coupled with concomitant electrophysiological recording (e.g., event-related potential, ERP) provides images of cortical activity with a high temporal resolution (milliseconds), and the evaluation of the time evolution of the global brain response to cognitive processing (Babiloni et al., 2009). Among the various different visual response ERP components elicited by the Posner paradigm, the endogenous ERP P3 regarding upstream processes has been demonstrated to reflect attention-related brain activity, that is, the resource allocation necessary for attention (e.g., stimulus evaluation) (Perchet and Garcia-Larrea, 2005; Polich, 2007; Neuhaus et al., 2009). In addition, the P3 component is also involved in conflict-related brain activity, this is, the attentional resources that are allocated to efficiently inhibit a response (Jonkman et al., 2003). A previous study reported that children with obesity had smaller P3 amplitudes than normal-weight ones when given a cognitive task involving attentional processes (i.e., the oddball paradigm) (Babiloni et al., 2009). The ERP component was thus used in the current study to better understand the cognitive neurophysiological mechanism of visuospatial attention in children with obesity.

In summary, compared to normal-weight controls, children with obesity seem to have problems in attention/attentional shifting (Cserjesi et al., 2007; Pauli-Pott et al., 2010; Davis and Cooper, 2011; Maayan et al., 2011; Wirt et al., 2015) and visuospatial abilities (Li et al., 2008; Jansen et al., 2011; Martin et al., 2016). To the best of our knowledge, no ERP study has used a paradigm that directly measures the shifting of visuospatial attention in children with obesity. Additionally, although numerous studies demonstrate that children with obesity show impairments which inhibit their prepotent mental and behavioral responses, there is a lack of research with regard to the attentional inhibition of such a group. The aim of this study was therefore to investigate the relationship between childhood obesity and executive functions, involving both visuospatial attention shifting and inhibitory control of attention, using the

Posner paradigm task. Based on the findings reviewed above, we hypothesized that children with obesity relative to healthy-weight cohorts would exhibit worse behavioral (i.e., slower reaction times and lower accuracy rates) performances and ERP abnormalities (i.e., prolonged P3 latencies and smaller P3 amplitudes) when performing the visuospatial attention task.

MATERIALS AND METHODS

Participants

Fifty-two children aged 9–10 years were recruited from mainstream classrooms in urban areas of Taiwan and categorized into obese ($n = 26$; eight girls; age: 114.58 ± 3.69 months) and healthy weight (control) ($n = 26$; eight girls; age: 113.73 ± 3.85 months) groups. Children in the obese group ($BMI = 27.39 \pm 1.62 \text{ kg/m}^2$) had a body mass index (BMI, calculated as weight/height^2) greater than the 95th percentile for their height and weight, and children in the control group ($BMI = 18.45 \pm 2.23 \text{ kg/m}^2$) had BMIs between the 5th and 85th percentiles according to Taiwanese norms, as seen in the related BMI-for-age growth chart (Ministry of Education, 2015). To exclude confounding effects of cardiorespiratory fitness on cognition (Hillman et al., 2008), all children had their cardiorespiratory fitness assessed using the PACER test, a multistage progressive 20-m shuttle run test, in the Brockport Physical Fitness Test Kit (Human Kinetics, Champaign, IL, USA). The number of shuttle runs was not significantly different [$t(50) = 0.35$, $p = 0.729$] between the obese (40.38 ± 25.72) and control (42.54 ± 18.32) groups. All the participants had normal or corrected-to-normal vision and were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). None of the children who took part in the current work had any clear signs of neurological disorders or behavioral problems, or special needs in education that would exclude them from this study. All the children were assessed using the Wechsler Intelligence Scale for Children-Revised (WISC-R), and fell within normal intelligence quotient (IQ) scores [obese: 104.56 ± 5.74 , control: 106.69 ± 10.53 , $t(50) = -0.41$, $p = 0.682$]. In addition, since children suffering from attention deficit hyperactivity disorder (ADHD) have shown abnormalities in behavioral response and ERPs during completion of the Posner paradigm (Perchet et al., 2001), parents and teachers were asked to complete a brief behavior rating scale (Dupaul et al., 1998) based on the DSM-IV criteria for ADHD, considering the children's behavior patterns in the last 6 months, and verify that they did not have DSM-IV ADHD (less than six inattention and six hyperactive/impulsive symptoms). Parental education level [obese vs. control groups: 14.96 ± 2.78 vs. 15.92 ± 3.24 years, $t(50) = 1.15$, $p = 0.256$] was not significantly different between the two groups. Prior to the beginning of the experiment, each child and his/her legal guardians provided written informed assent and consent, and this study was approved by the Institutional Ethics Committee.

The Posner Paradigm

The Posner paradigm was carried out with reference to a previous study, in which the cognitive task was used with subjects of a

similar age (Tsai et al., 2009). The cognitive task was presented on a computer screen, with a fixation cross ($0.5^\circ \times 0.5^\circ$) drawn in white on a black background at the beginning of the process. This served as a central fixation point, and was positioned midway between two empty white boxes (each $2 \text{ cm} \times 2 \text{ cm}$) on the same horizontal plane. The two empty boxes that were the potential locations for the target were arranged horizontally 1 cm from the fixation cross. The overall stimulus display remained on-screen until the end of the trial, except for the white fixation cross, as this was replaced by a yellow cue arrow during each trial. A trial started with a 3-s countdown followed 1000 ms later by the appearance of the two white stimulus boxes and the white fixation cross, and this was then followed 1000 ms later by the replacement of the fixation cross by the yellow cue arrow (1.5 cm in length), pointing to the right or left. After a further interval between cue onset and the appearance of the target for 350 ms, a green circle target stimulus with a diameter of 1.6 cm appeared in the center of the right or left white stimulus box. Upon detection of the green circle target the children were asked to press as quickly as possible the left “N” or right “M” button of the computer keyboard with the index or middle fingers of the dominant hand, and to avoid errors as much as possible. If the child did not respond to the target, the maximal inter-trial interval occurred 3 s after the target stimulus, and in such cases the system noted that there was a lack of response and a new trial was started. Each child completed 180 trials, and each experimental session was divided into two consecutive runs of 90 trials each, with a 3-min break after each block, when the child was allowed to rest but remained at the workstation. Each block of 90-trials consisted of three types of trials in a random order: (i) 54 valid trials (60%), where the target appeared in the stimulus box indicated by the cue, indicating a spatially ‘valid’ condition; (ii) 27 invalid trials (30%), where the target appeared in the opposite stimulus box to that indicated by the cue, indicating a spatially ‘invalid’ condition; and (iii) nine neutral trials (10%), where the target appeared without any cue (i.e., the non-cued condition). The probability with which the targets were presented in the left or right stimulus boxes was the same. For both valid and invalid trials, the direction to which the arrow pointed, left or right, was random and equally probable. After the button was pressed, the screen cleared and the next trial started 1500 ms later. Therefore, the inter-trial interval from cue to cue was variable, depending on the speed of the child's response.

Procedures

Before the cognitive task test the participants and their legal guardians completed all questionnaires, as mentioned above, to ascertain if they were qualified to take part in this study, and the participants' cardiorespiratory fitness was also assessed. On the second visit to the laboratory, the experimenter explained the procedure until the child was familiar with it. An electrocap and electro-oculographic (EOG) electrodes were attached to the child's head and face before the test. After all the equipment had been set up, the child was asked to sit in an adjustable chair in front of a computer screen (width = 43 cm), with this linked to an IBM compatible personal computer with a stimulation system (Neuroscan Ltd., EI Paso, USA). The overall stimulus display was

shown on a laptop computer screen located directly in front of the child, at face level and a distance of approximately 75 cm. All subjects simultaneously performed the Posner paradigm with concomitant electrophysiological recording. To familiarize children with the experimental procedure, a practice block of 10 trials was run before the beginning of the formal Posner paradigm test, during which they had to respond as quickly and accurately as possible, but without emphasizing one at the expense of the other (e.g., not to focus on speed to the detriment of accuracy). If they made more than 10% errors it was assumed that they could not understand the experimental procedure. In such cases the experimenter then explained the process again, and asked the child to continue practicing until they had less than 10% errors. The experimenter was seated next to the child to monitor his/her visual fixation. If the experimenter detected the child's eye movement away from the central stimulus during the response, they gave verbal encouragement to the child to look back at the screen. The formal test was administered once the child understood the whole experimental procedure. The cognitive experiment was administered in a sound-attenuated room with dimmed lights.

Psychophysiological Recording Methods

Electroencephalographic (EEG) activity as well as blinks and saccades were recorded from 18 sites (F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, and O2) using Ag/AgCl sintered electrodes embedded in an electrocap (Quik-Cap, Compumedics Neuroscan, Inc., El Paso, TX) according to the 10–20 system. Scalp locations were referred to linked left (A1) and right mastoids (A2), with AFz as the ground electrode. The adhesive electrodes that were placed on the supero-lateral right canthus and below and lateral to the left eye connected to the system reference for horizontal and vertical EOG (i.e., HEOG and VEOG) activity for eye movements. Electrode impedances were less than 10 k Ω . EEG data was acquired with an A/D rate of 500 Hz/channel and filter band-pass of 0.1–50 Hz, 60 Hz notch filter, and was written continuously to hard disk for off-line analysis using SCAN 4.3 analysis software (Compumedics Neuroscan, Inc., El Paso, TX, USA).

Data Processing

The subjects' behavioral performance was measured with the percentage of errors, as well as reaction times (RTs) to each target presentation. An error was recorded according to the following three standards: (1) anticipatory error: the participants responded sooner than 150 ms; (2) orientation error: a button-pressing error occurred (i.e., the response was not consistent with the location of the target); and (3) delay error: the participants responded more than 2000 ms after target onset. Some RTs were discarded when regarded as errors, and the rest of the RTs were grouped according to different conditions. The "strength of the inhibitory response effect" was calculated by subtracting the mean RT between valid and invalid trials (Tsai et al., 2009).

For the ERP components, the data used for the behavioral analysis was adopted to characterize the ERP elicited by the stimuli in order to truly correspond with the behavioral performance. Initially, each EEG epoch was visually inspected

and discarded from those including EEG artifacts (e.g., electromyogram exceeded 100 μ V peak-to-peak amplitude, VEOG, and HEOG) before averaging. The rest of the data was then averaged off-line, using a \pm 100 μ V automatic artifact rejection, and was constructed from both the valid and invalid conditions (excluding the non-cued condition, since this study is concerned with the strength of the inhibitory control effect, and the number of non-cued trials was small) over a 1550 ms epoch beginning 200 ms before cue stimulus onset. For target-elicited ERP components, the P3 mean amplitudes were calculated for 250–400 ms time intervals. These windows were determined from inspection of the group grand average waveforms, and were equivalent for ERP elicited by two conditions and subjects. Latencies were measured within the latency window for every child. The amplitude values for all ERP components, with reference to the 200 ms cue stimulus baseline, were determined within latency windows centered on the peak latency of the grand mean ERP.

Statistical Analysis

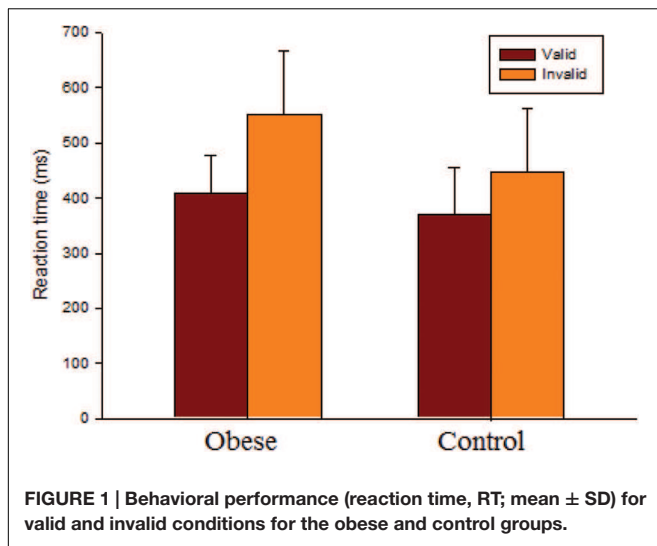
For the behavioral data, since the values of the accuracy rate were not normally distributed, statistical assessment of this was carried out with the Mann-Whitney non-parametric test. The results of the separate behavioral (e.g., RTs) and cognitive electrophysiological (e.g., P3 latency and amplitude) performances were analyzed statistically by a mixed design, factorial, and repeated-measures analysis of variance (RM ANOVA), with mean RTs of accepted trials serving as the dependent variable, with Group (obese vs. control groups) as the between-subjects factor, and Condition (valid vs. invalid, excluding the non-cued condition, as this study is concerned with the strength of the inhibitory response effect) as the within-subject factor. If any interactions between the group and condition factors were found, the strength of the inhibitory response effect between the two-group comparisons was further investigated. For the ERPs data, the P3 component (e.g., latency and amplitude) was also analyzed statistically by a mixed design, factorial, and RM ANOVA, with Group (obese vs. control groups) as the between-subjects factor, and Condition (valid vs. invalid conditions) and Electrode (Fz vs. Cz vs. Pz) as the within-subjects variables. Where a significant difference occurred, Bonferroni *post hoc* analyses were performed. Estimates of effect size, partial eta-square (η_p^2), were reported for significant main effects and interactions. The significance levels of the F ratios were adjusted with the Greenhouse–Geisser correction if the assumption of sphericity was violated. The significance level was set as $p < 0.05$.

RESULTS

Behavioral Performance

Accuracy Rate

The number of orientation errors did not show a significant difference between groups (obese: 4.19 ± 4.34 , control: 5.54 ± 3.89 ; Mann–Whitney non-parametric test, $p = 0.061$). The numbers of anticipatory (obese: 2.54 ± 2.79 , control: 3.62 ± 4.19 ; Mann–Whitney non-parametric test, $p = 0.344$)



and delay errors (obese: 0.19 ± 0.63 , control: 0.19 ± 0.57 ; Mann-Whitney non-parametric test, $p = 0.987$) also did not differ significantly between groups. In addition, no significant difference (obese: 0.04 ± 0.04 , control: 0.05 ± 0.04 ; Mann-Whitney non-parametric test, $p = 0.078$) was found between groups for the error rate [i.e., (orientation error + anticipatory error + delay error)/180].

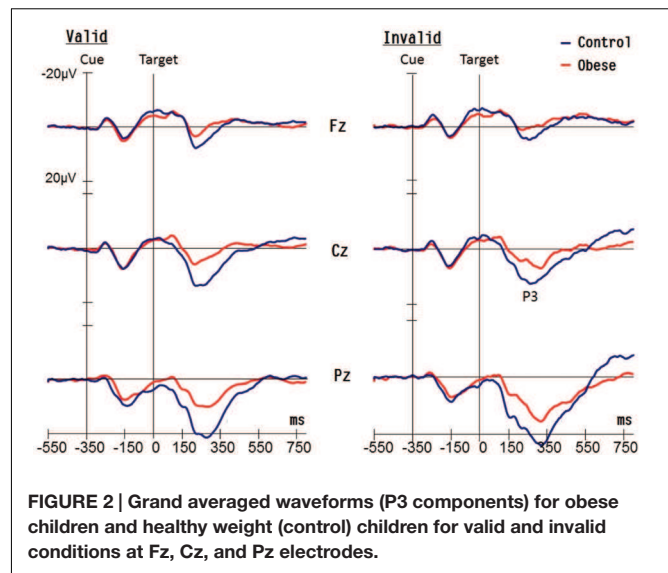
Reaction Time

As seen in **Figure 1**, the RM ANOVA on the RTs revealed significant main effects of Group [$F(1,50) = 7.93$, $p = 0.007$, $\eta_p^2 = 0.14$] and Condition [$F(1,50) = 120.12$, $p < 0.001$, $\eta_p^2 = 0.71$]. *Post hoc* analyses indicated that the obese group (479.58 ms) responded more slowly than the control group (408.35 ms) in both conditions, and that the valid condition (388.56 ms) showed a significant difference with regard to the invalid condition (499.37 ms) in both groups. These main effects were superseded by the Group \times Condition [$F(1,50) = 10.85$, $p = 0.002$, $\eta_p^2 = 0.18$] interaction. *Post hoc* analysis showed that the obese group only responded more slowly than the control group in the invalid condition (obese: 551.63 ms vs. control: 447.11 ms, $p = 0.002$), and not the valid one (obese: 407.52 ms vs. control: 369.59 ms, $p = 0.088$). The results indicated that the average value of the strength of the inhibitory response effect (invalid RT-valid RT) (obese: 144.11 ± 93.37 ms vs. control: 77.51 ± 43.74 ms) was significantly larger in the obese group as compared to the control group.

ERPs Performances

Target-P3 Latency

There was no effect of Group on the latency of the target-P3 components (see **Figure 2**). An effect of Condition [$F(1,50) = 74.06$, $p < 0.001$, $\eta_p^2 = 0.60$] was observed on the latency of the target-P3, with shorter latency in valid (250.64 ms) relative to invalid conditions (295.56 ms). There was an effect of Electrode [$F(2,100) = 47.53$, $p < 0.001$, $\eta_p^2 = 0.49$] on the latency of the target-evoked P3 with the following gradient:



Fz (299.39 ms) > Cz (274.49 ms) > Pz (245.41 ms). These main effects were superseded by the Condition \times Electrode [$F(2,100) = 6.68$, $p = 0.002$, $\eta_p^2 = 0.12$] interaction. *Post hoc* analysis showed that the gradient Fz > Cz > Pz was found in both valid and invalid conditions.

Target-P3 Amplitude

As illustrated in **Figure 2**, an effect of Group [$F(1,50) = 39.24$, $p < 0.001$, $\eta_p^2 = 0.44$] was observed on the amplitude of the target-P3 response, with smaller amplitudes in the obese group (8.32 μ V) in comparison with the control group (15.03 μ V). There was an effect of Electrode [$F(2,100) = 318.83$, $p < 0.001$, $\eta_p^2 = 0.86$] on the amplitude of the target-P3, with the following gradient: Pz (18.93 μ V) > Cz (10.97 μ V) > Fz (5.12 μ V). There was also a Group \times Electrode [$F(2,100) = 19.23$, $p < 0.001$, $\eta_p^2 = 0.28$] interaction explained by the following gradient of amplitude, and the discrepancy was greater in the control group when compared to the obese group for the three electrodes [Pz: 27.74 vs. 14.39 μ V, $p < 0.001$; Cz: 15.83 vs. 8.96 μ V, $p < 0.001$; Fz: 8.16 vs. 5.89 μ V, $p = 0.049$]. In addition, the interaction of Condition \times Electrode [$F(2,100) = 55.99$, $p < 0.001$, $\eta_p^2 = 0.53$] also achieved a significant difference. *Post hoc* analysis showed that the gradient, Pz > Cz > Fz, was found in both valid and invalid conditions.

DISCUSSION

The present study examined covert orienting of visuospatial attention, either valid or invalid with target presentation, in children with obesity and healthy weight, with the aim of better understanding the visuospatial attention abilities of the obese group. Although the anticipatory, delay, and orientation errors did not differ significantly between obese and control groups, consonant with our hypothesis, the obese group relative to control group showed significantly slower RTs in the invalid condition and a significantly weaker inhibitory control of

attention in the visuospatial attention task. In terms of cognitive electrophysiological performance, although the obese group, when compared to the control group, did not exhibit significant differences in the P3 latencies, they showed smaller P3 amplitudes elicited by the target stimuli. These findings show that children with obesity may not only have neuropsychological deficits in attentional inhibition, but could also have certain aberrances in visuospatial attention processing compared with controls, supporting the conclusions of previous fMRI studies that obesity is linked with dysregulated activation in a distributed network of areas involved in executive function/attention (Carnell et al., 2012; García-García et al., 2015).

Since the children with obesity showed comparable accuracy rates as the controls, the RT differences between the two groups could be attributed to the differences in processing time, and not due to any trade-off between speed and accuracy. The computerized visuospatial attention task in the current study is a SRT task which involves manual key presses and/or covert reorienting of visuospatial attention (Rosenthal et al., 2009) directed to two fixed locations in response to a target stimulus that is cued at one of the two spatial locations. The present finding suggests that children with obesity could be less efficient in dealing with the cue-target sequence supporting the previous finding that individuals with obesity showed longer RTs when performing a computerized cognitive task involving attentional responses (Babiloni et al., 2009). In addition, the obese group exhibited significantly slower RTs across all conditions relative to the control group when performing the visuospatial attention task, also supporting earlier studies investigating the relationship between childhood obesity and attentional problems (Cserjesi et al., 2007; Pauli-Pott et al., 2010; Davis and Cooper, 2011; Maayan et al., 2011; Wirt et al., 2015). For example, Cserjesi et al. (2007) found that children with obesity, compared to the control children, showed prolonged RTs on an attention endurance task, and the correlation analysis revealed a relationship between BMI/body weight and the results of the attention task in children. Nederkoorn et al. (2006) also found that more obese children/adolescents exhibited slower RTs when performing the stop-signal task needing attention requirement. Therefore, given the findings in the present and previous studies, it seems probable that the obese group showed poorer executive functions involving visuospatial attention relative to the control group.

Interference control, one of the inhibitory control abilities, includes cognitive inhibitory and attentional inhibition (Diamond, 2013). Individuals with obesity have significantly less gray matter volume in the brain areas (e.g., orbitofrontal cortex) involving response inhibition (Raji et al., 2010; Horstmann et al., 2011; Maayan et al., 2011). Previous studies have demonstrated that children with obesity showed deficits in cognitive inhibitory focus using the Go/No Go test (Pauli-Pott et al., 2010; Kamijo et al., 2012a,b; Wirt et al., 2014). The obese group in the current study also seemed to have a problem with regard to attentional inhibition, since the significantly larger RT performance in the strength of inhibitory response in this group demonstrated that these subjects' inhibitory control of attention was worse with regard to orienting attention toward the falsely indicated location than that seen with the healthy weight children. The findings

also indicate that children with obesity were much slower at modifying a movement after they had been primed to an invalid location (i.e., they were unable to complete the attentional shift as efficiently as the controls). In contrast, healthy weight children could be less reliant on the spatial information given by the cue when preparing their responses, and thus maintain attentional scanning of the whole field. Indeed, Li et al. (2008) found that children with higher BMI showed poorer performance in visuospatial organization, even controlling for parental socioeconomic status. Similar to the visuospatial attention task adopted in the present study, Pauli-Pott et al. (2010) used an incompatibility task (i.e., press left or right button depending on whether the arrow points right or left) to measure attention and the capability to resist interference and inhibit a preponderant response, and found that more obese children/adolescents showed lower inhibitory control performance due to lapses of attention. Likewise, Wirt et al. (2015) also adopted a cognitive flexibility task (i.e., press left or right button depending on the color of stimuli on the left and right sides) to assess the attention shifting/focus, and found that cognitive flexibility performance was associated with the children's body weight. The current findings also demonstrate that children with obesity, as compared to the healthy weight controls, suffered a reduction in the time efficiency of the central processing of cognitive functions associated with the disengagement of visuospatial attention. This indicates that weight status during childhood could be related to the visuospatial attention networks that have been implicated in executive functions (Bruce et al., 2010; Carnell et al., 2012). Accordingly, based on the findings outlined above, a consensus appears to have been reached in the literature concerning the negative association between childhood obesity and visuospatial attention information processing.

P3 is an ERP component typically associated with attentional stimulus evaluations, with P3 latency being related to the speed of cognitive stimulus processing and response selection, and P3 amplitude being proportional to the amount of attentional resources allocated to a task (Polich, 2007). There was no significant difference in the P3 latencies between the two groups in the current study, suggesting that the time needed for target stimulus evaluation and detection was comparable in both groups (Perchet and Garcia-Larrea, 2000). However, the obese group exhibited smaller P3 amplitudes, despite having similar intelligence and cardiorespiratory fitness levels as those in the control group, indicating that the children with obesity exhibited less efficient allocation of attentional resources or reduced attention focus compared to the healthy weight children when performing the visuospatial attention task. However, this finding is somewhat inconsistent with Kamijo et al. (2012b), which showed no significant between-group (obese vs. healthy weight) differences in the P3 amplitudes for both the Go and No Go tasks. However, in agreement with Carnell et al. (2012) it was noted in the current study that individuals with obesity showed less activation in areas associated with object processing and attention, potentially indicating a relative absence of objective evaluation of stimuli. Similarly, Babiloni et al. (2009) used an oddball paradigm to assess the attentional cortical response, and found that the amplitude of medial prefrontal

P3 sources (Brodmann area 9) was lower in the obese than normal-weight subjects in the food condition, and there was a negative correlation between the body fat percentage and P3 amplitude across conditions. These previous findings suggest that individuals with obesity show less activation in the prefrontal cortex implicated in the attention and cognitive control (Babiloni et al., 2009; Carnell et al., 2012). Nevertheless, in the present study, the effects of the between-group difference in the P3 amplitude were significant not only at the frontal cortex, but also at the central and parietal cortices when the obese and control groups performed a cognitive task involving attentional control (one of the interference control abilities). The results are partly in accordance with those in the earlier studies which found a significant discrepancy between the P3 amplitudes of obese and healthy weight children at the central (Cz) electrode when performing the auditory oddball task (i.e., children with obesity showed decreased P3 amplitude when compared to the healthy controls, Tascilar et al., 2011). In addition, when exposed to pictures of food in a eucaloric state or in response to anticipatory food, individuals with obesity showed obviously discrepant activation in the parietal lobe as compared to the lean individuals (Stice et al., 2008; Tregellas et al., 2011). The previous and present findings may thus, at least partly, reflect that individuals with obesity show impairments in the neural activity of the frontoparietal cortices with regard to activation, and diminished functional network connectivity in the brain areas involved in several brain circuits signaling perceptual processes, attention, executive and motor functions (Stice et al., 2008; Carnell et al., 2012; García-García et al., 2015).

It is worth noting with regard to attentional networks that obesity is related to a decrease in striatal dopaminergic receptors (Wang et al., 2001), and that dopaminergic transmission plays an important role in attentive neural processing (e.g., the strength of the P3 signals) (Neuhaus et al., 2009). Based on the present and previous findings, as presented in the paragraph above, the attenuated P3 amplitudes found for the obese group could be signs of weaker attentive neural processing due to fewer striatal dopaminergic receptors. In addition, the cognitive task adopted in the present study is a sequential task. Children with obesity performing the visuospatial attention task during the test had to learn a higher-order association in the SRT task which involves a distinct cortico-cerebellar/cortico-striatal network, with activation in the inferior parietal lobule identified with encoding an effector-independent description of successive locations (Grafton et al., 1998; Seidler et al., 2005). Poorer cognitive electrophysiological performance (e.g., P3 amplitude) in children with obesity in the present study seems to imply that such a group seems to have deficits in these neural networks. Also, the P3 amplitude and cognitive resource allocation may more or less reflect the physiological processes related to the RTs (Muller and Knight, 2002), and since the children with obesity in the present study exhibited smaller P3 amplitudes, this might result in longer RTs. Moreover, P3 has been suggested as an inhibition component in children (Jonkman et al., 2003), and children with higher BMI are associated with less activation in the brain regions involved in inhibitory control (Batterink et al., 2010). The smaller

P3 amplitudes found in the current study could also partly explain the poorer attentional inhibition in children with obesity.

Although some potentially confounding factors (e.g., parental education level and cardiorespiratory fitness) which could mediate the obesity-cognition association were rigorously controlled in the present study (Hillman et al., 2008; Li et al., 2008), there are still some potential limitations to its cross-sectional study design. Dual-process theories of attention (Corbetta and Shulman, 2002) propose that orienting of attention is controlled by dorsal and ventral networks (Butler et al., 2009), which might represent the endogenous, goal-directed attention orienting system and the exogenous, stimulus-driven attention orienting system, respectively (Perry and Zeki, 2000; Corbetta and Shulman, 2002). In the current study, the endogenous orienting task [i.e., a high probability (i.e., 60%) of valid precues] was adopted to give rise to the facilitatory effect (Muller and Rabbitt, 1989; Rafal and Henik, 1994). Therefore, children with obesity displayed a deficit in volitional/intentional orienting of visual attention, that is, in the dorsal attention network. As yet it remains unclear whether children with obesity also display a visuospatial attention deficit in shifts of automatic (exogenous) attention. In addition, individuals with metabolic syndrome (MS) have showed an approximately fourfold increased risk of lowered cognitive performance after adjusting for insulin resistance (IR) relative to those without MS (Fergenbaum et al., 2009). Children who are overweight/obese are prone to have dyslipidemia and MS (Casavalle et al., 2014). Children with obesity and IR show significantly different electrophysiological performance compared to those without IR when doing the auditory oddball task, suggesting that IR is an important risk factor leading to cognitive dysfunction in such children (Tascilar et al., 2011). Therefore, one avenue for future research is to examine the possibility of an interaction among childhood obesity, MS, IR, and neurocognitive functions. Further, obesity is determined by BMI and is associated with an excessive accumulation of peripheral fat. Although actual body fat content could not be reflected by BMI percentiles as measured from body weight (Wirt et al., 2014), in general, BMI percentile is a clinically meaningful measure for population research (Reilly et al., 2003), which enables comparability with other studies. Despite poorer performance in visuospatial attention in children with higher BMI in the present study, BMI may sometimes underestimate the relationship between cognition and weight status during childhood. Therefore, how the role of fat mass influences the relationship between obesity and executive functions in children remains an open question. Lastly, besides genetic predisposition, some life risk factors, such as preferences for a sedentary lifestyle (e.g., low levels of physical activity) and poor dietary behaviors (e.g., overeating or over-consuming fat or sugar) are demonstrated to contribute to childhood obesity (Reilly et al., 2005; Braet et al., 2007; Iannotti and Wang, 2013) which could impair cognitive functions (Bechara, 2005; Verdejo-Garcia et al., 2006). In the current study, the obese group exhibited similar cardiorespiratory fitness to the healthy weight group, which implies comparable levels of physical activity between the two. As such, the issue as to whether deficits in inattention

and attentional inhibition are the underlying mechanism of overeating in children needs to be investigated.

CONCLUSION

Children with obesity showed poorer behavioral (e.g., slower RTs and larger values of the strength of the inhibitory response effect) performances and aberrant neural activity (e.g., smaller P3 amplitudes) associated with cognitive information processing when doing the visuospatial attention task in the present study. Previous studies mostly examined childhood obesity with regard to the health consequences, such as diabetes, cardiorespiratory diseases, hypertension, or future risk of adult obesity. In contrast, only few studies have explored the possible relation between childhood obesity and cognitive functions, specifically with regard to executive function tasks. Executive function is responsible for adjusting behavior in relation to a situation which requires individuals to resist temptation, as well as for transmissions between the internal world and environmental challenges (Norman and Shallice, 1980/2000). The findings of the

current study extend those of previous works, and imply that deviant cognitive processing should also be taken into account as an obesity-related health issue, and thus how to treat both conditions (i.e., obesity and executive function impairment), rather than obesity in isolation, is an important issue in clinical practice.

AUTHOR CONTRIBUTIONS

Dr. C-LT designed the study, wrote the protocol, and the first draft of the manuscript. Dr. F-CC analyzed the data. Dr. C-YP helped explain results. Mrs. Y-TT helped collect data.

ACKNOWLEDGMENTS

We are grateful for the participation of students and staff in this research, which was supported by a grant from the National Science Council (NSC 98-2410-H-006-106-MY2 and NSC 102-2628-H-006-003-MY3) in Taiwan.

REFERENCES

- Babiloni, C., Del Percio, C., Valenzano, A., Marzano, N., De Rosas, M., Petito, A., et al. (2009). Frontal attentional responses to food size are abnormal in obese subjects: an electroencephalographic study. *Clin. Neurophysiol.* 120, 1441–1448. doi: 10.1016/j.clinph.2009.06.012
- Batterink, L., Yokum, S., and Stice, E. (2010). Body mass correlates inversely with inhibitory control in response to food among adolescent girls: an fMRI study. *Neuroimage* 52, 1696–1703. doi: 10.1016/j.neuroimage.2010.05.059
- Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat. Neurosci.* 8, 1458–1463. doi: 10.1038/nn1584
- Braet, C., Claus, L., Verbeken, S., and van Vlierberghe, L. (2007). Impulsivity in overweight children. *Eur. Child Adolesc. Psychiatry* 16, 473–483. doi: 10.1007/s00787-007-0623-2
- Bruce, A. S., Holsen, L. M., Chambers, R. J., Martin, L. E., Brooks, W. M., Zarcone, J. R., et al. (2010). Obese children show hyperactivation to food pictures in brain networks linked to motivation, reward and cognitive control. *Int. J. Obes.* 34, 1494–1500. doi: 10.1038/ijo.2010.84
- Butler, B. C., Lawrence, M., Eskes, G., and Klein, A. (2009). Visual search patterns in neglect: comparison of peripersonal and extrapersonal space. *Neuropsychologia* 47, 869–878. doi: 10.1016/j.neuropsychologia.2008.12.020
- Carnell, S., Gibson, C., Benson, L., Ochner, C. N., and Geliebter, A. (2012). Neuroimaging and obesity: current knowledge and future directions. *Obes. Rev.* 13, 43–56. doi: 10.1111/j.1467-789X.2011.00927.x
- Casavalle, P. L., Lifshitz, F., Romano, L. S., Pandolfo, M., Caamaño, A., Boyer, P. M., et al. (2014). Prevalence of dyslipidemia and metabolic syndrome risk factor in overweight and obese children. *Pediatr. Endocrinol. Rev.* 12, 213–223.
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Cserjési, R., Luminet, O., Poncelet, A., and Lénard, L. (2009). Altered executive function in obesity. Exploration of the role of affective states on cognitive abilities. *Appetite* 52, 535–539. doi: 10.1016/j.appet.2009.01.003
- Cserjési, R., Molnar, D., Luminet, O., and Lenard, L. (2007). Is there any relationship between obesity and mental flexibility in children? *Appetite* 49, 675–678. doi: 10.1016/j.appet.2007.04.001
- Davis, C. L., and Cooper, S. (2011). Fitness, fatness, cognition, behavior, and academic achievement among overweight children: do cross-sectional associations correspond to exercise trial outcomes? *Prev. Med.* 52, S65–S69. doi: 10.1016/j.ypmed.2011.01.020
- Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Dupaul, G. J., Power, T. J., Anastopoulos, A. D., and Reid, R. (1998). *ADHD Rating Scale II: Checklists, Norms, and Clinical Interpretation*. New York, NY: Guilford.
- Fergenbaum, J. H., Bruce, S., Lou, W., Hanley, A. J., Greenwood, C., and Young, T. K. (2009). Obesity and lowered cognitive performance in a Canadian first nation population. *Obesity* 17, 1957–1963. doi: 10.1038/oby.2009.161
- García-García, I., Jurado, M. Á., Garolera, M., Marqués-Iturria, I., Horstmann, A., Segura, B., et al. (2015). Functional network centrality in obesity: a resting-state and task fMRI study. *Psychiatry Res.* 233, 331–338. doi: 10.1016/j.psychres.2015.05.017
- Grafton, S. T., Hazeltine, E., and Ivry, R. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *J. Neurosci.* 18, 9420–9428.
- Hannon, T. S., Rao, G., and Arslanian, S. A. (2005). Childhood obesity and type 2 diabetes mellitus. *Pediatrics* 116, 473–480. doi: 10.1542/peds.2004-2536
- Hillman, C. H., Erickson, K. I., and Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nat. Rev. Neurosci.* 9, 58–65. doi: 10.1038/nrn2298
- Horstmann, A., Busse, F. P., Mathar, D., Müller, K., Lepsien, J., Schlögl, H., et al. (2011). Obesity-related differences between women and men in brain structure and goal-directed behavior. *Front. Hum. Neurosci.* 5:58. doi: 10.3389/fnhum.2011.00058
- Iannotti, R. J., and Wang, J. (2013). Trends in physical activity, sedentary behavior, diet, and BMI among US adolescents, 2001–2009. *Pediatrics* 132, 606–614. doi: 10.1542/peds.2013-1488
- Jansen, P., Schmelter, A., Kasten, L., and Heil, M. (2011). Impaired mental rotation performance in overweight children. *Appetite* 56, 766–769. doi: 10.1016/j.appet.2011.02.021
- Jonkman, L. M., Lansbergen, M., and Stauder, J. E. A. (2003). Developmental differences in behavioral and event-related brain responses associated with response preparation and inhibition in a go/nogo task. *Psychophysiology* 40, 752–761. doi: 10.1111/1469-8986.00075
- Kamijo, K., Khan, N. A., Pontifex, M. B., Scudder, M. R., Drollette, E. S., Raine, L. B., et al. (2012a). The relation of adiposity to cognitive control and scholastic achievement in preadolescent children. *Obesity* 20, 2406–2411. doi: 10.1038/oby.2012.112
- Kamijo, K., Pontifex, M. B., Khan, N. A., Raine, L. B., Scudder, M. R., Drollette, E. S., et al. (2012b). The association of childhood obesity to neuroelectric indices of inhibition. *Psychophysiol.* 49, 1361–1371. doi: 10.1111/j.1469-8986.2012.01459.x

- Kamijo, K., Pontifex, M. B., Khan, N. A., Raine, L. B., Scudder, M. R., Drollette, E. S., et al. (2014). The negative association of childhood obesity to cognitive control of action monitoring. *Cereb. Cortex* 24, 654–662. doi: 10.1093/cercor/bhs349
- Li, Y., Dai, Q., Jackson, J. C., and Zhang, J. (2008). Overweight is associated with decreased cognitive functioning among school-age children and adolescents. *Obesity* 16, 1809–1815. doi: 10.1038/oby.2008.296
- Liang, J., Matheson, B. E., Kaye, W. H., and Boutelle, K. N. (2014). Neurocognitive correlates of obesity and obesity-related behaviors in children and adolescents. *Int. J. Obes.* 38, 494–506. doi: 10.1038/ijo.2013.142
- Lokken, K. L., Boeka, A. G., Austin, H. M., Gunstad, J., and Harmon, C. M. (2009). Evidence of executive dysfunction in extremely obese adolescents: a pilot study. *Surg. Obes. Relat. Dis.* 5, 547–552. doi: 10.1016/j.soard.2009.05.008
- Maayan, L., Hoogendoorn, C., Sweat, V., and Convit, A. (2011). Disinhibited eating in obese adolescents is associated with orbitofrontal volume reductions and executive dysfunction. *Obesity* 19, 1382–1387. doi: 10.1038/oby.2011.15
- Martin, A., Booth, J. N., Young, D., Revie, M., Boyter, A. C., Johnston, B., et al. (2016). Associations between obesity and cognition in the pre-school years. *Obesity* 24, 207–214. doi: 10.1002/oby.21329
- Ministry of Education (2015). *Taiwanese Fitness Norm*. Available at: <http://www.fitness.org.tw/model08.php>
- Muller, H. J., and Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 315–330. doi: 10.1037/0096-1523.15.2.315
- Muller, N. G., and Knight, R. T. (2002). Age-related changes in fronto-parietal networks during spatial memory: an ERP study. *Brain Res. Cogn. Brain Res.* 13, 221–234. doi: 10.1016/S0926-6410(01)00119-7
- Nederkorn, C., Jansen, E., Mulken, S., and Jansen, A. (2006). Impulsivity predicts treatment outcome in obese children. *Behav. Res. Ther.* 45, 1071–1075. doi: 10.1016/j.brat.2006.05.009
- Neuhaus, A. H., Goldberg, T. E., Hassoun, Y., Bates, J. A., Nassauer, K. W., Sevy, S., et al. (2009). Acute dopamine depletion with branched chain amino acids decreases auditory top-down event-related potentials in healthy subjects. *Schizophr. Res.* 111, 167–173. doi: 10.1016/j.schres.2009.03.023
- Norman, D. A., and Shallice, T. (1980/2000). “Attention to action: willed and automatic control of behaviour,” in *Cognitive Neuroscience: A Reader*, ed. M. Gazzaniga (Oxford: Blackwell), 376–391.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Pauli-Pott, U., Albayrak, Ö., Hebebrand, J., and Pott, W. (2010). Association between inhibitory control capacity and body weight in overweight and obese children and adolescents: dependence on age and inhibitory control component. *Child Neuropsychol.* 16, 592–603. doi: 10.1080/09297049.2010.485980
- Perchet, C., and Garcia-Larrea, L. (2000). Visuospatial attention and motor reaction in children: an electrophysiological study of the ‘Posner’ paradigm. *Psychophysiology* 37, 231–241. doi: 10.1111/1469-8986.3720231
- Perchet, C., and Garcia-Larrea, L. (2005). Learning to react: anticipatory mechanisms in children and adults during a visuospatial attention task. *Clin. Neurophysiol.* 116, 1906–1917. doi: 10.1016/j.clinph.2005.03.022
- Perchet, C., Revol, O., Fournier, P., Manguiere, F., and Garcia-Larrea, L. (2001). Attention shifts and anticipatory mechanisms in hyperactive children: an ERP study using the Posner paradigm. *Biol. Psychiatry* 50, 44–57. doi: 10.1016/S0006-3223(00)01119-7
- Perry, R. J., and Zeki, S. (2000). The neurology of saccades and covert shifts in spatial attention: an event-related fMRI study. *Brain* 123, 2273–2288. doi: 10.1093/brain/123.11.2273
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Posner, M. I., and DiGirolamo, G. J. (1998). “Executive attention: conflict, target detection, and cognitive control,” in *The Attentive Brain*, ed. R. Parasuraman (Cambridge, MA: MIT Press), 401–423.
- Posner, M. I., Rothbart, M. K., and Sheese, B. E. (2007). Attention genes. *Dev. Sci.* 10, 24–29. doi: 10.1111/j.1467-7687.2007.00559.x
- Rafal, R. D., and Henik, A. (1994). “The neurology of inhibition: integrating controlled and automatic processes,” in *Inhibitory Processes in Attention, Memory, and Language*, eds D. Dagenback and T. H. Carr (San Diego, CA: Academic Press), 1–51.
- Raji, C. A., Ho, A. J., Parikshak, N. N., Becker, J. T., Lopez, O. L., Kuller, L. H., et al. (2010). Brain structure and obesity. *Hum. Brain Mapp.* 31, 353–364. doi: 10.1002/hbm.20870
- Reilly, J. J., Armstrong, J., Dorosty, A. R., Emmett, P. M., Ness, A., Rogers, I., et al. (2005). Early life risk factors for obesity in childhood: cohort study. *BMJ* 330: 1357. doi: 10.1136/bmj.38470.670903.E0
- Reilly, J. J., Methven, E., McDowell, Z. C., Hacking, B., Alexander, D., Stewart, L., et al. (2003). Health consequences of obesity. *Arch. Dis. Child.* 88, 748–752. doi: 10.1136/ad.88.9.748
- Riggs, N., Chou, C. P., Spruijt-Metz, D., and Pentz, M. A. (2010). Executive cognitive function as a correlate and predictor of child food intake and physical activity. *Child Neuropsychol.* 16, 279–292. doi: 10.1080/09297041003601488
- Robbins, T. W. (1998). “Dissociating executive functions of the prefrontal cortex,” in *The Prefrontal Cortex. Executive and Cognitive Function*, eds A. C. Roberts, T. W. Robbins, and L. Weiskrantz (Oxford: Oxford University Press), 117–131.
- Robertson, E. M. (2007). The serial reaction time task: implicit motor skill learning? *J. Neurosci.* 27, 10073–10075. doi: 10.1523/JNEUROSCI.2747-07.2007
- Rosenthal, C. R., Roche-Kelly, E. E., Husain, M., and Kennard, C. (2009). Response dependent contributions of human primary motor cortex and angular gyrus to manual and perceptual sequence learning. *J. Neurosci.* 29, 15115–15125. doi: 10.1523/JNEUROSCI.2603-09.2009
- Seidler, R. D., Purushotham, A., Kim, S. G., Ugurbil, K., Willingham, D., and Ashe, J. (2005). Neural correlates of encoding and expression in implicit sequence learning. *Exp. Brain Res.* 165, 114–124. doi: 10.1007/s00221-005-2284-z
- Skoranski, A. M., Most, S. B., Lutz-Stehl, M., Hoffman, J. E., Hassink, S. G., and Simons, R. F. (2013). Response monitoring and cognitive control in childhood obesity. *Biol. Psychol.* 92, 199–204. doi: 10.1016/j.biopsycho.2012.09.001
- Stice, E., Spoor, S., Bohon, C., Veldhuizen, M. G., and Small, D. M. (2008). Relation of reward from food intake and anticipated food intake to obesity: a functional magnetic resonance imaging study. *J. Abnorm. Psychol.* 117, 924–935. doi: 10.1037/a0013600
- Tascilar, M. E., Turkkahraman, D., Oz, O., Yucel, M., Taskesen, M., Eker, I., et al. (2011). P300 auditory event-related potentials in children with obesity: is childhood obesity related to impairment in cognitive functions? *Pediatr. Diabetes Metab. Res. Rev.* 12, 589–595. doi: 10.1111/j.1399-5448.2010.00748.x
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychol.* 135, 77–99. doi: 10.1016/j.actpsy.2010.02.006
- Tregellas, J. R., Wylie, K. P., Rojas, D. C., Tanabe, J., Martin, J., Kronberg, E., et al. (2011). Altered default network activity in obesity. *Obesity (Silver Spring)* 19, 2316–2321. doi: 10.1038/oby.2011.119
- Tsai, C. L., Pan, C. Y., Cherng, R. J., Hsu, Y. W., and Chiu, H. H. (2009). Mechanisms of deficit of visuospatial attention shift in children with developmental coordination disorder: a neurophysiological measures of the endogenous Posner paradigm. *Brain Cogn.* 71, 246–258. doi: 10.1016/j.bandc.2009.08.006
- Verdejo-Garcia, A., Bechara, A., Recknor, E. C., and Perez-Garcia, M. (2006). Executive dysfunction in substance dependent individuals during drug use and abstinence: an examination of the behavioural, cognitive and emotional correlates of addiction. *J. Int. Neuropsychol. Soc.* 12, 405–415. doi: 10.1017/S1355617706006046
- Wang, G. J., Volkow, N. D., Logan, J., Pappas, N. R., Wong, C. T., Zhu, W., et al. (2001). Brain dopamine and obesity. *Lancet* 357, 354–357. doi: 10.1016/S0140-6736(00)03643-6
- Wang, Y., and Lobstein, T. (2006). Worldwide trends in childhood overweight and obesity. *Int. J. Pediatr. Obes.* 1, 11–25. doi: 10.1080/17477160600586747

- Whitaker, R. C., Wright, J. A., Pepe, M. S., Seidel, K. D., and Dietz, W. H. (1997). Predicting obesity in young adulthood from childhood and parental obesity. *N. Engl. J. Med.* 337, 869–873. doi: 10.1056/NEJM199709253371301
- Wirt, T., Hundsdoerfer, V., Schreiber, A., Kesztyüs, D., Steinacker, J. M., and Komm mit in das gesunde Boot-Grundschule-Research Group. (2014). Associations between inhibitory control and body weight in German primary school children. *Eat. Behav.* 15, 9–12. doi: 10.1016/j.eatbeh.2013.10.015
- Wirt, T., Schreiber, A., Kesztyüs, D., and Steinacker, J. M. (2015). Early life cognitive abilities and body weight: cross-sectional study of the association of inhibitory control, cognitive flexibility, and sustained attention with BMI percentiles in primary school children. *J. Obes.* 2015:534651. doi: 10.1155/2015/534651

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Tsai, Chen, Pan and Tseng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Nature of Verbal Short-Term Impairment in Dyslexia: The Importance of Serial Order

Steve Majerus^{1*} and Nelson Cowan²

¹ Psychology and Neuroscience of Cognition Unit, Department of Psychology, University of Liège, Liège, Belgium,

² Department of Psychological Sciences, University of Missouri, Columbia, MO, USA

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Thomas Lachmann,
Kaiserslautern University
of Technology, Germany
G. Brian Thompson,
Victoria University of Wellington,
New Zealand

*Correspondence:

Steve Majerus
smajerus@ulg.ac.be

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 13 June 2016

Accepted: 20 September 2016

Published: 03 October 2016

Citation:

Majerus S and Cowan N (2016)
The Nature of Verbal Short-Term
Impairment in Dyslexia:
The Importance of Serial Order.
Front. Psychol. 7:1522.
doi: 10.3389/fpsyg.2016.01522

Verbal short-term memory (STM) impairment is one of the most consistent associated deficits observed in developmental reading disorders such as dyslexia. Few studies have addressed the nature of this STM impairment, especially as regards the ability to temporarily store serial order information. This question is important as studies in typically developing children have shown that serial order STM abilities are predictors of oral and written language development. Associated serial order STM deficits in dyslexia may therefore further increase the learning difficulties in these populations. In this mini review, we show that specific serial order STM impairment is frequently reported in both dyslexic children and adults with a history of dyslexia. Serial order STM impairment appears to occur for the retention of both verbal and visuo-spatial sequence information. Serial order STM impairment is, however, not a characteristic of every individual dyslexic subject and is not specific to dyslexia. Future studies need to determine whether serial order STM impairment is a risk factor which, in association with phonological processing deficits, can lead to dyslexia or whether serial order STM impairment reflects associated deficits causally unrelated to dyslexia.

Keywords: short-term memory, verbal, serial order, dyslexia, phonological

INTRODUCTION

Dyslexia is characterized by important and persisting difficulties in acquiring accurate and efficient reading abilities despite normal-range intellectual efficiency (Snowling, 2000). Although the precise underlying factors are still a matter of debate, input phonological processing difficulties are most frequently identified as being impaired in dyslexia, in addition to the difficulties in reading acquisition (Ramus et al., 2003, 2013; Serniclaes et al., 2004; Szenkovits and Ramus, 2005). These phonological processing difficulties are considered to prevent efficient mapping of phonemic and graphemic representations, leading to protracted reading development and slowed reading speed in adulthood, and this for languages with either consistent or inconsistent phonology-to-orthography mappings (Ziegler and Goswami, 2005). A further associated factor is verbal short-term memory (STM) impairment. Verbal STM capacity, as measured by digit span or non-word repetition, is typically reduced in children with dyslexia, and this reduction is still present in adults with a history of dyslexia (Brady et al., 1983; Avons and Hanna, 1995; Snowling et al., 1996). This deficit may represent a contributing factor to dyslexia, by reducing the amount of phonological and graphemic information that can be co-activated during the reading process at a given time, and this especially during the recoding reading process, when grapheme-phoneme mappings are not yet automatized

(Gathercole and Baddeley, 1993; Martinez Perez et al., 2012b). The purpose of this article is to review the implications for dyslexia of one aspect of STM, the memory for the serial order or serial positions of items, which has received an increasing research interest over the past 5 years. Word reading, especially at the beginning of reading acquisition, is a sequential process involving the extraction of an ordered mental representation of the letters from the printed word, and the construction of a temporary phonological sequences that matches the letter sequence.

THE PROBLEM OF OVERLAP BETWEEN VERBAL STM AND PHONOLOGICAL PROCESSING IN DYSLEXIA

The nature of the link between verbal STM impairment and dyslexia is complicated by the fact that mechanisms that allow information to be maintained for a given duration are themselves dependent upon access to long-term memory (LTM) representations. These LTM representations correspond to representations stored in the language system. Several studies showed that the likelihood of an item being recalled in a verbal STM task is determined partly by the nature of underlying linguistic representations: Words lead to higher recall performance than non-words (Hulme et al., 1991). Even for non-words, linguistic knowledge impacts STM in that non-words with a phonotactic structure more frequent in the language lead to higher recall performance than non-words of low phonotactic frequency (Gathercole et al., 1999; Majerus et al., 2004, 2012). Linguistic knowledge thus appears to be an important determinant of verbal STM. If linguistic representations are poorly developed, verbal STM performance will be directly impacted. In the case of dyslexia, this means that verbal STM impairment could be a consequence of the phonological processing impairment which characterizes dyslexia.

ITEM-ORDER DISTINCTION IN STM

Not all aspects of verbal STM, however, depend upon access to the language system. Mainly retention of item information appears to be influenced by long-term knowledge: Linguistic variables such as word frequency and semantic similarity determine the number of items recalled in a STM task, but not the number of items recalled in correct serial order (Nairne and Kelley, 2004). At a theoretical level, item information (i.e., the words of a STM list) is considered to be coded by temporarily activating the underlying language representations; serial order information is often considered to rely on distinct processing systems, based on temporal, spatial, or magnitude codes (Henson, 1998; Page and Norris, 1998; Burgess and Hitch, 1999, 2006; Brown et al., 2000; van Dijck and Fias, 2011; van Dijck et al., 2013). Developmental studies have shown that serial order STM abilities predict lexical and reading development independently of item STM abilities, and are the most robust predictor of lexical and reading abilities (Majerus et al., 2006a; Leclercq and Majerus, 2010; Martinez

Perez et al., 2012b). Two recent studies showed that serial order STM abilities assessed in children at third year of kindergarten predict their reading decoding abilities 1 and 2 years later (Martinez Perez et al., 2012b; Binamé and Poncet, 2016). It follows that the distinction between item and serial order STM abilities may be particularly useful for understanding the nature of verbal STM impairment in dyslexia. If verbal STM in dyslexia is simply a consequence of underlying phonological processing difficulties, then performance for item STM should be particularly impaired. If, on the other hand, there are additional, specific STM deficits in dyslexia, serial order STM should also be impaired.

ITEM VERSUS SERIAL ORDER STM IN DYSLEXIA

Next, we review the recent studies that have distinguished item and serial order STM abilities in dyslexia. These studies mainly involve adults with a history of dyslexia as there are currently very few studies that specifically investigated serial order STM in children with dyslexia.

Item versus Serial Order STM in Children with Dyslexia

A first study making an explicit distinction between item and serial order STM was conducted by Martinez Perez et al. (2012a) in children with dyslexia. The authors used tasks designed to maximize temporary retention of either item or serial order information (**Table 1**). The item STM task was a single non-word repetition task probing the retention of phonological item information. Serial order STM was assessed using a serial order reconstruction task for auditory sequences of familiar words. Martinez Perez et al. (2012a) observed both item and serial order STM deficits, with serial order STM deficits and item STM deficits appearing to be independent; the serial order STM deficit was observed relative to both chronological age and reading age matched control groups, whereas the item STM deficit was observed only relative to the chronological age matched control group. These results were partly replicated by Staels and Van den Broeck (2014) in multilingual children with a diagnosis of dyslexia, by showing also both item and serial order STM difficulties, but, contrary to the study by Martinez Perez et al. (2012a), the deficit in the serial order STM task appeared to be dependent upon the deficit in the item STM task.

Item versus Serial Order STM in Adults with Dyslexia

Studies in adults with a history of dyslexia have used similar study designs to distinguish item and serial order STM abilities. Hachmann et al. (2014) contrasted item and serial order probe recognition tasks, using both verbal and visual STM tasks. They observed specifically impaired serial order STM, and this interestingly for both verbal and visual modalities (**Table 1**). They did not observe verbal item STM impairment, while this impairment could have been expected given that verbal item

TABLE 1 | Synthetic presentation of studies included in the review, with description of study characteristics, tasks and main outcome.

Study	Age/N	Language	Source of diagnosis	Reading deficit	Item processing task	Serial order processing task	Outcome
Children							
Martinez Perez et al. (2012a)	10.30/22	French	Formal	-1.86 ^b	<i>Delayed item repetition task</i> : Recall of single, monosyllabic non-words after 6-s filled retention delay.	<i>Serial order reconstruction task</i> : Auditory sequences of familiar animal names (2-7 stimuli); reconstruction of the order of the animals using cards	Item STM: DYS = RA < CA Order STM: DYS < RA < CA
Staelen and Van den Broeck (2014)	10.75/36	Multilingual, Dutch-speaking schools	Formal	-1.81 ^a	<i>Delayed item repetition task</i> : Recall of single, monosyllabic non-words after 6-s filled retention delay	<i>Serial order reconstruction task</i> : Auditory sequences of familiar animal names (2-7 stimuli); reconstruction of the order of the animals using cards	Item STM: DYS < CA Order STM: DYS < CA
Staelen and Van den Broeck (2014)	Study 2 10.76/25	Dutch	Formal	-1.49 ^a	Not applicable	<i>Serial order learning task</i> : Hebb-repetition learning for syllable (verbal condition) and dot sequences (visuo-spatial condition); comparison of reproduction performance for repeating and non-repeating sequences	Hebb Verbal: DYS = CA Hebb Visuo-spatial: DYS = CA
Adults							
Bogaerts et al. (2015a)	21.22/25	Dutch	Formal	-1.64 ^a	Not applicable	<i>N-back task</i> : Continuous sequence of pictures of familiar objects; detection of items identical to those presented two items earlier in the sequence, with serial position confusions manipulated via foils occurring in 3-back positions.	DYS < CA
Bogaerts et al. (2015b)	Study 1 20.60/25 Study 2 21.35/18	Dutch	Formal Formal	-1.64 ^a -3.06 ^a	Not applicable	<i>Serial order learning task</i> : Hebb-repetition learning for syllable sequences; comparison of reproduction performance for repeating and non-repeating sequences.	Hebb Verbal: DYS < CA
Hachmann et al. (2014)	20.80/21	Dutch	Formal	-1.65 ^a	<i>Word probe recognition task</i> : 13 pictures depicting familiar objects presented sequentially, followed by an auditory word recognition probe. <i>Visual probe recognition task</i> : 13 nonsense drawings presented sequentially, followed by a nonsense drawing recognition probe.	<i>Digit serial order recognition</i> : Sequentially presented target and recognition probe lists of nine digits each <i>Visual serial order recognition</i> : Target and recognition probe sequences of four nonsense drawings each. Target and probe lists differed by the exchange of two adjacent serial positions	Verbal item STM: DYS = CA Verbal order STM: DYS < CA Visual item STM: DYS = CA Visual order STM: DYS < CA
Laasonen et al. (2012)	43.30/30	Finnish	Self-report	-2.19 ^b	Not applicable	<i>Multi-modal sensory sequence recognition</i> : Pairs of sequences mixing brief auditory and tactile stimuli (3-15 stimuli); non-matching pairs differed by the order of the stimuli within each sequence	Sequence STM : DYS < CA

(Continued)

TABLE 1 | Continued

Study	Age/N	Language	Source of diagnosis	Reading deficit	Item processing task	Serial order processing task	Outcome
Martinez Perez et al. (2013)	24.30/31	French	Formal	-1.80 ^a	<p><i>Study 1: Delayed item repetition task</i>: Recall of single, monosyllabic non-words after 8-s filled retention delay</p> <p><i>Study 2: Item and order reconstruction task</i>: Auditory sequences of familiar animal names (5–9 stimuli); reconstruction of the sequences using cards to be selected from a pool of target and non-target stimuli → wrongly selected cards = item error score</p> <p><i>Study 3: Item recognition task</i>: Auditory sequences of four non-words, followed by a single non-word recognition probe</p>	<p><i>Study 1: Serial order reconstruction task</i>: Auditory sequences of familiar animal names (6–9 stimuli); reconstruction of the order of the animals using cards.</p> <p><i>Study 2: Item and order reconstruction task</i> → wrongly placed cards = order error score</p> <p><i>Study 3: Serial order recognition task</i>: Auditory sequences of four non-words, followed by a two-non-word recognition probe, with non-words in the same order as in the target list or not</p>	<p>Study 1: Item STM: DYS < CA Order STM: DYS < CA</p> <p>Study 2: Item STM: DYS < CA Order STM: DYS < CA</p> <p>Study 3: Item STM: DYS = CA Order STM: DYS < CA</p>
Martinez Perez et al. (2015)	23.60/15	French	Formal	-2.10 ^b	<p><i>Item recognition task; verbal condition</i>: Array of four horizontally aligned words, followed by a non-word recognition probe; <i>visual condition</i>: Array of four horizontally aligned unfamiliar faces, followed by a face recognition probe</p>	<p><i>Serial order recognition task; verbal condition</i>: Array of four horizontally aligned words, followed by a recognition probe with two non-words aligned vertically; <i>visual condition</i>: Array of four horizontally aligned unfamiliar faces, followed by a recognition probe with two faces aligned vertically; participants decided whether the order of the stimuli in the probe display matched their order in the memory array</p>	<p>Verbal item STM: DYS < CA Verbal order STM: DYS < CA Visual item STM: DYS = CA Visual order STM: DYS < CA</p>
Romani et al. (2015)	27.7/44	English	Formal (23) Self-report (13) No diagnosis (8)	-3.42 ^a	Not applicable	<p><i>Serial order reconstruction task</i>: Simultaneous or sequential presentation of four Hindi characters or five Japanese characters, followed by a serial order reconstruction task using tiles</p>	<p>Sequential: DYS < CA Simultaneous: DYS = CA</p>
Staelen and Van den Broeck (2014)	Study 1 20.69/26	Dutch	Formal	-2.15 ^a	Not applicable	<p><i>Serial order learning task</i>: Hebb-repetition learning for syllable (verbal condition) and dot sequences (visuo-spatial condition); comparison of reproduction performance for repeating and non-repeating sequences</p>	<p>Hebb Verbal: DYS = CA Hebb Visuo-spatial: DYS = CA</p>

(Continued)

TABLE 1 | Continued

Study	Age/N	Language	Source of diagnosis	Reading deficit	Item processing task	Serial order processing task	Outcome
Szmaliec et al. (2011)	21, 19/16	Dutch	Formal	−2.65 ^a	Not applicable	Serial order learning task: Hebb-repetition learning for syllable (verbal condition) and dot sequences (visuo-spatial condition); comparison of reproduction performance for repeating and non-repeating sequences	Hebb Verbal: DYS < CA Hebb Visuo-spatial: DYS < CA
Wang et al. (2016)	Adults/28	English	Self-report	−1.15 ^a	Immediate serial recall task: recall of visually presented 6-word lists → Item score: all items produced on inclusion ('recall all words') and exclusion ('do not recall word in position X') conditions	Immediate serial recall task: → Order score: all items produced on inclusion divided by Item score	Item STM: DYS = CON Order STM: DYS = CON

Age/N: age in years and sample size of the DYS group; DYS: group with a formal or self-reported diagnosis of dyslexia; CA: control group matched on chronological age and non-verbal intelligence; RA: control group matched for reading age; CON: control group with no specific matching variable.
^aZ-score for performance (speed, combined speed-accuracy, or accuracy) of the dyslexic group on a non-word reading task; ^bZ-score for performance (speed or combined speed-accuracy) of the dyslexic group on a word reading task.

STM is considered to depend on access to underlying linguistic representations. However, the linguistic levels that are impaired in dyslexia are at the level of phonological rather than lexico-semantic representations (Ramus et al., 2013); the item verbal memory lists in the study by Hachmann et al. (2014) were comprised of sequences of pictures depicting familiar objects followed by an auditory probe word, inducing a strong lexico-semantic component for item maintenance and recognition. Martinez Perez et al. (2013) administered a range of item and serial order STM tasks to adults with a history of dyslexia, and they observed both impaired item STM and serial order STM performances. They did so on the basis of two kinds of results: Those from tasks specifically designed to dissociate item and serial order STM processes, as in Martinez Perez et al. (2012a), and from item and serial order errors observed in an item and serial order reconstruction task (Table 1). Item and serial order STM deficits appeared to be statistically independent and their correlation ($r = -0.07$) was non-significant after controlling for verbal and non-verbal intellectual efficiency. Note that the dyslexic participants were impaired in non-word and word item recall tasks, but not in a non-word item probe recognition task, suggesting that phonological item representations were sufficient for accurate recognition but not full reproduction. The finding of specific serial order impairment is also supported by a recent study by Bogaerts et al. (2015a), observing impaired performance in a sample of adult dyslexic participants in an N-back task requiring efficient maintenance and updating of serial order information. In another recent study, however, Wang et al. (2016) found no evidence for either verbal item or serial order STM impairment in undergraduate university students with a self-reported diagnosis of dyslexia. They used a process dissociation procedure to derive item and serial order STM estimates from performance in a word list immediate serial recall task; the task demands of this procedure may have contributed to these results, given that serial order recall is estimated in a rather indirect manner, by asking participants to recall all items except the item occurring in a specified serial position (Table 1).

Brain Correlates of Item and Serial Order STM in Dyslexia

A recent neuroimaging study in adults with a history of dyslexia sheds further light on the status of item and serial order STM in dyslexia. Martinez Perez et al. (2015) investigated the neural networks associated with item and serial order short-term probe recognition tasks for both verbal (words) and visuo-spatial (faces) stimuli (Table 1). Although at the behavioral level, item and serial order STM deficits of similar severity were observed for the verbal STM modality, they were associated with distinct neural networks. The dyslexic participants activated to a higher extent the left intraparietal cortex, the bilateral cingulate cortex and the right dorsolateral prefrontal cortex in the verbal item STM condition; this network has been associated with attentional control processes during STM tasks (Silton et al., 2010), and may have reflected the greater difficulty of this task in the dyslexic group relative to the control group. In the serial order STM condition the

dyslexic group activated to a lower extent a network centered around the right intraparietal sulcus, which had been identified in other studies to be specifically associated with serial order STM processes (Majerus et al., 2006b; see also Beneventi et al., 2009, for similar results comparing single letter versus letter sequence probe recognition tasks). Interestingly, the same network was also hypoactivated in the visual serial order STM condition, and was, at the behavioral level, associated with both serial order STM and reading impairment. The only condition where the dyslexic participants did not differ from controls, at both behavioral and neural levels, was the visual item STM condition.

Dyslexia and Serial Order Processing in Other Domains

Other studies have investigated serial order processing capacities in dyslexia using tasks that slightly differ from STM tasks. Szmalec et al. (2011) as well as Bogaerts et al. (2015b) showed that adults with dyslexia present difficulties in learning verbal and visuo-spatial sequences in Hebb repetition experiments involving the reproduction of repeating and novel sequences of supra-span length (Table 1). Staels and Van den Broeck (2015) could not replicate sequence learning difficulties in verbal or visuo-spatial Hebb learning experiments, but like Bogaerts et al. (2015b), they had observed impaired performance already for non-repeating, filler sequences, further highlighting the difficulties for temporary maintenance of serial order information in dyslexia. Other studies also point to serial order processing difficulties in dyslexia. Romani et al. (2015) required adult dyslexic participants to reproduce the order of presentation of visual characters (Table 1). They observed difficulties in reconstructing the order of the characters, especially when they were presented in a sequential rather than a simultaneous manner. Finally, Laasonen et al. (2012) observed deficits in adults with a history of dyslexia for STM for audio-tactile sequences (Table 1) with a strong serial order processing component, and these deficits correlated with performance on verbal STM tasks.

DISCUSSION

This mini-review of STM deficits in dyslexia reveals a number of important findings. First, all studies reviewed here, except for one, show verbal STM impairment in dyslexia, and these deficits persist until adulthood. Second, the verbal STM deficits cannot be explained only on the basis of underlying phonological processing impairment, given that both item STM, considered to depend most strongly on phonological processing, and serial order STM aspects appear to be impaired; importantly, serial order STM also appears to be impaired in the visuo-spatial STM domain, further ruling out the possibility that serial order STM would only be the consequence of verbal impairment.

At the same time, the level to which item and serial order STM deficits are independent in the verbal domain has been questioned, some studies showing independent verbal item and serial order STM impairment, while others do not. The use of bilingual populations with an emigration background

in some studies (Staels and Van den Broeck, 2014), making proper identification of dyslexia difficult, could have been one contributing factor to these inconsistent findings. Dyslexic group specificities may also be related to the absence of both item and serial order STM deficits reported by Wang et al. (2016). In that study, dyslexic participants were undergraduate students at a university with competitive access; dyslexic applicants with important STM difficulties and ensuing learning difficulties (Gathercole and Alloway, 2006) may have difficulties in reaching the academic grades necessary for entry at university, as acknowledged by the authors.

These observations, however, also stress the likely heterogeneity of dyslexia populations as regards the presence and severity of verbal STM impairment, and particularly serial order STM impairment. We further know that serial order STM impairment is not specific to dyslexia as it has also been observed in other developmental learning disorders such as dyscalculia (Attout and Majerus, 2015). We suggest here that poorly developed serial order STM abilities increase the risk of learning difficulties in different cognitive domains and situations (Leclercq and Majerus, 2010; Jaroslawska et al., 2016). If occurring at the same time as phonological processing difficulties, the serial order STM difficulties will put the child in a particularly difficult situation for efficiently learning and performing sequential mappings between orthography and phonology, leading to the phenotype of dyslexia. However, severe phonological processing difficulties may also lead on their own to a phenotype of dyslexia, even if serial order STM is not impaired.

A further question relates to domain-general factors that could explain serial order STM impairment. A number of studies have shown that children with dyslexia show difficulties in processing temporally organized information, at either fast or normal sequential presentation speeds (Laasonen et al., 2001, 2002; Romani et al., 2015). This finding is important as a number of theoretical models of serial order STM propose that serial order information is encoded using time-based codes (Burgess and Hitch, 1999, 2006; Brown et al., 2000); empirical evidence for this assumption has been provided recently (Hartley et al., 2016). Note, however, that time-based models of serial order STM represent only one among many different theoretical accounts of serial order STM (Hurlstone et al., 2014).

Another domain-general factor that needs to be considered is attentional impairment (Cowan, 1988, 2010, 2016). Sequentially presented items and their associations may depend on focus of attention capacity. Cowan et al. (2013) presented lists of three to nine words to adults with the task of determining the most interesting word from each list. A delayed test showed that associations had formed between adjacent list items when the lists were short enough to fit within the hypothesized scope of the focus of attention (Cowan, 2001) but not for longer lists. Attention could thus mediate memory of serial order. Some studies suggest that at least a subset of children with dyslexia present attentional deficits, and especially in the area of visual attention (Valdois et al., 2004). Lobier et al. (2014) observed reduced right superior parietal activation in dyslexic participants relative to controls like Martinez Perez et al. (2015),

and this for a task in which they had to categorize multiple verbal or non-verbal stimuli. However, the hypothesis of general attentional impairment does not fit with the dissociation observed between preserved item and impaired serial order visual STM tasks observed by Hachmann et al. (2014) and Martinez Perez et al. (2015). A related question here is whether the impairment observed for visual serial order STM tasks may have been driven by verbal encoding strategies. The fact that right intraparietal cortical areas, involved in non-verbal spatial and attentional processes, supported the serial order STM impairment in the study by Martinez Perez et al. (2015) speaks against this possibility; also, Hachmann et al. (2014) used difficult-to-verbalize nonsense drawings.

The mechanisms linking serial order STM and reading acquisition also need further exploration. Serial order STM may support ordered storage and output of letter-to-sound conversion processes during early reading acquisition and support the matching of letter serial positions within a letter string with those of visual word forms stored in LTM during visual word identification (Davis, 2010; Martinez Perez et al., 2012b). If serial order STM is causally involved in dyslexia, then dyslexic participants should have specific difficulties in these word decoding stages.

Future studies should focus more specifically on children populations with dyslexia, few studies exploring item and serial order STM in dyslexia having focused on children populations. Also, in order to determine the potential causal involvement of serial order STM deficits in dyslexia, longitudinal study

designs need to be used, in order to determine to what extent the STM deficits are predictive of the severity of later reading impairment as opposed to simply reflecting an associated deficit. Also, whether the serial order STM impairment precedes dyslexia or whether it arises at a later age, after the diagnosis of dyslexia, remains an open question. Importantly, future studies need to shed more light on the inconsistencies regarding the status of STM functioning in dyslexia observed in some studies. Given that these inconsistencies may be related to some degree of heterogeneity in dyslexic populations, population characteristics should be reported with as much detail as possible and include information about history of diagnosis, tests used to establish the diagnosis, linguistic, and socio-economic environment as well as a comprehensive characterization of both linguistic and non-linguistic cognitive abilities.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by grants T.1003.15 (Fonds de la Recherche Scientifique FNRS, Belgium) and PAI-IUAP P7/11 (Belgian Federal Science Policy) awarded to SM, and grant R01-HD21338 (NIH, USA) awarded to NC.

REFERENCES

- Attout, L., and Majerus, S. (2015). Working memory deficits in developmental dyscalculia: the importance of serial order. *Child Neuropsychol.* 21, 432–450. doi: 10.1080/09297049.2014.922170
- Avons, S. E., and Hanna, C. (1995). The memory-span deficit in children with specific reading disability: is speech rate responsible? *Br. J. Dev. Psychol.* 13, 303–311. doi: 10.1111/j.2044-835X.1995.tb00681.x
- Beneventi, H., Tonnessen, F. E., and Ersland, L. (2009). Dyslexic children show short-term memory deficits in phonological storage and serial rehearsal: an fMRI study. *Int. J. Neurosci.* 119, 2017–2043. doi: 10.1080/00207450903139671
- Binamé, F., and Poncelet, M. (2016). Order short-term memory capacity predicts nonword reading and spelling in first and second grade. *Read. Writ.* 29, 1–20. doi: 10.1007/s11145-015-9577-9
- Bogaerts, L., Szmalec, A., Hachmann, W. M., Page, M. P., and Duyck, W. (2015a). Linking memory and language: evidence for a serial-order learning impairment in dyslexia. *Res. Dev. Disabil.* 4, 106–122. doi: 10.1016/j.ridd.2015.06.012
- Bogaerts, L., Szmalec, A., Hachmann, W. M., Page, M. P., Woumans, E., and Duyck, W. (2015b). Increased susceptibility to proactive interference in adults with dyslexia? *Memory* 23, 268–277. doi: 10.1080/09658211.2014.882957
- Brady, S., Shankweiler, D., and Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *J. Exp. Child Psychol.* 35, 345–367. doi: 10.1016/0022-0965(83)90087-5
- Brown, G. D. A., Preece, T., and Hulme, C. (2000). Oscillator-based memory for serial order. *Psychol. Rev.* 107, 127–181. doi: 10.1037/0033-295X.107.1.127
- Burgess, N., and Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychol. Rev.* 106, 551–581. doi: 10.1037/0033-295X.106.3.551
- Burgess, N., and Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *J. Mem. Lang.* 55, 627–652. doi: 10.1016/j.jml.2006.08.005
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychol. Bull.* 104, 163–191. doi: 10.1037/0033-2909.104.2.163
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. doi: 10.1017/S0140525X01003922
- Cowan, N. (2010). The focus of attention as observed in visual working memory tasks: making sense of competing claims. *Neuropsychologia* 49, 1401–1406. doi: 10.1016/j.neuropsychologia.2011.01.035
- Cowan, N. (2016). Working memory maturation: can we get at the essence of cognitive growth? *Perspect. Psychol. Sci.* 11, 239–264. doi: 10.1177/1745691615621279
- Cowan, N., Donnell, K., and Sauls, J. S. (2013). A list-length constraint on incidental item-to-item associations. *Psychon. Bull. Rev.* 20, 1253–1258. doi: 10.3758/s13423-013-0447-7
- Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychol. Rev.* 117, 713–758. doi: 10.1037/a0019738
- Gathercole, S. E., and Alloway, T. P. (2006). Practitioner review: short-term and working memory impairments in neurodevelopmental disorders: diagnosis and remedial support. *J. Child Psychol. Psychiatry* 47, 4–15. doi: 10.1111/j.1469-7610.2005.01446.x
- Gathercole, S. E., and Baddeley, A. D. (1993). Phonological working memory: a critical building block for reading development and vocabulary acquisition? *Eur. J. Psychol. Educ.* 8, 259–272. doi: 10.1007/BF03174081

- Gathercole, S. E., Frankish, C. R., Pickering, S. J., and Peaker, S. (1999). Phonotactic influences on short-term memory. *J. Exp. Psychol. Hum. Learn. Mem.* 25, 84–95.
- Hachmann, W. M., Bogaerts, L., Szmalec, A., Woumans, E., Duyck, W., and Job, R. (2014). Short-term memory for order but not for item information is impaired in developmental dyslexia. *Ann. Dyslexia* 64, 121–136. doi: 10.1007/s11881-013-0089-5
- Hartley, T., Hurlstone, M. J., and Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: a model and tests of its stimulus-driven mechanism. *Cogn. Psychol.* 87, 135–178. doi: 10.1016/j.cogpsych.2016.05.001
- Henson, R. N. A. (1998). Short-term memory for serial order: the start-end model. *Cogn. Psychol.* 36, 73–137. doi: 10.1006/cogp.1998.0685
- Hulme, C., Maughan, S., and Brown, G. D. (1991). Memory for familiar and unfamiliar words: evidence for a long-term memory contribution to short-term memory span. *J. Mem. Lang.* 30, 685–701. doi: 10.1016/0749-596X(91)90032-F
- Hurlstone, M. J., Hitch, G. J., and Baddeley, A. D. (2014). Memory for serial order across domains: an overview of the literature and directions for future research. *Psychol. Bull.* 140, 339–373. doi: 10.1037/a0034221
- Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., and Holmes, J. (2016). Following instructions in a virtual school: does working memory play a role? *Mem. Cogn.* 44, 580–589. doi: 10.3758/s13421-015-0579-2
- Laasonen, M., Service, E., and Virsu, V. (2001). Temporal order and processing acuity of visual, auditory, and tactile perception in developmentally dyslexic young adults. *J. Cogn. Affect. Behav. Neurosci.* 1, 394–410. doi: 10.3758/CABN.1.4.394
- Laasonen, M., Service, E., and Virsu, V. (2002). Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. *Brain Lang.* 80, 340–354. doi: 10.1006/brln.2001.2593
- Laasonen, M., Virsu, V., Oinonen, S., Sandbacka, M., Salakari, A., and Service, E. (2012). Phonological and sensory short-term memory are correlates and both affected in developmental dyslexia. *Read. Writ.* 25, 2247–2273. doi: 10.1007/s11145-011-9356-1
- Leclercq, A. L., and Majerus, S. (2010). Serial order short-term memory predicts vocabulary development: evidence from a longitudinal study. *Dev. Psychol.* 46, 417–427. doi: 10.1037/a0018540
- Lobier, M. A., Peyrin, C., Pichat, C., Le Bas, J. F., and Valdois, S. (2014). Visual processing of multiple elements in the dyslexic brain: evidence for a superior parietal dysfunction. *Front. Hum. Neurosci.* 8:479. doi: 10.3389/fnhum.2014.00479
- Majerus, S., Martinez Perez, T., and Oberauer, K. (2012). Two distinct origins of long-term learning effects in verbal short-term memory. *J. Mem. Lang.* 66, 38–51. doi: 10.1016/j.jml.2011.07.006
- Majerus, S., Poncelet, M., Greffe, C., and Van der Linden, M. (2006a). Relations between vocabulary development and verbal short-term memory: the importance of short-term memory for serial order information. *J. Exp. Child Psychol.* 93, 95–119. doi: 10.1016/j.jecp.2005.07.005
- Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., et al. (2006b). The left intraparietal sulcus and verbal short-term memory: focus of attention or serial order? *Neuroimage* 32, 880–891. doi: 10.1016/j.neuroimage.2006.03.048
- Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., and Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: evidence from an incidental phonotactic learning paradigm. *J. Mem. Lang.* 51, 297–306. doi: 10.1016/j.jml.2004.05.002
- Martinez Perez, T., Majerus, S., Mahot, A., and Poncelet, M. (2012a). Evidence for a specific impairment of serial order short-term memory in dyslexic children. *Dyslexia* 18, 94–109. doi: 10.1002/dys.1438
- Martinez Perez, T., Majerus, S., and Poncelet, M. (2012b). The contribution of short-term memory for serial order to early reading acquisition: evidence from a longitudinal study. *J. Exp. Child Psychol.* 111, 708–723. doi: 10.1016/j.jecp.2011.11.007
- Martinez Perez, T., Majerus, S., and Poncelet, M. (2013). Impaired short-term memory for order in adults with dyslexia. *Res. Dev. Disabil.* 34, 2211–2233. doi: 10.1016/j.ridd.2013.04.005
- Martinez Perez, T., Poncelet, M., Salmon, E., and Majerus, S. (2015). Functional alterations in order short-term memory networks in adults with dyslexia. *Dev. Neuropsychol.* 40, 407–429. doi: 10.1080/87565641.2016.1153098
- Nairne, J. S., and Kelley, M. R. (2004). Separating item and order information through process dissociation. *J. Mem. Lang.* 50, 113–133. doi: 10.1016/j.jml.2003.09.005
- Page, M. P. A., and Norris, D. (1998). The primacy model: a new model of immediate serial recall. *Psychol. Rev.* 105, 761–781. doi: 10.1037/0033-295X.105.4.761-781
- Ramus, F., Marshall, C. R., Rosen, S., and van der Lely, H. K. (2013). Phonological deficits in specific language impairment and developmental dyslexia: towards a multidimensional model. *Brain* 136(Pt 2), 630–645. doi: 10.1093/brain/awt356
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126(Pt 4), 841–865. doi: 10.1093/brain/awg076
- Romani, C., Tsouknida, E., and Olson, A. (2015). Encoding order and developmental dyslexia: a family of skills predicting different orthographic components. *Q. J. Exp. Psychol.* 68, 99–128. doi: 10.1080/17470218.2014.938666
- Serniclaes, W., Van Heghe, S., Mousty, P., Carre, R., and Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *J. Exp. Child Psychol.* 87, 336–361. doi: 10.1016/j.jecp.2004.02.001
- Silton, R. L., Heller, W., Towers, D. N., Engels, A. S., Spielberg, J. M., Edgar, J. C., et al. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top-down attentional control. *Neuroimage* 50, 1292–1302. doi: 10.1016/j.neuroimage.2009.12.061
- Snowling, M. (2000). *Dyslexia*, 2nd Edn. Oxford: Blackwell Publishing.
- Snowling, M., Goulandris, N., and Defty, N. (1996). A longitudinal study of reading development in dyslexic children. *J. Educ. Psychol.* 88, 653–669. doi: 10.1037/0022-0663.88.4.653
- Staels, E., and Van den Broeck, W. (2014). Order short-term memory is not impaired in dyslexia and does not affect orthographic learning. *Front. Hum. Neurosci.* 8:732. doi: 10.3389/fnhum.2014.00732
- Staels, E., and Van den Broeck, W. (2015). No solid empirical evidence for the SOLID (serial order learning impairment) hypothesis of dyslexia. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 650–669. doi: 10.1037/xlm0000054
- Szenkovits, G., and Ramus, F. (2005). Exploring dyslexics' phonological deficit I: lexical vs sub-lexical and input vs output processes. *Dyslexia* 11, 253–268. doi: 10.1002/dys.308
- Szmalec, A., Loncke, M., Page, M. P., and Duyck, W. (2011). Order or disorder? Impaired Hebb learning in dyslexia. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1270–1279. doi: 10.1037/a0023820
- Valdois, S., Bosse, M. L., and Tainturier, M. J. (2004). The cognitive deficits responsible for developmental dyslexia: review of evidence for a selective visual attentional disorder. *Dyslexia* 10, 339–363. doi: 10.1002/dys.284
- van Dijck, J. P., Abrahamse, E. L., Majerus, S., and Fias, W. (2013). Spatial attention interacts with serial-order retrieval from verbal working memory. *Psychol. Sci.* 24, 1854–1859. doi: 10.1177/0956797613479610
- van Dijck, J. P., and Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- Wang, X., Xuan, Y., and Jarrold, C. (2016). Using a process dissociation approach to assess verbal short-term memory for item and order information in a sample of individuals with a self-reported diagnosis of dyslexia. *Front. Psychol.* 7:208. doi: 10.3389/fpsyg.2016.00208
- Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Majerus and Cowan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Spatial-Sequential Working Memory in Younger and Older Adults: Age Predicts Backward Recall Performance within Both Age Groups

Louise A. Brown*

School of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK

OPEN ACCESS

Edited by:

Snehlata Jaswal,
Indian Institute of Technology
Jodhpur, India

Reviewed by:

Chunyan Guo,
Capital Normal University, China
Colin Joseph Hamilton,
University of Northumbria
at Newcastle, UK

*Correspondence:

Louise A. Brown
l.brown@strath.ac.uk
orcid.org/0000-0003-3520-6175

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 30 June 2016

Accepted: 20 September 2016

Published: 04 October 2016

Citation:

Brown LA (2016) Spatial-Sequential Working Memory in Younger and Older Adults: Age Predicts Backward Recall Performance within Both Age Groups. *Front. Psychol.* 7:1514. doi: 10.3389/fpsyg.2016.01514

Working memory is vulnerable to age-related decline, but there is debate regarding the age-sensitivity of different forms of spatial-sequential working memory task, depending on their passive or active nature. The functional architecture of spatial working memory was therefore explored in younger (18–40 years) and older (64–85 years) adults, using passive and active recall tasks. Spatial working memory was assessed using a modified version of the Spatial Span subtest of the Wechsler Memory Scale – Third Edition (WMS-III; Wechsler, 1998). Across both age groups, the effects of interference (control, visual, or spatial), and recall type (forward and backward), were investigated. There was a clear effect of age group, with younger adults demonstrating a larger spatial working memory capacity than the older adults overall. There was also a specific effect of interference, with the spatial interference task (spatial tapping) reliably reducing performance relative to both the control and visual interference (dynamic visual noise) conditions in both age groups and both recall types. This suggests that younger and older adults have similar dependence upon active spatial rehearsal, and that both forward and backward recall require this processing capacity. Linear regression analyses were then carried out within each age group, to assess the predictors of performance in each recall format (forward and backward). Specifically the backward recall task was significantly predicted by age, within both the younger and older adult groups. This finding supports previous literature showing lifespan linear declines in spatial-sequential working memory, and in working memory tasks from other domains, but contrasts with previous evidence that backward spatial span is no more sensitive to aging than forward span. The study suggests that backward spatial span is indeed more processing-intensive than forward span, even when both tasks include a retention period, and that age predicts backward spatial span performance across the adult lifespan, within both younger and older adulthood.

Keywords: cognitive aging, ageing, spatial-sequential working memory, spatio-sequential, visual-spatial, visuo-spatial sketchpad, central executive attention, Corsi blocks test

INTRODUCTION

An important factor in spatial working memory performance is the degree of active processing involved in the task (Cornoldi and Vecchi, 2003). Passive storage involves retaining information which has not been modified after encoding, while active processing requires transforming, manipulating, or integrating information. Cornoldi and Vecchi (2003) argued that the degree

of processing must be conceived along an activity continuum. Visuo-spatial tasks which require storage and retrieval do require some active processing in the form of rehearsal. However, the degree of active processing is very low compared with a task which additionally requires active manipulation of the information before the participant can provide the appropriate response. Multiple component models of working memory (e.g., Baddeley and Hitch, 1974; Logie, 1995, 2003, 2011; Baddeley, 2000, 2007, 2012) conceive of such a distinction between storage and processing. These models comprise specialized verbal and visuo-spatial storage components (the *phonological loop* and *visuo-spatial sketchpad*, respectively), as well as domain-general processing capacity (the *central executive*). Specifically regarding storage and processing in the visuo-spatial domain, Logie's model of working memory comprises a passive visual store component, which temporarily retains visual detail (*visual cache*), and an active spatial rehearsal mechanism (*inner scribe*), which stores spatial-sequential codes. The inner scribe can also refresh the contents of the visual store, which is otherwise subject to rapid (2 s) decay (Logie, 2011). This active processing comes at a cognitive cost, however, and draws upon the capacity of the central executive (see also Rudkin et al., 2007; Vandierendonck, 2016).

Fluid cognitive abilities such as working memory are particularly vulnerable to age-related decline, and have been shown to be subject to linear decline throughout the adult lifespan, often from the early 20s (Park et al., 2002; Logie and Maylor, 2009; Johnson et al., 2010). Regarding the potential role of active, central executive processing in age-related decline, previous research has suggested that older adults exhibit specific deficits when active information processing is involved, while age differences are relatively minimal in passive tasks (Phillips and Hamilton, 2001; Bopp and Verhaeghen, 2005; Braver and West, 2008). Vecchi et al. (2005; see also Vecchi and Cornoldi, 1999; Mammarella et al., 2013) showed that active working memory tasks, in both the visuo-spatial and verbal domains, are subject to greater age-related deficits, and that age effects are typically seen earlier in the lifespan for active tasks relative to passive ones.

One validated task for assessing the performance of spatial-sequential working memory (Logie, 2011) is the Corsi blocks task (Milner, 1971; Corsi, 1972; De Renzi and Nichelli, 1975; Smyth and Scholey, 1994; Berch et al., 1998; Della Sala et al., 1999). Often described simply as *spatial span*, the task involves presenting a series of spatial sequences and asking participants to recall them either immediately or following a maintenance period. The sequences take the form of movements between various spatial locations. They may be presented to participants either via a computer screen, or using a board featuring an array of blocks that are tapped by the researcher. The latter is the case for the Spatial Span subtest of the Wechsler Memory Scale (3rd Edition; WMS-III, Wechsler, 1998). Particularly with a maintenance period inserted after presentation and before recall, the task requires active rehearsal of the information (Cornoldi and Vecchi, 2003; Logie, 2011).

The degree of active processing involved in the spatial span task is amenable to manipulation, however, with a typical comparison being whether sequences are recalled either in the

same order as the researcher (forward) or the opposite order (backward). Given the above findings comparing passive vs. active working memory tasks, one could predict greater age-related decline in a backward recall version of the spatial span task, compared with forward recall. However, the findings in this respect have been mixed. For example, Hester et al. (2004) investigated the effect of age on forward and backward spatial span and the age-related decline in performance was equivalent in both measures. Similarly, Wilde et al. (2004) analyzed data from the WMS-III (Wechsler, 1997) standardization sample ($n = 1,250$). While forward recall was performed better than backward recall overall in this study, the difference between the two recall types was not enhanced in older age. The authors concluded that backward spatial recall is no more age-sensitive than forward recall (see also Myerson et al., 2003). Indeed, Hester et al. proposed that the central executive component of working memory is recruited for successful performance in both versions of the task. This could be supported by the argument that spatial-sequential working memory, even in forward recall, requires central executive processing (Vecchi and Richardson, 2001; Logie, 2011).

In younger adults, Vandierendonck et al. (2004) compared the difficulty of forward and backward spatial span and found no consistent effect of recall format (see also Wilde and Strauss, 2002, for a clinical sample). However, a targeted interference paradigm was used across varying sequence lengths, designed to disrupt specific forms of processing during encoding and to determine the underlying processes involved. Concurrent spatial tapping was intended to disrupt spatial processing, while random interval generation was performed in order to suppress central executive functioning. Vandierendonck et al. (2004) showed that active spatial processing was required throughout the task, even at lower levels of complexity, demonstrating the critical nature of this processing resource (Vecchi and Richardson, 2001; Hamilton et al., 2003). However, additionally, specifically when reaching span and supra-span levels of complexity, domain-general central executive resources were increasingly employed at the most challenging levels of the task (Vandierendonck et al., 2004; see also Hamilton et al., 2003; Thompson et al., 2006; Logie, 2011). Importantly, though, both forward and backward recall drew upon central executive resources (Rudkin et al., 2007; Vandierendonck, 2016).

Clearly, there has been debate in the literature regarding the extent of active processing required by different spatial-sequential working memory tasks, and particularly regarding the possibility that the more active backward spatial span task is especially impaired by aging. The current study assessed whether or not a differential age-related decline would be evident between conditions of relatively passive and active recall (forward vs. backward spatial span). Phase one assessed recall of sequences in the same order as presentation and required no manipulation of the material, only active spatial rehearsal. Phase two of the task, on the other hand, assessed recall of sequences in the reverse order of presentation and therefore required active rehearsal as well as active manipulation of the information prior to recall. If aging differentially affects more active spatial working memory

tasks, then a greater age-related deficit would be predicted when backward recall is performed.

Regarding the potential for differential use of visuo-spatial sketchpad resources by younger and older adults, a targeted interference paradigm was additionally employed in the current study. This was to assess the extent to which younger and older adults each rely upon the visual cache and inner scribe mechanisms when performing a spatial-sequential working memory task. As discussed above, one would predict that spatial processing would be employed throughout successful task performance and that a spatial interference task would therefore be disruptive to the span level achieved. However, as current cognitive aging theory predicts less specialized cognitive processing with aging, and more generalized compensatory processing (Reuter-Lorenz and Park, 2014), it is possible that older adults may show a different pattern of interference effects than younger adults. For example, they may show less spatial interference, and/or more interference from the visuo-spatial sketchpad resource which is less specialized for this task (the visual cache). Indeed, Fiore et al. (2012) showed that updating in visuo-spatial working memory is age-sensitive, and suggested that older adults engage in less active rehearsal in spatial working memory than younger adults. This was on the basis of greater age effects at early sequence items, in conjunction with intact recency effects in older adults. On the other hand, Jenkins et al. (2000; see also Jenkins et al., 1999) assessed the effect of a visuo-spatial interference task on performance of a spatial working memory task, in which spatial locations were to be recalled (without sequential order). This research showed that, although capacity was reduced with aging, a visuo-spatial concurrent task (tapping on individual colored locations) was no more disruptive to older than younger adults. However, note that the memory task in these studies was not spatial-sequential, and the interference task was not specific to disrupting either the visual cache or inner scribe components of the visuo-spatial sketchpad. Thus, regarding the functional architecture of the visuo-spatial sketchpad, the hypothesis for the current study was that this may be affected by aging, and that differential visual and/or spatial interference effects may be observed across younger and older adults as a consequence of different processing abilities.

In summary, the first key aim of this research was to establish whether or not the degree of active processing in a spatial-sequential working memory task affects the extent of age-related decline observed. The second aim was to establish whether or not the functional architecture underlying spatial working memory may be subject to age-related change. Specifically, this study investigated whether or not there are age-related differences in the reliance on the visual store and inner scribe mechanisms of working memory.

MATERIALS AND METHODS

Design

This study took the form of a 2 (age group; younger, older) \times 2 (recall type; forward, backward – repeated measures) \times 3 (interference; control, visual, spatial – between participants)

mixed factorial design. Task performance was assessed using a mean span measure of capacity (described below).

Participants

This study was carried out in accordance with the recommendations of the Ethics of Research on Human Participants, Glasgow Caledonian University, with written informed consent from all subjects, in accordance with the Declaration of Helsinki. There were 75 younger (18–40 years) and 75 older (64–85 years) participants. The younger group comprised 30 males and 45 females with a mean age of 27.93 ($SD = 5.98$). Their mean number of years of formal education was 17.40 ($SD = 3.17$). The older participants comprised 32 males and 43 females, with a mean age of 73.62 years ($SD = 6.11$) and a mean number of years of education of 11.63 ($SD = 2.61$). The older adults were screened for signs of cognitive impairment, and were required to achieve a score of 25 from the possible 30 in the Mini-Mental State Examination (MMSE; Folstein et al., 1975). The mean MMSE score was 28.15 ($SD = 1.40$). **Table 1** presents the participant demographics by age group and interference condition. None of the participants had carried out the task before, and they each received a small participation fee.

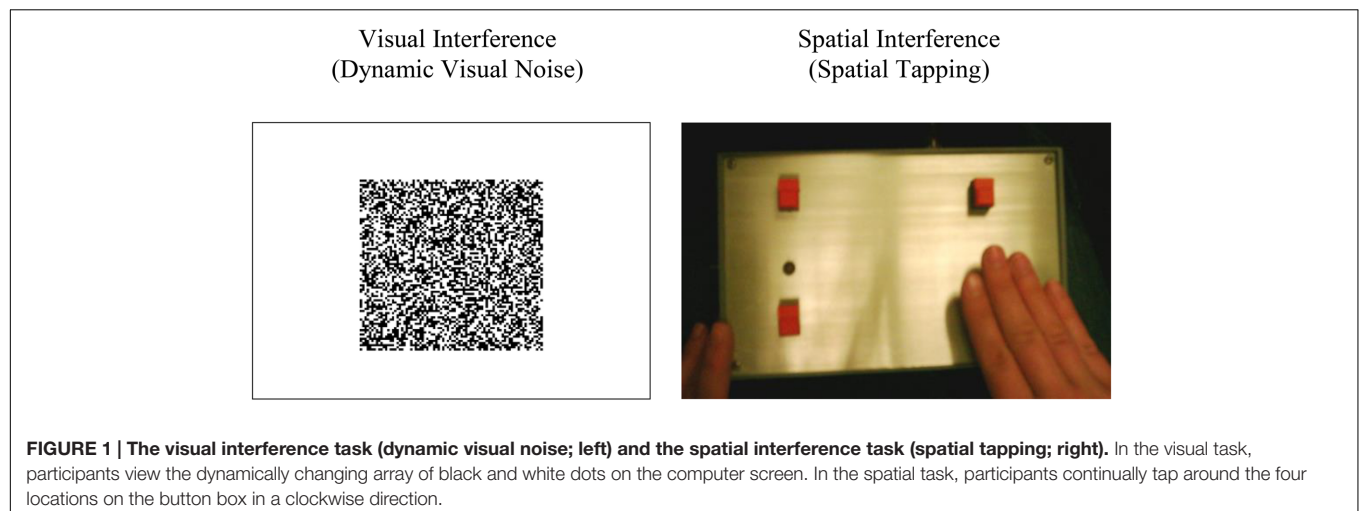
Materials

A version of the Corsi blocks test was used for measuring spatial working memory span. This was a modified version of the Spatial Span subtest of the WMS-III (Wechsler, 1998). The WMS-III Spatial Span board features 10 irregularly spaced blue cubes set upon a white rectangular board, with each cube featuring an identifying number on only the researcher's side. The board measures approximately 28 cm \times 21.5 cm, and the cubes measure 3 cm³. The standard task comprises eight sequence levels (ranging from 2 to 9 blocks in length) with two trials at each level. In order to enhance the sensitivity of the task, one additional sequence was created per sequence level, thus allowing three trials at each level. New sequences were generated by adopting a selection without replacement procedure for each of the numbers 1–10. Sequences were then fixed, and each new sequence was administered after the standard two. Consistent with the protocol of the standard Spatial Span test, in the backward recall phase of the task the administration order of the forward phase was reversed within each level, such that the third (new) sequences in each level were administered first. Additionally, so that the sequences were not identical to those administered in the forward recall phase, the order of each sequence was reversed for the backward recall phase, again, as in the standard Spatial Span procedure. Finally, in order to increase working memory demand by requiring the use of rehearsal, the task incorporated a 10 s delay period (retention interval) between presentation and recall.

For some participants, either visual or spatial interference took place during the 10 s retention interval of the spatial working memory task. The visual interference took the form of dynamic visual noise (DVN; **Figure 1**), which interferes specifically with the operation of the passive visual store in the visuo-spatial sketchpad (Quinn and McConnell, 1996a,b) and reduces visual imagery and working memory for visual details (e.g., McConnell and Quinn, 2004; Dean et al., 2008; Darling et al., 2009; Borst

TABLE 1 | Means (with standard deviations) for each participant group's demographic data.

	Younger			Older		
	Control	Visual interference	Spatial interference	Control	Visual interference	Spatial interference
Age	28.16 (± 6.26)	28.08 (± 5.68)	27.56 (± 6.21)	74.44 (± 6.21)	73.08 (± 5.58)	73.33 (± 6.67)
Sex (M:F)	10:15	12:13	8:17	9:16	9:16	14:11
Years education	17.36 (± 3.11)	17.28 (± 3.12)	17.56 (± 3.40)	11.84 (± 2.58)	11.60 (± 2.69)	11.44 (± 2.65)
MMSE	—	—	—	28.20 (± 1.26)	28.08 (± 1.41)	28.16 (± 1.57)



et al., 2012). The DVN was a computer-generated display of small black and white 'dots' which randomly change between black and white in an even and continuous fashion across the array. The array measured 320 pixels \times 320 pixels (or approximately 12 cm²), and comprised 80 \times 80 dots (6400, each 16 pixels in area), which randomly changed between black and white in a continuous, evenly distributed fashion. The rate of change was relatively high, at 30% (1920 dots) per second (McConnell and Quinn, 2000; Dean et al., 2005). The spatial interference was a manual spatial tapping task, with movements to known, predictable locations, which interferes specifically with spatial rehearsal and the inner scribe of working memory, and with minimal central executive involvement (Farmer et al., 1986; Smyth and Pendleton, 1989; Quinn, 1994; Della Sala et al., 1999; Darling et al., 2009). A handheld box was used for this (Figure 1), measuring 21.5 cm \times 13 cm \times 7.5 cm and featuring four buttons (each 1.2 cm \times 1.6 cm \times 1.1 cm) spaced in a rectangular formation (11 cm \times 5.5 cm). An electronic counter was linked to the handheld box for the purpose of calculating the total number of taps per trial.

Procedure

Participants were randomly allocated to one of the three interference conditions. Each participant sat at a desk opposite the researcher, and the Spatial Span board was positioned between them. From the perspective of the participant, a laptop was positioned to the left of the Spatial Span board. For those in the spatial interference condition, a handheld button box was

also placed on the participant's lap. The first phase of the task measured spatial working memory span with forward recall, and the second phase measured backward recall (as per the standard Spatial Span test; see also Park et al., 2002).

The participant was first instructed to touch the same blocks that the researcher touched, in the same order. Depending on participant choice, one or two practice trials were completed prior to beginning the experimental trials. For a given trial under control conditions, the procedure was as follows: the researcher tapped out the sequence, at a rate of approximately one tap per second, before immediately pressing the button of the mouse, which produced a tone from the laptop; the participant then viewed the blank laptop screen for a period of 10 s, until the word *recall* was presented; the participant then attempted to touch the same blocks in the same order as the researcher. In the visual interference condition, the procedure was the same except that, during the 10 s retention interval, the participant viewed DVN on the screen. In the spatial interference condition, the procedure was the same as in the control condition except that the participant was also required to tap around the buttons on the handheld box in clockwise order with their preferred hand, at their own pace, during the 10 s retention interval. The participant was specifically instructed to view the blank screen and not the handheld box. Some time was provided to allow these participants to familiarize themselves with the spatial tapping task before combining it with the memory task. The electronic counter, which was linked to the box, recorded the number of times the buttons had been tapped within each trial.

Regarding memory task performance, the researcher recorded the numbers of the blocks that had been tapped and in which order, and provided performance feedback to the participant. The task continued either until all available trials had been completed, or the participant failed to recall correctly at least one of the three trials from a given level of complexity. Spatial working memory span with forward recall was taken to be the mean size of the last three correctly recalled sequences in this phase.

The second phase of the procedure then measured spatial working memory span with backward recall. The researcher informed participants that the task would be carried out again, except that they were now required to try to reproduce the sequence in the reverse order, beginning with the last cube and working backward. Again, following at least one practice trial, the experimental trials were administered under the same conditions as in forward recall, either until all trials had been administered, or until the participant had failed to recall correctly at least one trial from a given level. The mean size of the last three correctly recalled sequences was calculated¹.

Analyses

The *mean span* data were first analyzed using a 2 (age group) \times 2 (recall type) \times 3 (interference) mixed factorial Analysis of Variance (ANOVA). Follow-up tests were either planned comparisons or Bonferroni-corrected *t*-tests, as appropriate and specified below. Data were then also analyzed using a series of linear regression analyses, in which only the control and spatial interference conditions were included². For the regression analyses, collinearity diagnostics were within acceptable levels [all variance inflation factor (VIF) < 1.47; all tolerance values > 0.68].

RESULTS

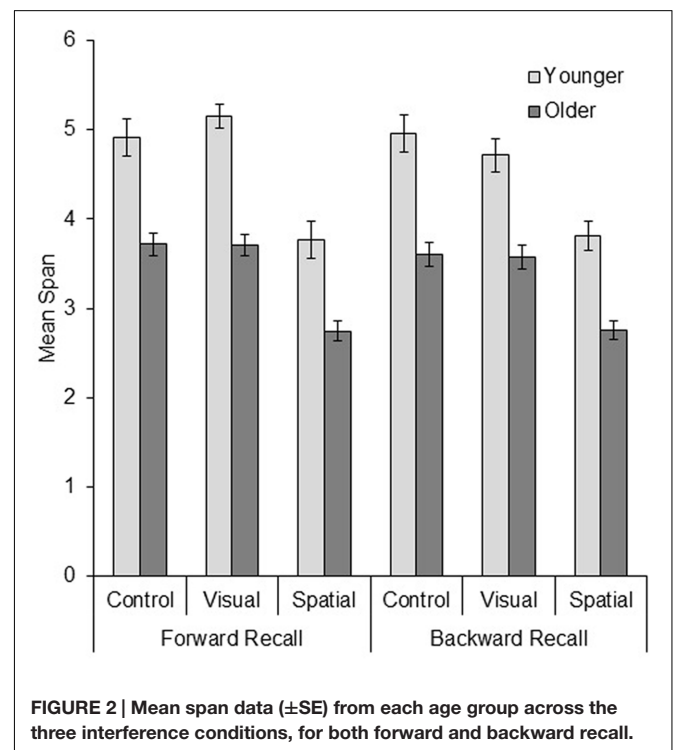
Regarding performance of the spatial interference (tapping) task, during forward recall, the mean number of taps per trial in the younger adults was 15.24 ($SD = 4.32$), and in older adults this was 13.88 ($SD = 3.15$). During backward recall, the mean number of taps in the younger adults was 16.29 ($SD = 5.08$), and in older adults this was 14.60 ($SD = 3.68$). A mixed factorial ANOVA revealed only a significant effect of recall type, $F(1,48) = 20.37$, $MSE = 0.96$, $p < 0.001$, $\eta_p^2 = 0.30$, in which the number of taps was slightly higher in backward recall ($M = 15.44$, $SD = 4.47$) than in forward recall ($M = 14.56$, $SD = 3.81$). All other effects were not significant (all $p > 0.19$). It is possible that the tapping rate increased slightly with practice. However, it is notable that there were no reliable age effects, and no interaction between the two variables. Particularly as the difference between the two recall types was very small (approximately one tap per trial), tapping rate will therefore not be considered further.

¹ For 4 of the 25 older participants in the spatial interference condition in forward recall, and 6 of the 25 older participants in the spatial interference condition in backward recall, only one or two successful trials had taken place. In these cases, the mean was taken of the available successful trials.

² However, the same pattern of findings is observed when including the visual interference condition data (collapsed with the control condition). Note also that, for the linear regression analyses, the older adult group was reduced from 50 to 49 participants, due to one missing datapoint.

The mean spatial working memory span data are presented in **Figure 2**. The data pattern shows that spatial working memory capacity appears lower for older compared with younger adults, and when the task was carried out alongside the spatial interference condition as compared with both the control and visual interference conditions.

A mixed factorial ANOVA indeed revealed significant effects of age group, $F(1,144) = 106.12$, $MSE = 1.03$, $p < 0.001$, $\eta_p^2 = 0.42$, with younger adults outperforming older adults, and interference, $F(2,144) = 33.86$, $MSE = 1.03$, $p < 0.001$, $\eta_p^2 = 0.32$. Follow-up planned comparisons revealed no significant difference between the control and visual interference conditions, $t(98) = 0.07$, $p = 0.94$, but that performance was poorer with spatial than visual interference, $t(98) = 5.65$, $p < 0.001$. There were also non-significant trends for the main effect of recall type, $F(1,144) = 3.11$, $MSE = 0.228$, $p = 0.080$, $\eta_p^2 = 0.02$, and the interaction between interference and recall type, $F(2,144) = 2.86$, $MSE = 0.228$, $p = 0.061$, $\eta_p^2 = 0.04$. The interactions between age group and recall type, $F(1,144) = 0.11$, $MSE = 0.228$, $p = 0.74$, $\eta_p^2 = 0.001$, between age group and interference, $F(2,144) = 0.50$, $MSE = 1.03$, $p = 0.61$, $\eta_p^2 = 0.007$, and the three-way interaction, $F(2,144) = 1.49$, $MSE = 0.228$, $p = 0.23$, $\eta_p^2 = 0.02$, were clearly not significant. Bonferroni-corrected paired *t*-tests were used to investigate the trend for the interaction between interference and recall type, analyzing the effect of recall type within each interference condition (with data collapsed across age group). This trend appears to have been driven by an effect of recall type being significant only in the visual interference condition, $t(49) = 3.23$, $p < 0.01$ (all other $p > 0.72$), and may have been influenced by a slightly raised score



in the visual interference condition in younger adults' forward recall.

The ANOVA, then, clearly shows no differential effects of recall type either by interference or age group. To supplement this analysis, however, and to discover the predictors of forward and backward recall performance within each age group, a series of linear regression analyses were then carried out. The data from each age group were therefore analyzed individually, in order to establish which factors were predictive of performance in each age group, using each outcome measure (forward and backward mean span). The relevant correlation matrices are presented in **Tables 2** and **3**. **Table 2** shows that, in younger adults, age is positively related with years of education, as the youngest adults would not yet have completed their education. More interestingly, and in line with the effect of interference described above, the presence of spatial interference was significantly related to lower capacity, in both forward and backward recall. Regarding age, there were non-significant trends for this to be negatively correlated with both forward ($p = 0.067$) and backward ($p = 0.084$) recall. Finally, forward and backward spatial span were positively correlated. **Table 3** highlights that, in older adults, increased age was associated with lower scores on the MMSE. The presence of spatial interference was again significantly related to lower forward and backward span scores in older people, and the two methods of recall were also positively correlated. Backward span was positively related to years of education and MMSE score and, while there was clearly no significant relationship between age and forward span, backward span showed a non-significant trend for a negative association with age ($p = 0.067$).

Linear regression analysis was first carried out on the younger adult data, to establish the predictors of spatial working memory capacity with forward recall. Age, sex, years of education, and interference (control or spatial) were entered into the analysis. The model was significant, $F(4,45) = 5.80$, $p = 0.001$, and predicted 28% of the variance in forward recall ($R = 0.58$, adjusted $R^2 = 0.28$, $SE = 1.00$). However, as shown in **Table 4**, only interference ($\beta = -0.49$, $p < 0.001$) significantly predicted forward recall. For the backward recall data, the model was again significant, $F(4,45) = 6.34$, $p < 0.001$, and predicted 30% of the variance ($R = 0.60$, adjusted $R^2 = 0.30$, $SE = 0.93$). However, this time, both interference ($\beta = -0.54$, $p < 0.001$) as well as age ($\beta = -0.34$, $p = 0.025$) served as significant predictors.

In the older adults, the same variables were entered into a linear regression analysis, along with the additional MMSE variable. Again, the model was significant, $F(5,43) = 7.99$, $p < 0.001$, and predicted 42% of the variance in forward recall ($R = 0.69$, adjusted $R^2 = 0.42$, $SE = 0.58$). Only interference ($\beta = -0.64$, $p < 0.001$) significantly predicted forward recall, although a non-significant trend can be noted in relation to the MMSE scores ($\beta = 0.23$, $p = 0.067$). The significant model of the backward recall data, $F(5,43) = 11.39$, $p < 0.001$, predicted 52% of the variance in performance ($R = 0.76$, adjusted $R^2 = 0.52$, $SE = 0.50$). Both interference ($\beta = -0.64$, $p < 0.001$) as well as age ($\beta = -0.23$, $p = 0.044$) significantly predicted performance. Thus, in both younger and older adults, age is a significant predictor of, specifically, backward recall.

Additionally, however, sex ($\beta = -0.24$, $p = 0.025$) and years of education ($\beta = 0.21$, $p = 0.043$) also significantly contributed to the model of backward span in older people, suggesting that better performance was associated with males, and a higher number of years of education.

DISCUSSION

This study investigated spatial-sequential working memory performance in younger and older adults. The working memory task varied according to the more passive (forward) or active (backward) nature of the recall format. Furthermore, interference tasks carried out during the retention interval in the tasks were intended to disrupt temporary visual storage (the operation of the visual cache) or spatial processing (inner scribe functioning). The study was aimed at establishing whether or not aging is associated with a greater decline in active vs. passive spatial working memory, and the extent to which the functional architecture of the visuo-spatial sketchpad is affected by age. Initial analyses showed that recall type did not reliably affect performance either in younger or older adults, and that specialized spatial processing, as indicated by spatial interference effects, appears to be used by both age groups in both passive and active spatial span tasks. However, supplementary linear regression analyses within each age group, which included additional demographic variables, showed that age significantly predicts specifically backward spatial span performance within both younger and older adults, suggesting that the more active recall format is more sensitive to the aging process throughout the adult lifespan.

In terms of the overall effect of aging on spatial working memory capacity, the results clearly showed a reliable age-related deficit, in both forward and backward recall. Robbins et al. (1998) argued that spatial short-term memory performance, as measured by the Corsi with immediate recall, exhibits little decline in older age. In contrast, Moffat et al. (2001; see also Jenkins et al., 2000) concluded that spatial working memory is markedly impaired by the aging process. The present results support the latter suggestion, that spatial working memory capacity is reliably reduced with aging. Note, however, that the task presently employed was designed to place significant demands on spatial working memory, due to the requirement to maintain the sequences over a 10 s delay period before recall. As noted previously, active processing in the form of spatial rehearsal is already necessary for successful performance of a spatial working memory task which features a delay period. However, in the present study, the extent of active processing was directly assessed, in order to compare active rehearsal with the requirement also to manipulate the information prior to recall (Cornoldi and Vecchi, 2003). The latter was expected to draw more heavily upon the resources of the domain-general central executive in the working memory system (Baddeley, 2007; Rudkin et al., 2007; Logie, 2011; Vandierendonck, 2016).

Passive vs. Active Processing

There has been debate regarding the potential role of relatively active processing in age-related cognitive decline. Previous

TABLE 2 | Pearson correlation coefficients between each variable included in the younger adult linear regression analyses.

	1	2	3	4	5	6
(1) Age	—	0.19	0.53***	−0.05	−0.26	−0.25
(2) Sex (0 = M, 1 = F)		—	0.03	0.08	−0.17	−0.13
(3) Years education			—	0.03	−0.25	−0.07
(4) Interference (0 = control, 1 = spatial tapping)				—	−0.49***	−0.52***
(5) Mean span – Forward					—	0.76***
(6) Mean span – Backward						—

NB: All correlations based upon $N = 50$. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 3 | Pearson correlation coefficients between each variable included in the older adult linear regression analyses.

	1	2	3	4	5	6	7
(1) Age	—	0.03	−0.05	−0.40**	−0.09	−0.05	−0.26
(2) Sex (0 = M, 1 = F)		—	−0.02	−0.08	−0.22	0.11	−0.12
(3) Years education			—	0.16	−0.06	0.16	0.29*
(4) MMSE				—	−0.01	0.25	0.32*
(5) Interference (0 = control, 1 = spatial tapping)					—	−0.64***	−0.58***
(6) Mean span – Forward						—	0.68***
(7) Mean span – Backward							—

NB: All correlations based upon $N = 49$. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 4 | Linear regression models predicting the forward and backward spatial working memory recall of younger and older adults.

		<i>B</i>	<i>SE B</i>	β	<i>t(p)</i>
Younger					
Forward	Age	−0.04	0.03	−0.20	−1.36 (0.18)
	Sex	−0.21	0.30	−0.09	−0.71 (0.48)
	Years education	−0.05	0.05	−0.13	−0.90 (0.37)
	Interference	−1.14	0.29	−0.49	−4.01 (<0.001)
Backward	Age	−0.06	0.03	−0.34	−2.32 (0.025)
	Sex	−0.06	0.28	−0.02	−0.20 (0.84)
	Years education	0.04	0.05	0.13	0.89 (0.34)
	Interference	−1.19	0.26	−0.54	−4.50 (<0.001)
Older					
Forward	Age	0.00	0.01	−0.004	−0.03 (0.97)
	Sex	−0.03	0.17	−0.02	−0.15 (0.88)
	Years education	0.02	0.03	0.08	0.71 (0.48)
	MMSE	0.12	0.07	0.23	1.88 (0.067)
	Interference	−0.97	0.17	−0.64	−5.65 (<0.001)
Backward	Age	−0.03	0.01	−0.23	−2.07 (0.044)
	Sex	−0.34	0.15	−0.24	−2.32 (0.025)
	Years education	0.06	0.03	0.21	2.09 (0.043)
	MMSE	0.09	0.06	0.17	1.53 (0.134)
	Interference	−0.91	0.15	−0.64	−6.14 (<0.001)

research has suggested that active tasks, which involve transforming, manipulating, or integrating information (Cornoldi and Vecchi, 2003) are more sensitive to aging (Vecchi and Cornoldi, 1999; Vecchi et al., 2005; Cansino et al., 2013). Indeed, in a meta-analysis, Bopp and Verhaeghen (2005) demonstrated progressively larger age-related deficits

depending on the extent of active processing required in a task, from simple (forward) storage span, to backward span, and finally to processing intensive working memory tasks such as sentence span. An interesting finding resulting from the initial ANOVA was that there was no differential effect of aging upon task performance in the forward and

backward recall conditions. This supports the results of Hester et al. (2004), who found no interaction between recall type and age in performance of the WMS-III Spatial Span subtest. Additionally, the findings of Vandierendonck et al. (2004) were supported, as they observed no effect of recall type in performance of the Corsi in younger adults. In the present study, the lack of reliable main effect may have been due to the significant maintenance period, which meant that, even in forward recall, the material required active rehearsal to avoid decay prior to the recall stage (Cornoldi and Vecchi, 2003; Baddeley, 2007; Logie, 2011; Vandierendonck, 2016).

However, further linear regression analyses within each age group indicated that the backward recall task may indeed require more active processing than the forward task. Within both age groups, linear regression analyses showed that age significantly predicted backward spatial span, but not forward span. This supports the idea that working memory is vulnerable to decline from early in the adult lifespan, given that age was also predictive of performance within those aged only 18–40 years (Park et al., 2002; Logie and Maylor, 2009; Johnson et al., 2010). Theoretically, the difference between the two tasks was the requirement to draw upon the central executive to re-order the sequences. Thus, specifically central executive functioning may be at least partly responsible for the age-related decline in spatial working memory from early adulthood. Cornoldi and Mammarella (2008) recently argued that forward and backward spatial span differ regarding the underlying resources. They analyzed the effects of recall type across two spatial ability groups, which were categorized as low and high, on the basis of a spatial processing (mental rotation) task. Performance in forward recall was found to be better than in backward recall, but only in the low spatial ability participants, indicating that backward recall does involve additional processing. This additional processing could potentially be more complex spatial processing, given the distinction in these participant groups on the basis of spatial ability. However, this could also be domain-general processing, which their spatial processing task likely also has in common with backward spatial span.

The central executive has been assumed to underlie, at least in part, the processing difference between forward and backward recall. However, particularly as a large amount of variance remained unexplained in the linear regression models, it is important to consider other potential mechanisms underlying the age-related decline in spatial-sequential working memory. Also, Belleville et al. (1998) directly investigated the possibility that central executive manipulation of the contents of working memory may underlie age-related deficits in capacity, and found no evidence for a general central executive deficit in aging. One candidate mechanism is processing speed (Phillips and Hamilton, 2001). Articulatory suppression typically does not affect spatial working memory span (e.g., Vandierendonck et al., 2004), supporting the claim that the task does not rely on verbal working memory. However, Smyth and Scholey (1996) observed that articulation rate reliably predicts spatial working memory span, with or without sequential order, and concluded

that the likely source of this shared variance is cognitive processing speed. Certainly, one influential theory of cognitive aging is that processing speed underlies most of the variance in cognitive functioning in older age (Salthouse, 1996), and this has been shown to be important specifically in visuo-spatial cognition in older adults (Brown et al., 2012; Guest et al., 2015). In the context of the present task, processing speed could be crucial to task performance during sequence encoding, manipulation (in the backward recall), rehearsal, and also in the recall stage. Clearly, there are numerous opportunities in the task for slowed processing to affect performance.

In addition to the effect of interference (discussed below), there were two other significant predictors of backward spatial span in older adults; sex and years of education. Although not specifically expected to predict performance in older people, previous research has identified that males are superior to females specifically within an active visuo-spatial working memory task (Vecchi and Girelli, 1998; see also Kaufman, 2007). Cansino et al. (2013) recently used verbal and visuo-spatial *n-back* tasks of different levels of demand (1- or 2-back) to assess the age-related decline in working memory in 1,500 participants across the adult lifespan. Not only did they show that age effects begin as early as in the third decade of life, but the effects begin earliest for more demanding working memory tasks, and the decline begins earlier in women than men for visuo-spatial working memory. Thus, although it is recommended to interpret the present evidence with caution, particularly as the spatial interference condition had slightly more older males than older females, it is interesting that the finding does relate to existing evidence. A further possible source of the age-related decline in spatial-sequential working memory is strategy, when considering the predictive power of years of education in backward recall in older adults (Orsini et al., 1987). This is increasingly being addressed in the cognitive aging literature, in the context of compensation and lifestyle factors that are being taken into account in current influential perspectives (Bailey et al., 2008; Reuter-Lorenz and Park, 2014). As noted earlier, Fiore et al. (2012) suggested that older adults may not use active spatial rehearsal to the same extent as younger adults. Although the present interference effects suggests that, on average, older adults were using active rehearsal, given the overall age-related deficit that was observed, it would be useful to establish whether or not older adults can benefit from strategy training in a spatial-sequential task.

Functional Architecture of Visuo-Spatial Working Memory

Clear, specific interference effects were observed in both younger and older adults when they were performing the spatial interference task, but not the visual interference task. This provides further evidence that spatial span relies upon spatial processing, and therefore the active spatial rehearsal mechanism of the visuo-spatial sketch pad, but not the visual storage resource within working memory (i.e., visual cache; Logie, 2011; see also Mammarella et al., 2013). Importantly, as both age groups

exhibited this effect, the evidence suggests that both age groups typically use the most relevant form of working memory rehearsal when performing the task. This is in line with Jenkins et al. (2000), who administered a visuo-spatial interference task in conjunction with a spatial working memory task and found the same interference effects in both younger and older adults. However, the present results develop upon this previous evidence by having incorporated more specific visual and spatial interference tasks, as opposed to one more general visuo-spatial interference task (see also Jenkins et al., 1999). The evidence also suggests that spatial span does not rely upon the operation of the passive visual store in working memory (Logie, 2011; Mammarella et al., 2013).

In terms of the use of visual and spatial working memory across the two age groups, then, younger and older adults have been shown to use the same spatial strategy which may be assumed to be the most effective one for task completion. Both age groups were shown not to rely upon visual working memory and this is beneficial to overall performance on the task. This is because the retention of a visual image may have allowed for successful recall of the appropriate block locations of a given sequence, but it would not have been conducive to recalling the sequential order. Interestingly, however, there was a potential effect of DVN specifically in the backward recall task, although the interaction between recall type and interference was not reliable. St Clair-Thompson and Allen (2013) investigated the effects of recall type and visual interference (DVN) on digit recall, and found that DVN presented during recall (but not encoding) disrupted specifically backward recall. They argued that visual imagery is a strategy more likely to be used during backward recall. It is therefore possible that the potential effect of DVN on backward recall in the current study is indicative of visual processing being used to some extent to aid with the more challenging backward recall. That is, some participants may have tried to rely upon the visual image of the layout, while manipulating the sequential order, at least in some of the more demanding levels (Vandierendonck et al., 2004). This possibility would be a useful avenue for future research, for understanding the potentially greater involvement of visual store of the visuo-spatial sketchpad in spatial working memory with backward recall, or other more demanding conditions that push working memory beyond capacity limits (Logie, 2011).

In future investigations of the effect of active processing in spatial-sequential working memory, it would be useful to take into account a number of methodological factors that could influence the findings. The present study used the standard spatial span approach by asking participants first to complete the forward task version, followed by the backward recall task (e.g., Park et al., 2002). Rowe et al. (2008) argued that younger adults' boost in performance under ascending compared with descending sequence presentation formats indicates an important role for practice in spatial span. As younger adults typically begin the task well below their capacity, they gain more practice at the smaller sequence lengths, thus, younger

adults' performance may be overestimated relative to that of older adults, who receive less practice. In the present context, this suggests not only that it would be useful to assess the effect of recall type with a counterbalanced administration procedure, but also that it would be interesting to observe the potential effects of controlling for the extent of task practice. Another issue raised by Wilde and Strauss (2002), is that the standard spatial span task presents the same sequences, but in reverse order, for the backward recall task version. Although rather unlikely, it is possible that participants may store memory traces of the stimuli, particularly if the first task version does not progress very far, which can often be the case in older adults. Thus, administering entirely new sequences in the backward span task would be useful in order to control for this potential issue.

CONCLUSION

This research has shown that spatial-sequential working memory is subject to age-related decline. Although a subject of debate, backward recall, which is assumed to require more active, central executive processing, does appear to be more sensitive to aging in spatial working memory, in the context that age significantly predicted backward recall performance within both the younger and older adult groups. Additionally, the functional architecture of the visuo-spatial sketchpad was shown not to be affected by age when performing forward and backward spatial-sequential working memory tasks, at least in the present conditions. Both age groups were shown to rely upon the most appropriate, specialized processing for task completion (active spatial rehearsal), as a spatial interference task exhibited specific interference effects in both age groups, across both the passive and active recall task versions. While younger and older adults appear to engage in active spatial rehearsal during a spatial-sequential span task, backward spatial span may indeed offer a more sensitive measure of spatial working memory performance across the adult lifespan.

AUTHOR CONTRIBUTIONS

LB is responsible for designing this research, acquiring, analyzing, and interpreting the data, drafting the manuscript, and is accountable for all aspects of the work.

ACKNOWLEDGMENTS

This research was carried out while completing a Ph.D. at Glasgow Caledonian University, under the supervision of Dr Douglas Forbes and Dr Jean McConnell. I also gratefully acknowledge Dr Elaine Niven (University of Dundee), Dr Simon Hunter (University of Strathclyde), and the two reviewers, for their useful comments on earlier versions of this manuscript.

REFERENCES

- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2
- Baddeley, A. D. (2007). *Working Memory, Thought and Action*. Oxford: Oxford University Press.
- Baddeley, A. D. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422
- Baddeley, A. D., and Hitch, G. J. (1974). Working memory. *Psychol. Learn. Motiv.* 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1
- Bailey, H., Dunlosky, J., and Hertzog, C. (2008). Does differential strategy use account for age-related deficits in working-memory performance? *Psychol. Aging* 24, 82–92. doi: 10.1037/a0014078
- Belleville, S., Rouleau, N., and Caza, N. (1998). Effect of normal aging on the manipulation of information in working memory. *Mem. Cogn.* 26, 572–583. doi: 10.3758/BF03201163
- Berch, D. B., Krikorian, R., and Huha, E. M. (1998). The Corsi block-tapping task: methodological and theoretical considerations. *Brain Cogn.* 38, 317–338. doi: 10.1006/brcg.1998.1039
- Bopp, K. L., and Verhaeghen, P. (2005). Aging and verbal memory span: a meta-analysis. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, 223–233. doi: 10.1093/geronb/60.5.P223
- Borst, G., Ganis, G., Thompson, W. L., and Kosslyn, S. M. (2012). Representations in mental imagery and working memory: evidence from different types of visual masks. *Mem. Cogn.* 40, 204–217. doi: 10.3758/s13421-011-0143-7
- Braver, T. S., and West, R. L. (2008). “Working memory, executive control, and aging,” in *The Handbook of Aging and Cognition*, 3rd Edn, eds F. I. M. Craik and T. A. Salthouse (New York, NY: Psychology Press), 311–372.
- Brown, L. A., Brockmole, J. R., Gow, A. J., and Deary, I. J. (2012). Processing speed and visuospatial executive function predict visual working memory ability in older adults. *Exp. Aging Res.* 38, 1–19. doi: 10.1080/0361073X.2012.636722
- Cansino, S., Hernández-Ramos, E., Estrada-Manilla, C., Torres-Trejo, F., Martínez-Galindo, J. G., Ayala-Hernández, M., et al. (2013). The decline of verbal and visuospatial working memory across the adult life span. *Age* 35, 2283–2302. doi: 10.1007/s11357-013-9531-1
- Cornoldi, C., and Mammarella, I. C. (2008). A comparison of backward and forward spatial spans. *Q. J. Exp. Psychol.* 61, 674–682. doi: 10.1080/17470210701774200
- Cornoldi, C., and Vecchi, T. (2003). *Visuo-Spatial Working Memory and Individual Differences*. Hove: Psychology Press.
- Corsi, P. M. (1972). *Human Memory and the Medial Temporal Region of the Brain*. Ph.D thesis, McGill University, Montreal, QC.
- Darling, S., Della Sala, S., and Logie, R. H. (2009). Dissociation between appearance and location within visuo-spatial working memory. *Q. J. Exp. Psychol.* 62, 417–425. doi: 10.1080/17470210802321984
- De Renzi, E., and Nichelli, P. (1975). Verbal and non-verbal short-term memory impairment following hemispheric damage. *Cortex* 11, 341–354. doi: 10.1016/S0010-9452(75)80026-8
- Dean, G. M., Dewhurst, S. A., Morris, P. E., and Whittaker, A. (2005). Selective interference with the use of visual images in the symbolic distance paradigm. *J. Exp. Psychol. Learn.* 31, 1043–1068. doi: 10.1037/0278-7393.31.5.1043
- Dean, G. M., Dewhurst, S. A., and Whittaker, A. (2008). Dynamic visual noise interferes with storage in visual working memory. *Exp. Psychol.* 55, 283–289. doi: 10.1027/1618-3169.55.4.283
- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., and Wilson, L. (1999). Pattern span: a tool for unwinding visuo-spatial memory. *Neuropsychologia* 37, 1189–1199. doi: 10.1016/S0028-3932(98)00159-6
- Farmer, E. W., Berman, J. V. F., and Fletcher, Y. L. (1986). Evidence for a visuospatial scratch-pad in working memory. *Q. J. Exp. Psychol.* 38A, 675–688. doi: 10.1080/14640748608401620
- Fiore, F., Borella, E., Mammarella, I. C., and De Beni, R. (2012). Age differences in verbal and visuo-spatial working memory updating: evidence from analysis of serial position curves. *Memory* 20, 14–27. doi: 10.1080/09658211.2011.628320
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). “Mini-Mental State”: a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6
- Guest, D., Howard, C., Brown, L. A., and Gleeson, H. (2015). Aging and the rate of visual information processing. *J. Vis.* 15:10. doi: 10.1167/15.14.10
- Hamilton, C. J., Coates, R. O., and Heffernan, T. (2003). What develops in visuo-spatial working memory development? *Eur. J. Cogn. Psychol.* 15, 43–69. doi: 10.1080/09541440303597
- Hester, R. L., Kinsella, G. J., and Ong, B. (2004). Effect of age on forward and backward span tasks. *J. Int. Neuropsychol. Soc.* 10, 475–481. doi: 10.1017/S1355617704104037
- Jenkins, L., Myerson, J., Hale, S., and Fry, A. F. (1999). Individual and developmental differences in working memory across the life span. *Psychon. Bull. Rev.* 6, 28–40. doi: 10.3758/BF03210810
- Jenkins, L., Myerson, J., Joerding, J. A., and Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychol. Aging* 15, 157–175. doi: 10.1037/0882-7974.15.1.157
- Johnson, W., Logie, R. H., and Brockmole, J. R. (2010). Working memory tasks differ in factor structure across age cohorts: implications for dedifferentiation. *Intelligence* 38, 513–528. doi: 10.1016/j.intell.2010.06.005
- Kaufman, S. B. (2007). Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity? *Intelligence* 35, 211–223. doi: 10.1016/j.intell.2006.07.009
- Logie, R. H. (1995). *Visuo-Spatial Working Memory*. Hove: Lawrence Erlbaum Associates.
- Logie, R. H. (2003). Spatial and visual working memory: a mental workspace. *Psychol. Learn. Motiv.* 42, 37–78. doi: 10.1016/S0079-7421(03)01002-8
- Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Curr. Dir. Psychol. Sci.* 20, 240–245. doi: 10.1177/0963721411415340
- Logie, R. H., and Maylor, E. A. (2009). An internet study of prospective memory across adulthood. *Psychol. Aging* 24, 767–774. doi: 10.1037/a0015479
- Mammarella, I. C., Borella, E., Pastore, M., and Pazzaglia, F. (2013). The structure of visuospatial memory in adulthood. *Learn. Individ. Differ.* 25, 99–110. doi: 10.1016/j.lindif.2013.01.014
- McConnell, J., and Quinn, J. G. (2000). Interference in visual working memory. *Q. J. Exp. Psychol.* 53A, 53–67. doi: 10.1080/713755873
- McConnell, J., and Quinn, J. G. (2004). Complexity factors in visuo-spatial working memory. *Memory* 12, 338–350. doi: 10.1080/09658210344000035
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *Br. Med. Bull.* 27, 272–277.
- Moffat, S. D., Zonderman, A. B., and Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiol. Aging* 22, 787–796. doi: 10.1016/S0197-4580(01)00251-2
- Myerson, J., Emery, L., White, D. A., and Hale, S. (2003). Effects of age, domain, and processing demands on memory span: evidence for differential decline. *Aging Neuropsychol. Cogn.* 10, 20–27. doi: 10.1076/anec.10.1.20.13454
- Orsini, A., Grossi, D., Capitani, E., Laiacina, M., Papagno, C., and Vallar, G. (1987). Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *Ital. J. Neurol. Sci.* 8, 537–548. doi: 10.1007/BF02333660
- Park, C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., and Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychol. Aging* 17, 299–320. doi: 10.1037/0882-7974.17.2.299
- Phillips, L. H., and Hamilton, C. (2001). “The working memory model in adult aging research,” in *Working Memory in Perspective*, ed. J. Andrade (Hove: Psychology Press), 101–125.
- Quinn, J. G. (1994). Towards clarification of spatial processing. *Q. J. Exp. Psychol.* 47A, 465–480. doi: 10.1080/14640749408401120
- Quinn, J. G., and McConnell, J. (1996a). Irrelevant pictures in visual working memory. *Q. J. Exp. Psychol.* 49A, 200–215. doi: 10.1080/713755613
- Quinn, J. G., and McConnell, J. (1996b). Indications of the functional distinction between the components of visual working memory. *Psychol. Beitr.* 38, 355–367.
- Reuter-Lorenz, P. A., and Park, D. C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychol. Rev.* 24, 355–370. doi: 10.1007/s11065-014-9270-9
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., et al. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive ageing. *J. Int. Neuropsychol. Soc.* 4, 474–490. doi: 10.1017/S1355617798455073
- Rowe, G., Hasher, L., and Turcotte, J. (2008). Age differences in visuospatial working memory. *Psychol. Aging* 23, 79–84. doi: 10.1037/0882-7974.23.1.79

- Rudkin, S. J., Pearson, D. G., and Logie, R. H. (2007). Executive processes in visual and spatial working memory tasks. *Q. J. Exp. Psychol.* 60, 79–100. doi: 10.1080/17470210600587976
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403
- Smyth, M. M., and Pendleton, L. R. (1989). Working memory for movements. *Q. J. Exp. Psychol.* 41A, 235–250. doi: 10.1080/14640748908402363
- Smyth, M. M., and Scholey, K. A. (1994). Interference in immediate spatial memory. *Mem. Cogn.* 22, 1–13. doi: 10.3758/BF03202756
- Smyth, M. M., and Scholey, K. A. (1996). The relationship between articulation time and memory performance in verbal and visuospatial tasks. *Br. J. Psychol.* 87, 179–191. doi: 10.1111/j.2044-8295.1996.tb02584.x
- St Clair-Thompson, H. L., and Allen, R. J. (2013). Are forward and backward recall the same? A dual-task study of digit recall. *Mem. Cogn.* 41, 519–532. doi: 10.3758/s13421-012-0277-2
- Thompson, J. M., Hamilton, C. J., Gray, J. M., Quinn, J. G., Mackin, P., Young, A. H., et al. (2006). Executive and visuospatial sketchpad resources in euthymic bipolar disorder: implications for visuospatial working memory architecture. *Memory* 14, 437–451. doi: 10.1080/09658210500464293
- Vandierendonck, A. (2016). Modality independence of order coding in working memory: evidence from cross-modal order interference at recall. *Q. J. Exp. Psychol.* 69, 161–179. doi: 10.1080/17470218.2015.1032987
- Vandierendonck, A., Kemps, E., Fastame, M. C., and Szmalec, A. (2004). Working memory components of the Corsi blocks task. *Br. J. Psychol.* 95, 57–79. doi: 10.1348/000712604322779460
- Vecchi, T., and Cornoldi, C. (1999). Passive storage and active manipulation in visuo-spatial working memory: further evidence from the study of age differences. *Eur. J. Cogn. Psychol.* 11, 391–406. doi: 10.1080/713752324
- Vecchi, T., and Girelli, L. (1998). Gender differences in visuo-spatial processing: the importance of distinguishing between passive storage and active manipulation. *Acta Psychol.* 99, 1–16. doi: 10.1016/S0001-6918(97)00052-8
- Vecchi, T., and Richardson, J. T. E. (2001). Measures of visuospatial short-term memory: the knox cube imitation test and the corsi blocks test compared. *Brain Cogn.* 46, 255–316. doi: 10.1016/S0278-2626(01)80086-5
- Vecchi, T., Richardson, J. T. E., and Cavallini, E. (2005). Passive storage versus active processing in working memory: evidence from age-related variations in performance. *Eur. J. Cogn. Psychol.* 17, 521–539. doi: 10.1080/09541440440000140
- Wechsler, D. (1997). *Wechsler Memory Scale: Administration and Scoring Manual*, 3rd Edn. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1998). *Wechsler Memory Scale: Administration and Scoring Manual*, 3rd Edn. London: Psychological Corporation.
- Wilde, N., and Strauss, E. (2002). Functional equivalence of WAIS-III/WMS-III digit and spatial span under forward and backward recall conditions. *Clin. Neuropsychol.* 16, 322–330. doi: 10.1076/clin.16.3.322.13858
- Wilde, N. J., Strauss, E., and Tulskey, D. S. (2004). Memory span on the wechsler scales. *J. Clin. Exp. Neuropsychol.* 26, 539–549. doi: 10.1080/13803390490496605

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Brown. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Differences in Verbal and Visuospatial Forward and Backward Order Recall: A Review of the Literature

Enrica Donolato^{1*}, David Giofrè² and Irene C. Mammarella¹

¹ Department of Developmental and Social Psychology, University of Padova, Padova, Italy, ² Department of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK

OPEN ACCESS

Edited by:

Snehlata Jaswal,
LM Thapar School of Management,
India

Reviewed by:

Emily M. Elliott,
Louisiana State University, USA
Rick Thomas,
Georgia Institute of Technology, USA

*Correspondence:

Enrica Donolato
enrica.donolato@gmail.com

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 12 July 2016

Accepted: 12 April 2017

Published: 04 May 2017

Citation:

Donolato E, Giofrè D and
Mammarella IC (2017) Differences in
Verbal and Visuospatial Forward and
Backward Order Recall: A Review of
the Literature. *Front. Psychol.* 8:663.
doi: 10.3389/fpsyg.2017.00663

How sequential, verbal and visuospatial stimuli are encoded and stored in memory is not clear in cognitive psychology. Studies with order recall tasks, such as the digit, and Corsi span, indicate that order of presentation is a crucial element for verbal memory, but not for visuospatial memory. This seems to be due to the different effects of forward and backward recall in verbal and visuospatial tasks. In verbal span tasks, performance is worse when recalling things in backward sequence rather than the original forward sequence. In contrast, when it comes to visuospatial tasks, performance is not always worse for a modified backward sequence. However, worse performance in backward visuospatial recall is evident in individuals with weak visuospatial abilities; such individuals perform worse in the backward version of visuospatial tasks than in the forward version. The main aim of the present review is to summarize findings on order recall in verbal and visuospatial materials by considering both cognitive and neural correlates. The results of this review will be considered in the light of the current models of WM, and will be used to make recommendations for future studies.

Keywords: order recall, verbal working memory, visuospatial working memory, short-term memory, neural correlates

INTRODUCTION

The ability to process serially ordered information is fundamental to many aspects of our lives, including spelling and orientation to a new environment. However, the cognitive mechanisms underlying encoding and recall of verbal and visuospatial sequences are still not fully understood.

One of the processes involved in serial recall is short-term memory (STM), which allows individuals to hold a small amount of information for a short period of time. Verbal STM is generally tested with the digit span task (DST) that involves recalling sequences of digits, while the ability to retrieve visuospatial information is typically tested with the Corsi span task (CST) that involves recalling sequences of blocks (Berch et al., 1998). In both verbal and visuospatial span tasks, participants may be asked to recall the information in either forward or backward order. In the DST, performance is usually worse in the backward version of the task (Baddeley, 1986; Li and Lewandowsky, 1995), while recall of the forward and backward versions of the CST is much the same for most subjects (Wilde and Strauss, 2002; Cornoldi and Mammarella, 2008). Although, these results give the impression that forward and backward verbal and visuospatial span tasks likely measure different constructs, experimental and neural correlated findings regarding

serial recall tasks are not consistent, suggesting a need for further research. It is also important to carefully review the research and literature on this subject to-date. Hence, here we will summarize the findings from studies of forward and backward recall in verbal and visuospatial domains as it relates to both cognitive and neural correlates.

Selection of Studies

We conducted a literature search via PsychINFO, Web of Science and Google Scholar electronic databases. We used the following search keywords: serial/order recall, forward/backward span/recall, DST and CST, verbal/visuospatial, STM/WM. We searched for these terms in titles, in abstracts, and in the keyword lists themselves. Titles and abstracts were screened for appropriateness and independently reviewed for relevance. Papers published from January 1960 to September 2016 have been considered. 132 manuscripts were originally selected for scrutiny; ultimately, only 54 met our inclusion criteria and were considered in the present review.

Papers were considered for inclusion if they covered: (i) behavioral or neural correlates for forward and/or backward recall in the verbal and/or visuospatial domain; and (ii) the impact of verbal and/or visuospatial memory capacity. Studies focusing purely on theoretical models for memory systems were not considered.

Similarities and Differences in Order Recall in Verbal or Visuospatial Domains

Studies in order recall on verbal and visuospatial domains to-date have used different methodologies and had different aims.

Forward and backward serial position curves have been analyzed by considering primacy and recency effects in verbal WM (**Table 1**, verbal WM). As for backward recall, findings have shown a qualitative change in serial position curves characterized by an increased recency effect and a decreased primacy effect (Li and Lewandowsky, 1993, 1995; Hulme et al., 1997, see also Penney, 1989 for a review). Likewise, in forward and backward visuospatial tasks, both primacy and recency effects also occur (Farrand and Jones, 1996; Farrand et al., 2001).

Several studies adopted the dual task paradigm when examining verbal tasks. In the literature, the most commonly used examples of dual tasks are *articulatory suppression*, and *irrelevant speech* (i.e., concurrent irrelevant sounds). These dual tasks effectively impair memory performance (see Baddeley and Hitch, 1974). In two separate studies, Bireta et al. (2010) and Guérard et al. (2012) explored these effects using similar methods and procedures, but came to different conclusions. In both studies, all the above-mentioned effects were confirmed on forward recall, but only Guérard et al. (2012) found the dual tasks had an impact on backward recall as well. Ritchie et al. (2015) conducted a meta-analysis of 16 experiments focusing on secondary tasks. Irrelevant speech was observed to have a weak effect on task performance. In contrast, articulatory suppression had very large effects, and seemed to disrupt recall for both first and late responses (i.e., primacy and recency), in forward and backward recall alike.

As for the visuospatial domain, studies tend to take one of two approaches: comparing extreme groups, or analyzing clinical populations (**Table 1**, visuospatial WM). For example, adults with high spatial abilities demonstrated very similar performance in the forward and backward versions of the CST (Wilde and Strauss, 2002), whereas participants with low spatial abilities demonstrated lower performances in backward recall (Cornoldi and Mammarella, 2008). This finding was also confirmed in children with non-verbal learning disability who had severe problems in the spatial domain (Cornoldi et al., 2003; Mammarella and Cornoldi, 2005; Garcia et al., 2014).

Other research directly compared the verbal and visuospatial domains (**Table 1**, verbal and visuospatial WM) by using the dual task paradigm. Research has shown that a serial secondary task (i.e., spatial tapping) interferes with recall of spatial information (Jones et al., 1995; Vandierendonck et al., 2004). Further, the presence of a verbal secondary task affects both verbal and visuospatial recall when the secondary task requires the manipulation of ordered information. However, when the secondary task requires the manipulation of unordered materials, visuospatial performance is not affected (Depoorter and Vandierendonck, 2009). This result was used by the researchers as evidence of the existence of a cross modal interference. Others argued that the effect seen in Depoorter and Vandierendonck's research was probably due to the specific manipulation used in the study (Logie et al., 2016). However, in further research a different manipulation was employed, and the results confirmed the existence of cross modal interference between verbal and spatial recall performance (Vandierendonck, 2016). This effect was also confirmed using visuospatial materials only (i.e., the CST), showing that both forward and backward recall were affected by the presence of a verbal secondary tasks (Higo et al., 2014).

Neural Correlates

Although, there is not a general consensus on a specific WM model, neuroscience studies can help to shed further light on the effect of order recall in verbal and visuospatial domains. In fact, event-related potentials (ERPs) and functional Magnetic Resonance Imaging (fMRI) have also been used to investigate neural correlates of forward and backward recall in the verbal and visuospatial domains (see **Table 2**).

ERPs were used in the backward DST under two conditions; digits were aurally presented and were followed by a second set that either corresponded to the reverse order (correct condition), or by a second set in which an incorrect digit was included in the list (incorrect condition) (Lefebvre et al., 2005; Marchand et al., 2006). The findings showed a positive P2 and P3 in the correct condition, and conversely showed a prolonged positive slow wave for the incorrect condition; this suggests that the two conditions are associated with different patterns of activation. Another study compared forward and backward recall, showing the presence of high negative correlations between P3 latency and the DST (Walhovd and Fjell, 2002). Finally, research has shown that the amplitudes of the P3a and P3b ERPs are reduced during backward recall in verbal but not in visuospatial tasks (Nulsen et al., 2010). This finding, suggesting a different pattern

TABLE 1 | Studies measuring forward and/or backward recall in verbal and/or visuospatial working memory.

Study	N	Age (Mean)	Gender (M/F)	Tasks	Behavioral measures	Main results	Order recall	
							F	B
STUDIES ON VERBAL WORKING MEMORY								
Li and Lewandowsky, 1993	N = ranged from 12 to 30 (Study 1–4)	Undergraduate students	NA	List of words (Study 1, 2 and 4) or letters (Study 3) and distractor item(s) that involved adding 3 to a random number and saying the result aloud.	Correct responses (correct item reported in its initial serial position regardless of the number and the nature of any preceding responses).	Forward but not backward recall was disrupted by an intralist distraction task. This effect was greatest for the first and last items and tended to decrease for the last items.	✓	✓
Li and Lewandowsky, 1995	N = ranged from 10 to 15	Undergraduate students	NA	Consonants (Study 1 and 3) or words (Study 2) and distractor item(s) (Study 1 and 2).	Correct responses (correct item reported in its initial serial position regardless of the number and the nature of any preceding responses).	Differences between forward and backward recall are interpreted as supporting the existence of two different retrieval processes for forward and backward order.	✓	✓
Hulme et al., 1997	N = 28 (Study 4)	Undergraduate students	Study 4: 24 F, 4 M	List of short high-frequency words vs. short low-frequency words.	Proportion of correct recall.	Word frequency had an increasing effect across serial positions in forward but not in backward recall.	✓	✓
Bireta et al., 2010	N = 20	Undergraduate students	NA	Short vs. long words (Study 1); letters with no noise vs. irrelevant speech (Study 2); similar- vs. dissimilar-sounding letters (Study 3); words associated with quiet vs. concurrent AS (Study 4).	Proportion of correct responses.	Word length effect, irrelevant speech effect, acoustic confusion effect, and concurrent articulation effect were less strong in the backward than in the forward recall, or even disappeared.	✓	✓
Guérard et al., 2012	N = ranged from 20 to 40	Undergraduate students	NA	Phonologically similar vs. monosyllabic words (Study 1A and 1B) associated with quiet vs. irrelevant speech (Study 2); open vs. closed pool of words associated with quiet vs. concurrent AS (Study 3); short vs. long word condition (Study 4); monosyllabic vs. polysyllabic words (Study 5).	Proportion of correct responses.	A replication of the Bireta et al. study (2010) revealed word length effect, irrelevant speech effect, acoustic confusion effect and concurrent articulation effect in both forward and backward recall.	✓	✓

(Continued)

STUDIES ON VISUOSPATIAL WORKING MEMORY

Study	N	Age (Mean)	Gender (M/F)	Tasks	Behavioral measures	Main results	Order recall
STUDIES ON VISUOSPATIAL WORKING MEMORY							
Vandierendonck et al., 2004	N = 25 (Study 1 and 2)	Undergraduate students	NA	Computerized CST with randomly selected block sequences of different lengths; Single vs. dual-task situations (AS, matrix-tapping, random-interval generation and fixed-interval generation) (Study 1 and 2).	Proportion of correct locations or not (Study 1 and 2).	The random interval generation task condition coincided with worse performance in backward than in forward order recall. Concurrently performing the matrix-tapping task impaired memory performance in both recall directions. Articulatory suppression did not affect memory performance on forward recall, but on backward recall, it impaired memory for longer sequences on backward recall. The standardized difference between forward and backward spatial span was 0.11 [−1.83, 2.04].	✓ ✓
Cornoldi and Mammarella, 2008	N = 40	19.05–19.23 years (HSA-LSA group)	HSA: 2M, 18 F LSA: 1M, 19 F	CST using ascending vs. descending format and different sequence lengths.	Order scores in terms of mean percentages of correct responses.	Participants with low-spatial-abilities (LSA) showed specific impairment in the backward recall direction; no differences between forward and backward recall in the high-spatial-ability (HAS) group.	✓ ✓
Garcia et al., 2014	N = 45	8–10 years	NLD: 10 M, 5 F RD: 6 M, 9 F TD: 8M, 7 F	CST, Color task, Color-location binding task.	Percentages of correct responses.	The non-verbal learning disability (NLD) group performed worse than the typical development (TD) group in the CST, and particularly in backward recall. The standardized difference between forward and backward spatial span was −0.50 [−1.21, 0.24].	✓ ✓
Higo et al., 2014	N ranged from 21 to 24 (Study 1 and 2)	22.7 years–22.9 years (Study 1 and 2)	Study 1: 11 M, 13 F; Study 2: 12 M, 9 F	CST associated with secondary tasks (control, serial AS or ST) (Study 2); item sequence length manipulation (Study 1).	Correct trials, position error, order errors (Study 1 and 2).	Spatial tapping (ST) interfered equally with both forward and backward recall, while serial articulatory suppression (AS) increased the number of order errors in backward recall, suggesting a stronger order representation for backward recall. The standardized difference between forward and backward spatial span was 0.40 [−0.01, 0.82].	✓ ✓

(Continued)

TABLE 1 | Continued

Study	N	Age (Mean)	Gender (M/F)	Tasks	Behavioral measures	Main results	Order recall	
							F	B
STUDIES ON VERBAL AND VISUOSPATIAL WORKING MEMORY								
Jones et al., 1995	N ranged from 20 to 36 (Study 1, 2 and 3)	Undergraduate students	NA	Sequences of dots (Study 1); Verbal (consonant) vs. spatial memory (dots) task associated with verbal or spatial interference (spatial manual interference, AS, irrelevant speech, respectively for Study 2- 3-4) and steady state vs. changing state condition.	Serial order errors (Study 1, 2).	Similar position curves, more errors with increasing list length, and more errors when retention interval increased, in forward verbal and spatial span tasks.	✓	
Farrand and Jones, 1996	N = 18 (Study 1, 2, 3 and 4)	NA	NA	Spatial task (dot), Visual-Verbal Task (sequence of consonants) and Auditory-Verbal Task (sequence of recorded consonants) (Study 1); Only Visual-Verbal Task (sequence of consonants) (Study 2 and 3); Only Spatial task (dot) (Study 4).	Serial position curves, serial order errors, city block errors (Study 4).	Performance on forward and backward recall for verbal and spatial stimuli was similar when only the order of information had to be recalled (Study 1). When both item and order information had to be recalled, performance in backward recall was significantly worse than in forward recall, suggesting a similarity between verbal and spatial recall.	✓	✓
Wilde and Strauss, 2002	N = 44	M = 37.1 years (SD = 13.9)	26 M, 18 F	Digit span and Spatial span tests.	Digit span and Spatial span scores.	Better performance in forward than in backward recall in the Digit span but not in the Spatial span tasks.	✓	✓
Mammarella and Cornoldi, 2005	N ranged from 36 to 46 (Study 1 and 2)	3rd-, 4th-, and 5th-grade schoolchildren (Study 1); nd- and 5th- grade schoolchildren (Study 2)	Study 1: NLD: 12 M, 6 F and CG: 12 M, 6 F Study 2: NA (Study 1-2)	Digit span and CST test.	Correct responses (Study 1 and 2).	In two studies the NLD group showed significantly lower scores in the backward than in the forward CST. In the Digit span task, both groups performed worse in backward recall. In the first study, the standardized difference between forward and backward spatial span was −0.48 [−1.13, 0.24]. In the second study, the standardized difference was 0.17 [−0.61, 0.93] and 0.18 [−0.63, 0.98] for children in the second and in the fifth grade respectively.	✓	✓

(Continued)

TABLE 1 | Continued

Study	N	Age (Mean)	Gender (M/F)	Tasks	Behavioral measures	Main results	Order recall	
							F	B
Depoorter and Vanderendonck, 2009	N ranged from 17 to 20 (Study 1, 2, 3, 4)	19.5–24.7 years	Study 1: 6 M, 11 F; Study 2: 5 M, 12 F; Study 3: 1 M, 19 F	Primary visuospatial task (visuospatial squares and lines task) associated with secondary visuospatial item or order task (Study 1); Primary verbal task (letters) associated with secondary verbal item or order task (Study 2); Primary visuospatial and verbal task associated with secondary verbal and visuospatial task (Study 3 and 4).	Correct responses (Study 1 and 2) and ratio measures (Study 2).	Worse performance in primary task recall when both the primary and the secondary task involved order, irrespective of whether input was verbal or visuospatial.	✓	
Vanderendonck, 2016	N ranged from 22 to 24 (Study 1 and 2)	18–22/32 years (Study 1 and 2)	Study 1: 2 M, 22 F; Study 2: 20 M, 12 F	Verbal (one-syllable Dutch words) vs. visuospatial (dot task) as primary vs. secondary task with order vs. item recall stimuli (Study 1). Verbal serial primary task associated with secondary visuospatial task in different conditions (perfect embedded trials, primary task recall for all trials and embedded task recall) (Study 2).	Serial recall tasks: proportion of elements recalled in correct relative position (scoring between 0 and 1); item recall tasks were scored 0 or 1 (incorrect vs. correct).	Visuospatial, but not verbal serial recall was more impaired when the embedded task was an order recall task than when it was an item recall task (Exp. 1), suggesting that order is at least partly coded on a different, modality-independent level.	✓	

Note: AS, articulatory suppression; DWN, dynamic visual noise; HSA, high spatial abilities; LSA, low spatial abilities; NA, not available; NLD, non-verbal learning disability; RD, reading disability; ST, spatial tapping; TD, typical development. Papers in this Table are arranged by year of publication.

TABLE 2 | Studies using fMRI or ERP measuring forward and/or backward recall in verbal and/or visuospatial working memory.

Study	N	Age (Mean)	Gender (M/F)	Tasks	Measures		Main results	Order recall	
					fMRI	ERP		F	B
STUDIES ON VERBAL WORKING MEMORY									
Walhovd and Fjell, 2002	N = 71 divided in 5 age group	21.8–94.7 years	35 F, 36 M	Digit span task.	✓		Significant negative correlations were present between P3 latency and digit span scores. Moreover, a multiple regression analysis controlling for age yielded a significant relationship between P3 latency and both the digit total and forward span tasks.	✓	✓
Lefebvre et al., 2005	N = 20	22.2 years	11 F, 9 M	Digit span backwards task consisting of different set sizes was aurally presented, followed by a second set that either corresponded to the reverse order of the first set (correct condition) or had one digit in the sequence replaced by an incorrect digit (incorrect condition). The participants have to indicate whether the trial was correct or incorrect.	✓		Positive component (P2/P3) was elicited in the correct condition, whereas a comparatively robust and prolonged positive slow wave was elicited in the incorrect condition. Moreover, the second wave and the difference between its amplitude in the incorrect and correct conditions decreased as working memory load increased.		✓
Sun et al., 2005	N = 22 (Young adults: 11; Older adults: 11)	Young adults: 19–27 years; Older adults: 60–72 years	Young adults: 5 F, 5 M; Older adults: 5 F, 5M	Digit recall task.	✓		In young adults, the left occipital visual region and the left PFC were more activated during backward recall task than forward recall task, while the GFI was more activated in forward than backward direction. In older adults, more regions and especially the frontal cortex showed a greater activation in backward than forward digit recall, while the GFI was more activated during backward recall.	✓	✓
Majerus et al., 2006	N = 18	20–28 years	NA	Sequential presentation of four words with probe recognition for word identity or word order.	✓		An increase in lateral inferior parietal activation was observed when comparing order and item conditions. However, activation within the left IPS was observed in both the order and item tasks, while the left IPS was activated during order than item conditions.	✓	
Marchand et al., 2006	N = 20	22.2 years	11 F, 9 M	Digit span backwards task consisting of different set sizes was aurally presented, followed by a second set that either corresponded to the reverse order of the first set (correct condition) or had one digit in the sequence replaced by an incorrect digit (incorrect condition). The participants have to indicate whether the trial was correct or incorrect.	✓		A prolonged positive slow wave (PSW) peaking between 450 and 750 ms. was elicited to incorrect condition trials.		✓

(Continued)

TABLE 2 | Continued

Study	N	Age (Mean)	Gender (M/F)	Tasks	Measures		Main results	Order recall	
					fMRI	ERP		F	B
Marshuetz et al., 2006	N = 12	20–27 years	6 F, 6 M	Five consonants were presented, followed by two probe letters from the memory set. The probe letters appeared either in the original order or were transposed and were separated by zero to three positions in the memory set. Participants had to indicate whether the items were in the correct order.	✓		The activation in the left parietal cortex showed a systematic decrease, with increasing inter-item probe distance.	✓	
Majerus et al., 2008	N = 22	HPB: 19.42 years; LPB: 19.62 years	HPB: 9 F, 2 M; LPB: 7 F, 4 M	Visual and sequential presentation of four words, followed by a maintenance phase, and a retrieval phase in which probe words were matching or not the target information.	✓		High proficiency bilinguals presented a greater activation in the lateral orbito-frontal and superior frontal gyri associated with order condition. High and low proficiency groups showed similar activation of fronto-parietal and fronto-temporal networks for order and item conditions. However, for order conditions, the high proficiency group showed a larger activation in the left orbito-frontal cortex during encoding, while the low proficiency group showed a greater activation in the bilateral superior frontal cortex during retrieval.	✓	
Rossi et al., 2013	N = 22	10 years	11 F, 11 M	Digit span task.	✓		The gray matter volume of the left AIC and the frontal and prefrontal regions were negatively correlated with backward memory span recall. On the other hand, the gray matter volume of the inferior parietal lobe was negatively correlated with the forward memory span. Moreover, smaller gray matter volumes in the left AIC region, the inferior and the superior frontal gyrus, were present in children who performed better on the backward span task.	✓	✓
Manan et al., 2014	N = 54	G1: 20–29 years; G2: 30–39 years; G3: 40–49 years; G4: 50–65 years	10 F, 44 M	Backward repeat task (BRT) and forward repeat task (FRT) where verbal stimuli consisted of a series of natural recorder speech words.	✓		Backward recall is associated with the activation of parietal and frontal regions compared to forward recall. Moreover, brain changes related to hemispheric laterality were detected in relation to age.	✓	✓
Yang et al., 2015	N = 72	7.13–16.84 years	32 F, 40 M	Digit span test.	✓		Backward digit span recall was associated with the activation of the right DLPFC, the FEF, the frontal operculum cortex, anterior insular cortex, and the dACC, while forward digit recall was associated with precuneus and lateral visual areas. Finally, the dACC region was positively related to backward span task but negatively related to forward span task.	✓	✓

(Continued)

TABLE 2 | Continued

Study	N	Age (Mean)	Gender (M/F)	Tasks	Measures		Main results	Order recall	
					fMRI	ERP		F	B
STUDIES ON VISUOSPATIAL WORKING MEMORY									
Croizé et al., 2004	N = 8	28 years	4 F, 6 M	In the task, subjects decided whether two dots were symmetrical or not with respect to a central fixation cross. Dots were presented either simultaneously ("immediate condition") or successively ("memorization" of the position of a flashed dot and a "delayed comparison" with a second displayed dot).	✓		The M4 activation observed by using magnetoencephalography (MEG) in the "memorization" condition was localized in the right premotor region, in direct spatial correspondence with the precise activation focus revealed by fMRI, along the premotor sulcus and at the intersection with the superior frontal sulcus.	✓	
Toepper et al., 2010	N = 19	26.1 years	10 F, 9M	Modified version of the CST. Participants were asked to reproduce each sequence by deciding which between two alternative response options.	✓		When compared to the baseline activity, the right hippocampus showed more activation during the CST condition. Moreover, increased activity within parietal, frontal and occipital areas was observed during the encoding stage.	✓	
Toepper et al., 2014	N = 36 (young: 18; old: 18)	young: 26.8 years, old: 71.7 years	young: 9 F, 9 M; old: 9 F, 9 M	Modified version of the CST where participants were asked to reproduce each sequence by pressing buttons in the correct order.	✓		When older individuals were compared to younger participants, lower right-dorsolateral prefrontal activation and lower functional connectivity between this area and the bilateral orbitofrontal cortex was observed. Moreover, the older high-performance group showed higher right dorsolateral and anterior prefrontal cortex activation and higher functional connectivity between these regions when compared to the older low-performance group.	✓	
STUDIES ON VERBAL AND VISUOSPATIAL WORKING MEMORY									
Nulsen et al., 2010	N = 24	23 years	18 F, 6 M	Digit and spatial span tasks. Participants had to decide whether a sequence of stimuli shown during the test phase matched or dis not match the initial studied items.		✓	A large effect on reversal recall was observed specifically on the digits task, with a smaller effect seen in the spatial span tasks. Moreover, the mean amplitude of the P3a component was larger to a statistically significant degree in the forward condition for the digit task, compared to the backward condition. The standardized difference between forward and backward spatial span was 0.26 [−0.15, 0.66].	✓	✓
Chen et al., 2011	N = 12	19–24 years	Study: 7 F, 5 M	For processing components: lexical decision and symmetry decision tasks; for storage components: letters and location tasks, in which subjects attempted to retain a sequence of 4 items for later recall.	✓		For both verbal and spatial versions of complex working memory span tasks, an increased activity in lateral prefrontal, anterior cingulate, and parietal cortices during the Encoding, Maintenance, and Coordination phase of task performance was seen. In addition, overlapping activity in the anterior prefrontal and medial temporal lobe regions was associated with both verbal and spatial recall from working memory.	✓	

(Continued)

TABLE 2 | Continued

Study	N	Age (Mean)	Gender (M/F)	Tasks	Measures		Main results	Order recall	
					fMRI	ERP		F	B
Nagel et al., 2013	N = 67	10–16 years	32 F, 35 M	Block-design, verbal and spatial WM 2-back task that differed only by way of task instruction.	✓		While verbal WM was associated with significant left hemispheric lateralization in the frontal and parietal lobes, the spatial WM involved right hemisphere lateralization in the frontal and temporal areas. Moreover, increased adolescent age was associated with less activity in the default mode brain network during the verbal WM, and associated with greater activity in the task-positive posterior parietal cortex during spatial WM task.		✓

Note: ACC, anterior cingulate cortex; AIC, anterior insular cortex; CWMs, complex working memory span; dACC, dorsal anterior cingulate gyrus; DLPFC, dorsolateral prefrontal cortex; FEF, frontal eye fields; GF, inferior frontal gyrus; HPB, High proficiency bilinguals; IPS, left intraparietal sulcus; LPB, low proficiency bilinguals; NA, not available; PFC, prefrontal cortex; PPC, posterior parietal cortex. Papers in this Table are arranged by year of publication.

of activation between verbal and spatial backward spans, seems to indicate a reduction of attentional resources in the verbal backward span (see Nulsen et al., 2010).

As for fMRI studies, in the verbal domain, two studies compared the recognition of ordered and item information. In the ordered condition, results show a greater bilateral activation in the intraparietal sulcus (IPS) and in the premotor frontal areas (Henson et al., 2000; Marshuetz et al., 2000), supporting the idea that ordered material requires more attentional resources. However, a study by Majerus et al. (2006) failed to find a consistent differential activation in the left IPS, indicating that the difference between order and item condition is related to a specific network that links the left and right IPS with the right dorsal premotor cortex and the superior cerebellum. Interestingly, bilingual individuals with a high level of proficiency in both languages demonstrated greater activation in the lateral orbito-frontal region and in the superior frontal gyri associated with the updating of ordered information (Majerus et al., 2008), confirming the presence of different patterns of activation for order and item encoding.

Studies comparing different patterns of activation in forward and backward recall are mainly focused on verbal material. Research has shown the involvement of different neural correlates in forward and backward recall of digits (Manan et al., 2014). Another study showed that backward digit recall was associated with a higher activation of the left occipital visual region and the left prefrontal cortex (PFC) in young adults (Sun et al., 2005), supporting the idea of the involvement of visuospatial processing during backward verbal tasks (e.g., Larrabee and Kane, 1986; Hoshi et al., 2000). Moreover, young adults showed a greater activation in the inferior frontal gyrus in both forward and backward recall. The activation was associated with a limited overlap, providing evidence in favor of a distinction between forward and backward recall activation patterns (Sun et al., 2005). Furthermore, the central executive seems to be highly taxed during backward digit recall (Carlesimo et al., 1994). These results are in line with previous findings suggesting that the backward DST was associated with the activation of regions that are also involved in tasks requiring high cognitive control. Such activated regions include the right dorsolateral PFC, the frontal eye field, the frontal operculum cortex, the anterior insular cortex and the dorsal anterior cingulate cortex (dACC) (Yang et al., 2015). Intriguingly, activation of the dACC region was positively related to the backward span task but negatively related to the forward one (Yang et al., 2015). Finally, results with child subjects revealed the presence of distinct negative correlations between the forward/backward DST and the gray matter volume of some brain areas, such as the left AIC region, the inferior frontal gyrus and the superior frontal gyrus (Rossi et al., 2013).

While research on verbal material is plentiful in the literature, fMRI studies on the visuospatial domain are mainly focused on the forward span. In a study, when participants were asked to decide whether two dots were symmetrical or not, the results revealed that in the memorization condition, where participants had to judge whether the symmetry of the second dot's position related to the memorized position of the first dot, the right

premotor region was activated (Croizé et al., 2004). Two other studies involved a modified version of the CST, and showed the involvement of the hippocampus in the encoding of spatial locations (Toepper et al., 2010). In addition, age effects were observed in the right-dorsolateral prefrontal cortex, which was found to be less activated in the older group compared to the younger one (Toepper et al., 2014).

Finally, in two studies comparing verbal and visuospatial domains the results seemed to favor a distinction between the two domains (Chein et al., 2011; Nagel et al., 2013). For example, Nagel et al. (2013) considered the difference between the verbal and visuospatial domains from a developmental point of view, suggesting that increased adolescent age was associated with less activity in the default mode brain network (i.e., a brain network more commonly active at rest and deactivated during task) during a verbal WM task. In contrast, increased adolescent age was associated with greater activity in the posterior parietal cortex during a spatial WM task.

Implications for the Working Memory Models

The presence of different results for order recall in verbal and visuospatial domains is considered as evidence in support of several existing theoretical models of WM. Baddeley's WM model postulates the existence of two domain-specific subsystems involved in the storage of verbal and visuospatial information: the *phonological loop* and the *visuospatial sketchpad*, respectively. These two components are linked with the *central executive system* that integrates and manipulates information (e.g., Baddeley, 1986). In this model, the *phonological loop* explains several phenomena affecting serial recall in verbal STM, such as the influence of word length, articulatory suppression, phonological similarity and item similarity. The decline in performance in the backward span is also interpreted in relation to the central executive's taxed resources (Baddeley, 1986). However, a limit of this model is that it fails to explain the results observed in visuospatial tasks and lacks a clear distinction between recalling sequential ordered information, and recalling unordered information.

Alternative models of WM propose a modality-independent view, with no distinction between verbal and visuospatial input. This approach is supported by the similar serial position curves detected in the verbal and visuospatial domains, and the shared memory resources for maintaining information in a given order (Engle, 1996; Cowan, 1999, 2005; Oberauer, 2009). It has also been suggested that the difference between verbal and visuospatial span tasks in forward and backward directions is associated with dissimilar retrieval demands: while participants use blocks to give their answers in the CST, in the DST the digits are not presented during the retrieval phase. Thus, verbal tasks would seem to require the recall of both items and order information, while visuospatial tasks would only require the latter (Farrand and Jones, 1996). Similarities between serial order and position effects in the verbal and spatial domains can be explained by assuming that order is treated similarly across different domains (Smyth, 1996). A comparable view is based

on the assumption that there is a modality-independent process for serial order retention, and a domain-specific process for item retention (Depoorter and Vandierendonck, 2009).

Another hypothesis postulates that verbal and visuospatial forward serial recall measures the "passive" STM component, while backward recall involves executive control resources (Carlesimo et al., 1994; Hester et al., 2004). Developmental studies, combined with research in which clinical samples were considered, have helped to clarify this hypothesis. For example, a greater involvement of executive control in backward serial recall has been demonstrated in typically-developing children (Alloway et al., 2009), and in children with ADHD or learning disabilities (Cornoldi et al., 2013a,b; Giofrè et al., 2016), but not in adults (Rosen and Engle, 1997).

Concerning the visuospatial domain, a model has been proposed (Logie, 1995; Darling et al., 2007) that distinguishes between the *visual cache*, linked with the temporary storage of static visual information, and the *inner scribe*, involved in the dynamic processing of sequences of movement. According to this model, the maintenance of sequential information is crucial in spatial processes. This is in contrast with other models which are based on the assumption that visuospatial processes tend to lose sequential information in favor of simultaneously presented information (Paivio, 1971).

Finally, a model distinguishing between a visual component and two spatial subcomponents involving spatial-sequential and spatial-simultaneous processes has been proposed (Lecerf and De Ribaupierre, 2005; Mammarella et al., 2008, 2013). This is supported by findings in different groups of children with developmental disorders (Mammarella et al., 2003, 2006; Lanfranchi et al., 2015), and in healthy adults (Mammarella et al., 2013). In this view, Mammarella and Cornoldi (2005) suggested that differences in forward and backward spatial recall are due not only to the involvement of the executive control, but also to involvement of spatial-sequential and spatial-simultaneous processing. Research has also shown that backward recall in the CST requires less executive control and more spatial processing, supporting the idea that backward recall involves a modality-independent order coding system (Higo et al., 2014). This hypothesis is supported by evidence suggesting that the backward CST involves both visuospatial processing and executive control (Vandierendonck et al., 2004; Vandierendonck, 2016).

We decided to investigate this result further by analyzing papers including both versions of the CST. Nine papers included in this review deal with both versions of the spatial span. Among these studies, two did not report means or effect sizes, making it impossible to calculate effect sizes (Farrand and Jones, 1996; Farrand et al., 2001); one reported data from a clinical sample (Wilde and Strauss, 2002), and one compared participants with high vs. low spatial abilities (Cornoldi and Mammarella, 2008). However, five studies reported descriptive statistics (i.e., Vandierendonck et al., 2004; Mammarella and Cornoldi, 2005; Nulsen et al., 2010; Garcia et al., 2014; Higo et al., 2014) and seven effect sizes were extracted from these studies (see also **Tables 1, 2**). When considering these effects together, and assuming random effects, the overall effect is $d_{unb} = 0.039 [-0.20, 0.28]$. This finding seems to indicate the difference between forward

and backward spatial span is very small and not statistically significant.

CONCLUSIONS

In the present paper, we reviewed findings on forward and backward recall in the verbal and visuospatial domains, considering the contribution of experimental and neuroscience studies.

The evidence from the cognitive studies is quite clear. Regarding the verbal domain, the verbal recall task is often characterized by a clear difference between the forward and the backward version of the span, with lower performance in the latter. In the visuospatial domain—at least when typically developing children or healthy adults are considered—it is more difficult to detect differences between recall of the forward and backward versions of the task.

Overall, experimental studies do not provide a clear support for any theoretical model described above. Advances in technical and quantitative methods of neuroscience over the past years have aided and propelled analysis in various fields of psychology. Neuroscientific studies cited in this review have indicated that verbal recall, in the backward order in particular, seems to require greater cognitive resources (Manan et al., 2014). In addition, different brain areas are activated in verbal and visuospatial tasks (Sun et al., 2005; Yang et al., 2015). These findings support modality independent models of WM, and in fact verbal performance and visuospatial performance is always clearly distinguishable.

Unfortunately, to-date no study has compared forward and backward recall in verbal and visuospatial domains in relation to neural correlates. A promising future line of research would involve studies that examine the simultaneous storage of information derived from different modalities. In fact, future efforts should directly compare the neural correlates of forward

and backward recall in verbal and visuospatial domains, and do so within a single study. Furthermore, other techniques should be used in order to collect further evidence. Potentially, future studies could employ other psycho-physiological measures such as eye movements, or neuroimaging techniques such as transcranial magnetic stimulation or magneto-encephalography. These kinds of additional analytical measures could allow researchers to reach clearer results. Moreover, methodologies and the types of tasks used in future studies should be consistent and comparable.

Ultimately, only few developmental studies have been carried out to-date; therefore, how the serial recall of verbal and spatial information develops is not yet completely clear. A deeper understanding of such changes could in turn help in improving our understanding of currently existing theoretical models.

Despite some shortcomings, the findings collectively gathered in this review are both comprehensive and beneficial to those currently researching in this field. The take home messages from these reviews are as follows: (1) verbal and spatial WM modalities seem to be distinct; (2) there is overwhelming evidence for a distinction between the forward and backward digit span; and (3) there is no clear evidence for a distinction between forward and backward spatial span.

AUTHOR CONTRIBUTIONS

ED, ICM, and DG contributed to the study concept and design. ED, ICM, and DG wrote the paper. All authors approved the final version of the manuscript for submission.

ACKNOWLEDGMENTS

We would like to thank David Recine for his editing and helpful comments.

REFERENCES

- Alloway, T. P., Gathercole, S. E., Kirkwood, H., and Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Dev.* 80, 606–621. doi: 10.1111/j.1467-8624.2009.01282.x
- Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon Press.
- Baddeley, A. D., and Hitch, G. J. (1974). "Working memory," in *The Psychology of Learning and Motivation Vol. 8*, ed G. H. Bower (New York, NY: Academic Press), 47–89.
- Berch, D. B., Krikorian, R., and Huha, E. M. (1998). The Corsi-block tapping task: methodological and theoretical considerations. *Brain Cogn.* 38, 317–338. doi: 10.1006/brcg.1998.1039
- Bireta, T. J., Fry, S. E., Jalbert, A., Neath, I., Surprenant, A. M., Tehan, G., et al. (2010). Backward recall and benchmark effects of working memory. *Mem. Cognit.* 38, 279–291. doi: 10.3758/MC.38.3.279
- Carlesimo, G. A., Fadda, L., Lorusso, S., and Caltagirone, C. (1994). Verbal and spatial memory spans in Alzheimer's and multi-infarct dementia. *Acta Neurol. Scand.* 89, 132–138. doi: 10.1111/j.1600-0404.1994.tb01648.x
- Chein, J. M., Moore, A. B., and Conway, A. R. (2011). Domain-general mechanisms of complex working memory span. *Neuroimage* 54, 550–559. doi: 10.1016/j.neuroimage.2010.07.067
- Cornoldi, C., Giofrè, D., Calgaro, G., and Stupiggia, C. (2013a). Attentional WM is not necessarily specifically related with fluid intelligence: the case of smart children with ADHD symptoms. *Psychol. Res.* 77, 508–515. doi: 10.1007/s00426-012-0446-8
- Cornoldi, C., and Mammarella, I. C. (2008). A comparison of backward and forward spatial spans. *Q. J. Exp. Psychol.* 61, 674–682. doi: 10.1080/17470210701774200
- Cornoldi, C., Orsini, A., Cianci, L., Giofrè, D., and Pezzuti, L. (2013b). Intelligence and working memory control: evidence from the WISC-IV administration to Italian children. *Learn. Individ. Differ.* 26, 9–14. doi: 10.1016/j.lindif.2013.04.005
- Cornoldi, C., Venneri, A., Marconato, F., Molin, A., and Montinari, C. (2003). A rapid screening measure for the identification of visuospatial learning disability in schools. *J. Learn. Disabil.* 36, 299–306. doi: 10.1177/00222194030360040201
- Cowan, N. (1999). "An embedded-processes model of working memory," in *Models of Working Memory: Mechanisms of Active Maintenance and Control*, eds A. Miyake and P. Shah (Cambridge: Cambridge University Press), 62–101.
- Cowan, N. (2005). *Working Memory Capacity*. Hove: Psychology Press.
- Croizé, A. C., Ragot, R., Garnero, L., Ducorps, A., Pelegrini-Issac, M., Dauchot, K., et al. (2004). Dynamics of parietofrontal networks underlying visuospatial short-term memory encoding. *Neuroimage* 23, 787–799.

- Darling, S., Della Sala, S., and Logie, R. H. (2007). Behavioural evidence for separating components within visuo-spatial working memory. *Cogn. Process.* 8, 175–181. doi: 10.1007/s10339-007-0171-1
- Depoorter, A., and Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *Q. J. Exp. Psychol.* 62, 531–549. doi: 10.1080/17470210801995002
- Engle, R. W. (1996). “Working memory and retrieval: an inhibition-resource approach,” in *Working Memory and Human Cognition*, ed J. T. E. Richardson (Oxford: Oxford University Press), 89–119.
- Farrand, P., and Jones, D. M. (1996). Direction of report in spatial and verbal serial short-term memory. *Q. J. Exp. Psychol. A* 49, 140–158.
- Farrand, P., Parmentier, F. B. R., and Jones, D. M. (2001). Temporal-spatial memory: retrieval of spatial information does not reduce recency. *Acta Psychol.* 106, 285–301. doi: 10.1016/S0001-6918(00)00054-8
- Garcia, R. B., Mammarella, I. C., Tripodi, D., and Cornoldi, C. (2014). Visuospatial working memory for locations, colours, and binding in typically developing children and in children with dyslexia and non-verbal learning disability. *Br. J. Dev. Psychol.* 32, 17–33. doi: 10.1111/bjdp.12019
- Giofrè, D., Stoppa, E., Ferioli, P., Pezzuti, L., and Cornoldi, C. (2016). Forward and backward digit span difficulties in children with specific learning disorder. *J. Clin. Exp. Neuropsychol.* 38, 478–486. doi: 10.1080/13803395.2015.1125454
- Guérard, K., Saint-Aubin, J., Burns, S. C., and Chamberland, C. (2012). Revisiting backward recall and benchmark memory effect: a reply to Bireta et al. (2010). *Mem. Cognit.* 40, 388–407. doi: 10.3758/s13421-011-0156-2
- Henson, R. N., Burgess, N., and Frith, C. D. (2000). Recoding, storage, rehearsals, and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia* 38, 426–440. doi: 10.1016/S0028-3932(99)00098-6
- Hester, R. L., Kinsella, G. J., and Ong, B. (2004). Effect of age on forward and backward span tasks. *J. Int. Neuropsychol. Soc.* 10, 475–481. doi: 10.1017/S1355617704104037
- Higo, K., Minamoto, T., Ikeda, T., and Osaka, M. (2014). Robust order representation is required for backward recall in the Corsi blocks task. *Front. Psychol.* 5:1285. doi: 10.3389/fpsyg.2014.01285
- Hoshi, Y., Oda, I., Wada, Y., Ito, Y., Yamashita, Y., Oda, M., et al. (2000). Visuospatial imagery is a fruitful strategy for the digit span backward task: a study with near-infrared optical tomography. *Cogn. Brain Res.* 9, 339–342. doi: 10.1016/S0926-6410(00)00006-9
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., and Stuart, G. (1997). Word frequency effects on short-term memory tasks: evidence for a reintegration process in immediate serial recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 1217–1232. doi: 10.1037/0278-7393.23.5.1217
- Jones, D., Farrand, P., Stuart, G., and Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 21:1008. doi: 10.1037/0278-7393.21.4.1008
- Lanfranchi, S., De Mori, L., Mammarella, I. C., Carretti, B., and Vianello, R. (2015). Spatial-sequential and spatial-simultaneous working memory in individuals with Williams syndrome. *Am. J. Intellect. Dev. Disabil.* 120, 193–202. doi: 10.1352/1944-7558-120.3.193
- Larrabee, G. J., and Kane, R. L. (1986). Reversed digit repetition involves visual and verbal processes. *Int. J. Neurosci.* 30, 11–15. doi: 10.3109/00207458608985649
- Lecerf, T., and De Ribaupierre, A. (2005). Recognition in a visuospatial memory task: the effect of presentation. *Eur. J. Cogn. Psychol.* 17, 47–75. doi: 10.1080/09541440340000420
- Lefebvre, C. D., Marchand, Y., Eskes, G. A., and Connolly, J. F. (2005). Assessment of working memory abilities using an event-related brain potential (ERP)-compatible digit span backward task. *Clin. Neurophysiol.* 116, 1665–1680. doi: 10.1016/j.clinph.2005.03.015
- Li, S.-C., and Lewandowsky, S. (1993). Intralist distracters and recall direction: constraints on models of memory for serial order. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 895–908. doi: 10.1037/0278-7393.19.4.895
- Li, S.-C., and Lewandowsky, S. (1995). Forward and backward recall: different retrieval processes. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 837–847. doi: 10.1037/0278-7393.21.4.837
- Logie, R. H. (1995). *Visuo-Spatial Working Memory*. Hove: Erlbaum.
- Logie, R. H., Saito, S., Morita, A., Varma, S., and Norris, D. (2016). Recalling visual serial order for verbal sequences. *Mem. Cognit.* 44, 590–607.
- Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A. L., Martinez, T., Schmidt, C., et al. (2008). Neural networks for short-term memory for order differentiate high and low proficiency bilinguals. *Neuroimage* 42, 1698–1713. doi: 10.1016/j.neuroimage.2008.06.003
- Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., et al. (2006). The left intraparietal sulcus and verbal short-term memory: focus of attention or serial order? *Neuroimage* 32, 880–891. doi: 10.1016/j.neuroimage.2006.03.048
- Mammarella, I. C., Borella, E., Pastore, M., and Pazzaglia, F. (2013). The structure of visuospatial memory in adulthood. *Learn. Individ. Differ.* 25, 99–110. doi: 10.1016/j.lindif.2013.01.014
- Mammarella, I. C., and Cornoldi, C. (2005). Sequence and space: the critical role of a backward spatial span in the working memory deficit of visuospatial learning disabled children. *Cogn. Neuropsychol.* 22, 1055–1068. doi: 10.1080/02643290442000509
- Mammarella, I. C., Cornoldi, C., Pazzaglia, F., Toso, C., Grimaldi, M., and Vio, C. (2006). Evidence for a double dissociation between spatial-simultaneous and spatial-sequential working memory in visuospatial (nonverbal) learning disabled children. *Brain Cogn.* 62, 58–67. doi: 10.1016/j.bandc.2006.03.007
- Mammarella, I. C., Pazzaglia, F., and Cornoldi, C. (2008). Evidence for different components in children's visuospatial working memory. *Br. J. Dev. Psychol.* 26, 337–355. doi: 10.1348/026151007X236061
- Mammarella, N., Cornoldi, C., and Donadello, E. (2003). Visual but not spatial working memory deficit in children with spina bifida. *Brain Cogn.* 53, 311–314. doi: 10.1016/S0278-2626(03)00132-5
- Manan, H. A., Franz, E. A., Yusoff, A. N., and Mukari, S. Z. M. S. (2014). Age-related brain activation during forward and backward verbal memory tasks. *Neurol. Psychiatry Brain Res.* 20, 76–86. doi: 10.1016/j.npbr.2014.08.001
- Marchand, Y., Lefebvre, C. D., and Connolly, J. F. (2006). Correlating digit span performance and event-related potentials to assess working memory. *Int. J. Psychophysiol.* 62, 280–289. doi: 10.1016/j.ijpsycho.2006.05.007
- Marshuetz, C., Reuter-Lorenz, P. A., Smith, E. E., Jonides, J., and Noll, D. C. (2006). Working memory for order and the parietal cortex: an event-related functional magnetic resonance imaging study. *Neuroscience* 139, 311–316.
- Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., and Chenevert, T. L. (2000). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *J. Cogn. Neurosci.* 12, 130–144. doi: 10.1162/08989290051137459
- Nagel, B. J., Herting, M. M., Maxwell, E. C., Bruno, R., and Fair, D. (2013). Hemispheric lateralization of verbal and spatial working memory during adolescence. *Brain Cogn.* 82, 58–68. doi: 10.1016/j.bandc.2013.02.007
- Nulsen, C. E., Fox, A. M., and Hammond, G. R. (2010). Processes contributing to forward and backward span: an ERP investigation. *Neuroreport* 21, 298–302. doi: 10.1097/WNR.0b013e32833730f0
- Oberauer, K. (2009). “Design for a working memory,” in *The Psychology of Learning and Motivation Vol. 51B*, ed H. Ross (London: Academic Press), 45–100.
- Paivio, A. (1971). *Imagery and Verbal Processes*. New York, NY: Holt, Reinhart & Winston.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Mem. Cognit.* 17, 398–422. doi: 10.3758/BF03202613
- Ritchie, G., Tolan, G. A., Tehan, G., Goh, H. E., Guérard, K., and Saint-Aubin, J. (2015). Phonological effects in forward and backward serial recall: qualitative and quantitative differences. *Can. J. Exp. Psychol.* 69:95. doi: 10.1037/cep0000028
- Rosen, V. M., and Engle, R. W. (1997). Forward and backward serial recall. *Intelligence* 25, 37–47. doi: 10.1016/S0160-2896(97)90006-4
- Rossi, S., Lubin, A., Simon, G., Lanoë, C., Poirel, N., Cachia, A., et al. (2013). Structural brain correlates of executive engagement in working memory: children's inter-individual differences are reflected in the anterior insular cortex. *Neuropsychologia* 51, 1145–1150. doi: 10.1016/j.neuropsychologia.2013.03.011
- Smyth, M. M. (1996). Serial order in spatial immediate memory. *Q. J. Exp. Psychol. A* 49, 159–177. doi: 10.1080/713755615
- Sun, X., Zhang, X., Chen, X., Zhang, P., Bao, M., Zhang, D., et al. (2005). Age-dependent brain activation during forward and backward digit recall revealed by fMRI. *Neuroimage* 26, 36–47. doi: 10.1016/j.neuroimage.2005.01.022

- Toepper, M., Markowitsch, H. J., Gebhardt, H., Beblo, T., Bauer, E., Woermann, F. G., et al. (2014). The impact of age on prefrontal cortex integrity during spatial working memory retrieval. *Neuropsychologia* 59, 157–168. doi: 10.1016/j.neuropsychologia.2014.04.020
- Toepper, M., Markowitsch, H. J., Gebhardt, H., Beblo, T., Thomas, C., Gallhofer, B., et al. (2010). Hippocampal involvement in working memory encoding of changing locations: an fMRI study. *Brain Res.* 1354, 91–99. doi: 10.1016/j.brainres.2010.07.065
- Vandierendonck, A. (2016). Modality independence of order coding in working memory: evidence from cross-modal order interference at recall. *Q. J. Exp. Psychol.* 69, 161–179. doi: 10.1080/17470218.2015.1032987
- Vandierendonck, A., Kemps, E., Fastame, M. C., and Szmalec, A. (2004). Working memory components of the Corsi block task. *Br. J. Psychol.* 95, 57–79. doi: 10.1348/000712604322779460
- Walhovd, K. B., and Fjell, A. M. (2002). The relationship between P3 and neuropsychological function in an adult life span sample. *Biol. Psychol.* 62, 65–87. doi: 10.1016/S0301-0511(02)00093-5
- Wilde, N., and Strauss, E. (2002). Functional equivalence of WAIS-III/WMS-III digit and spatial span under forward and backward recall conditions. *Clin. Neuropsychol.* 16, 322–330. doi: 10.1076/clin.16.3.322.13858
- Yang, Z., Jutagir, D. R., Koyama, M. S., Craddock, R. C., Yan, C. G., Shehzad, Z., et al. (2015). Intrinsic brain indices of verbal working memory capacity in children and adolescents. *Dev. Cogn. Neurosci.* 15, 67–82. doi: 10.1016/j.dcn.2015.07.007

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Donolato, Giofrè and Mammarella. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Commentary: Distinct neural mechanisms for remembering when an event occurred

Sarah DuBrow^{1*} and Lila Davachi^{1,2}

¹ Department of Psychology, New York University, New York, NY, USA, ² Center for Neural Science, New York University, New York, NY, USA

Keywords: temporal order memory, temporal context, sequence memory, recency discrimination, hippocampus

A commentary on

Distinct neural mechanisms for remembering when an event occurred

by Jenkins, L. J., and Ranganath, C. (2016). *Hippocampus* 26, 554–559. doi: 10.1002/hipo.22571

OPEN ACCESS

Edited by:

Snehlata Jaswal,
L M Thapar School of Management,
India

Reviewed by:

Marc Howard,
Boston University, USA
Christopher MacDonald,
Massachusetts Institute of
Technology, USA
Andy C. H. Lee,
University of Toronto Scarborough,
Canada
Gabriel Radvansky,
University of Notre Dame, USA

*Correspondence:

Sarah DuBrow
sdubrow@nyu.edu

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 10 June 2016

Accepted: 30 January 2017

Published: 21 February 2017

Citation:

DuBrow S and Davachi L (2017)
Commentary: Distinct neural
mechanisms for remembering when
an event occurred.
Front. Psychol. 8:189.
doi: 10.3389/fpsyg.2017.00189

Memory for the relative order of events is a critical feature of episodic remembering that is thought to rely on hippocampal processes (Eichenbaum, 2013; Davachi and DuBrow, 2015). However, there are multiple ways in which the hippocampus may support order memory. Here, we review a recent fMRI paper by Jenkins and Ranganath (2016) investigating two potential memory mechanisms that may support recency discrimination. To briefly summarize, participants were scanned while encoding sequences of object images and were subsequently tested on which of two objects had been presented more recently. The authors examined neural patterns during encoding that predicted later recency judgments and found evidence that item strength and context differentiation support order memory. Our goal here is to provide a theoretical perspective on these and related findings to highlight how numerous mechanisms may support order memory and how fMRI can be leveraged to test competing theories.

Perhaps the most intuitive way to evaluate the order of two items is to compare how strong they are in memory. Since memory strength decays over time, an item's current strength can provide an estimate of how much time passed since it was encountered (Hinrichs, 1970). To determine which of two items occurred more recently, one strategy might be to simply select the one that has the higher activation strength (Hintzman, 2005, c.f. Hintzman, 2010). Jenkins and Ranganath (2016) found evidence in line with a strength-based temporal representation in the prefrontal (PFC) and medial temporal lobe cortices including the perirhinal cortex, which has been consistently implicated in encoding item strength (Aggleton and Brown, 1999; Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007). Specifically, these regions showed greater activation during the initial encoding of items later endorsed as more recent regardless of their true temporal position. While these results are consistent an item-strength comparison account of recency judgments, an alternative retrieval process called scanning could show similar effects at encoding. Backwards scanning models propose that memoranda are sequentially sampled from the end until reaching an item with a sufficient match to one of the recency probes (Hacker, 1980; Howard et al., 2015). Thus, if the more recent item was not encoded strongly enough, it could be bypassed in favor of the stronger, earlier item, consistent with the findings of Jenkins and Ranganath.

Another possibility is that recency judgments could be supported by a comparison of the contexts associated with the objects during encoding. Prominent memory theories propose that items are bound to a temporal context representation that gradually changes over time (Howard and Kahana, 2002; Polyn et al., 2009). Jenkins and Ranganath suggest that this representation may be used to guide recency judgments, presumably by a process that compares the retrieved

contexts of the two items and selects the item whose associated context is most similar to the current state. The more differentiated the two retrieved contexts are, the easier it should be to make the comparison. The authors find evidence for this “context differentiation” account in bilateral hippocampus as well as in regions of the medial and anterior PFC. Specifically, the more dissimilar the fMRI patterns were during the encoding of the two items, the better performance was on later recency discrimination. Assuming that pattern dissimilarity reflects a change in the intervening context above and beyond differences between the items themselves, this suggests that context differentiation leads to better order memory because the items’ contexts are more discriminable at retrieval. Note, however, that neural patterns at retrieval were not examined.

The item strength and context differentiation accounts of order memory are similar in that they are both based on estimating and comparing distances of the two items between encoding and retrieval. However, there are at least two other major classes of theories of temporal representation—those that are based on the absolute time or position at which an event occurred (i.e., location-based) and those that are based on relative time or position (Friedman, 1993). One example of a model that encodes relative position is associative chaining, in which each item in a sequence is directly linked to its neighbors (e.g., Lewandowsky and Murdock, 1989). The temporal context theories described above are actually closely related to associative chaining. However, rather than employing direct item-item links, temporal context theory proposes that neighboring items are linked indirectly through their associated context representation (Howard and Kahana, 2002). These associations allow an item’s retrieved context to elicit retrieval of nearby items that share a similar temporal context. Thus, an alternate account of temporal context in recency judgments might predict that retrieving the context associated with the recency items may lead to the reactivation of the intervening sequence, since the intervening items share context with both recency probes. These sequential associations may in turn provide the relative order information necessary to make accurate recency judgments. Note, a similar retrieval process applied to a location-based temporal representation (e.g., Howard et al., 2015) could also retrieve sequential associations with more absolute temporal precision.

There is evidence supporting this associative account of recency discrimination in episodic memory from both behavioral and fMRI work. Behaviorally, intervening boundaries have been shown to disrupt associative binding (Zwaan and Radvansky, 1998; Ezzyat and Davachi, 2011) and impair order memory (DuBrow and Davachi, 2013; Horner et al., 2016). There is also evidence that, when making recency judgments, the intervening sequence is incidentally reactivated (DuBrow and Davachi, 2013, 2014). Importantly, in this design, hippocampal pattern *similarity* was related to successful recency judgments (DuBrow and Davachi, 2014) in contrast to the hippocampal pattern *dissimilarity* reported by Jenkins and Ranganath. One possibility is that these conflicting results may be due to differences in the processes engaged during encoding. DuBrow

and Davachi promoted the use of associative encoding, which has been shown to influence behavioral and neural order memory effects (Konishi et al., 2006; Jonker and Macleod, 2016). In contrast, the use of a single stimulus category in Jenkins and Ranganath may have promoted a differentiation strategy. Indeed, hippocampal differentiation of items that share similar features or associates has been shown to lead to better memory (LaRocque et al., 2013; Hulbert and Norman, 2015; Schlichting et al., 2015; Favila et al., 2016). Thus, it is not clear to what extent hippocampal patterns in these studies indexed context *per se*, as opposed to processes that either promote maintenance (pattern similarity) or differentiation (pattern dissimilarity).

The study by Jenkins and Ranganath is an important contribution to the literature on temporal memory (for recent reviews, see Howard and Eichenbaum, 2013; Eichenbaum, 2014; Davachi and DuBrow, 2015; Ranganath and Hsieh, 2016). Together with previous data, this work suggests that no singular mechanism supports all order memory, but instead multiple temporal representations and retrieval mechanisms may coexist. This work also highlights the importance of considering how distinct cognitive processes that can be localized to the same brain region may give rise to similar behaviors, in this case successful order memory, despite different underlying mechanisms. Indeed, while a wealth of data implicates the hippocampus in temporal memory, the mechanisms attributed to it have been wide ranging and include each class of theories discussed above—relative distance supported by context differentiation (Manns et al., 2007; Ezzyat and Davachi, 2014; Jenkins and Ranganath, 2016), relative order supported by sequential binding (Tubridy and Davachi, 2011; Schapiro et al., 2012; DuBrow and Davachi, 2014), and location information supported by positional coding (Hsieh et al., 2014; Kalm and Norris, 2014). In the Jenkins and Ranganath study alone, the hippocampus showed both item strength and context differentiation effects at encoding. Another recent study showed that associative and distance-based order judgments engaged the hippocampus equally (Lieberman et al., 2016). Moving forward by using explicit models to compare the predictions of different temporal memory theories will help specify the precise computational role(s) of a given brain region (e.g., Kalm and Norris, 2014). In addition, collecting data during both encoding and retrieval would allow the underlying *representation* of temporal order to be evaluated separately from the decision *process*, and in turn, capture more individual variability in order memory judgments. Manipulating access to temporal information within the same study will also be necessary to determine whether different mechanisms could be employed adaptively depending on available sources of information and current retrieval goals. For example, lengthening the interval between items may flip the relative reliance on associative vs. distance-based information and may be indexed by the influence of pattern similarity vs. dissimilarity, respectively, on accuracy. Ultimately, examining whether and how these processes may tradeoff at different timescales and under various encoding and retrieval conditions will be critical for establishing a comprehensive model of temporal memory.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Aggleton, J. P., and Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behav. Brain Sci.* 22, 425–444. doi: 10.1017/S0140525X99002034
- Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Curr. Opin. Neurobiol.* 16, 693–700. doi: 10.1016/j.conb.2006.10.012
- Davachi, L., and DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. *Trends Cogn. Sci.* 19, 92–99. doi: 10.1016/j.tics.2014.12.004
- Diana, R. A., Yonelinas, A. P., and Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn. Sci.* 11, 379–386. doi: 10.1016/j.tics.2007.08.001
- DuBrow, S., and Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *J. Exp. Psychol. Gen.* 142, 1277–1286. doi: 10.1037/a0034024
- DuBrow, S., and Davachi, L. (2014). Temporal memory is shaped by encoding stability and intervening item reactivation. *J. Neurosci.* 34, 13998–14005. doi: 10.1523/JNEUROSCI.2535-14.2014
- Eichenbaum, H. (2013). Memory on time. *Trends Cogn. Sci.* 17, 81–88. doi: 10.1016/j.tics.2012.12.007
- Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nat. Rev. Neurosci.* 15, 732–744. doi: 10.1038/nrn3827
- Eichenbaum, H., Yonelinas, A., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annu. Review Neurosci.* 30, 123–152. doi: 10.1146/annurev.neuro.30.051606.094328
- Ezzyat, Y., and Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychol. Sci.* 22, 243–252. doi: 10.1177/0956797610393742
- Ezzyat, Y., and Davachi, L. (2014). Similarity Breeds Proximity: Pattern Similarity within and across Contexts Is Related to Later Mnemonic Judgments of Temporal Proximity. *Neuron* 81, 1179–1189. doi: 10.1016/j.neuron.2014.01.042
- Favila, S. E., Chanales, A. J. H., and Kuhl B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nat. Commun.* 7:11066. doi: 10.1038/ncomms11066
- Friedman, W. J. (1993). Memory for the time of past events. *Psychol. Bull.* 113, 44–66. doi: 10.1037/0033-2909.113.1.44
- Hacker, M. J. (1980). Speed and accuracy of recency judgments for events in short-term memory. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 651–675. doi: 10.1037/0278-7393.6.6.651
- Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychol. Rev.* 77, 223–233. doi: 10.1037/h0029101
- Hintzman, D. L. (2005). Memory strength and recency judgments. *Psychon. Bull. Rev.* 12, 858–864. doi: 10.3758/BF03196777
- Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Mem. Cognit.* 38, 102–115. doi: 10.3758/MC.38.1.102
- Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., and Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition* 154, 151–164. doi: 10.1016/j.cognition.2016.05.013
- Howard, M. W., and Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *J. Exp. Psychol. Gen.* 142, 1211–1230. doi: 10.1037/a0033621
- Howard, M. W., and Kahana, M. J. (2002). A distributed representation of temporal context. *J. Math Psychol.* 46, 269–299. doi: 10.1006/jmps.2001.1388
- Howard, M. W., Shankar, K. H., Aue, W. R., and Criss, A. H. (2015). A distributed representation of internal time. *Psychol. Rev.* 122, 24–53. doi: 10.1037/a0037840
- Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., and Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron* 81, 1165–1178. doi: 10.1016/j.neuron.2014.01.015
- Hulbert, J. C., and Norman, K. A. (2015). Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cereb. Cortex* 25, 3994–4008. doi: 10.1093/cercor/bhu284
- Jenkins, L. J., and Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. *Hippocampus* 26, 554–559. doi: 10.1002/hipo.22571
- Jonker, T. R., and Macleod, C. M. (2016). Not all order memory is equal: test demands reveal dissociations in memory for sequence information. *J. Exp. Psychol. Learn. Mem. Cogn.* 43, 177–188. doi: 10.1037/xlm0000301
- Kalm, K., and Norris, D. (2014). The representation of order information in auditory-verbal short-term memory. *J. Neurosci.* 34, 6879–6886. doi: 10.1523/JNEUROSCI.4104-13.2014
- Konishi, S., Asari, T., Jimura, K., Chikazoe, J., and Miyashita, Y. (2006). Activation shift from medial to lateral temporal cortex associated with recency judgements following impoverished encoding. *Cereb. Cortex* 16, 469–474. doi: 10.1093/cercor/bhi126
- LaRocque, K., Smith, M., Carr, V. A., Witthoft, N., Grill-Spector, K., and Wagner, A. D. (2013). Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *J. Neurosci.* 33, 5466–5474. doi: 10.1523/JNEUROSCI.4293-12.2013
- Lewandowsky, S., and Murdock, B. B. (1989). Memory for serial order. *Psychol. Rev.* 96, 25–57. doi: 10.1037/0033-295X.96.1.25
- Lieberman, J. S., Kyle, C. T., Schedlbauer, A., Stokes, J., and Ekstrom, A. D. (2016). A tale of two temporal coding strategies: common and dissociable brain regions involved in recency versus associative temporal order retrieval strategies. *J. Cogn. Neurosci.* doi: 10.1162/jocn_a_01081. [Epub ahead of print].
- Manns, J. R., Howard, M. W., and Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron* 56, 530–540. doi: 10.1016/j.neuron.2007.08.017
- Polyn, S. M., Norman, K. A., and Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychol. Rev.* 116, 129–156. doi: 10.1037/a0014420
- Ranganath, C., and Hsieh, L.-T. (2016). The hippocampus: a special place for time. *Ann. N.Y. Acad. Sci.* 1369, 93–110. doi: 10.1111/nyas.13043
- Schapiro, A., Kustner, L., and Turk-Browne, N. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr. Biol.* 22, 1622–1627. doi: 10.1016/j.cub.2012.06.056
- Schlichting, M. L., Mumford, J. A., and Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* 6:8151. doi: 10.1038/ncomms9151
- Tubridy, S., and Davachi, L. (2011). Medial temporal lobe contributions to episodic sequence encoding. *Cereb. Cortex* 21, 272–280. doi: 10.1093/cercor/bhq092
- Zwaan, R. A., and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol. Bull.* 123, 162–185. doi: 10.1037/0033-2909.123.2.162

FUNDING

This work is supported by the National Institute of Mental Health (RO1-MH074692) and the National Science Foundation (DGE 0813964).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 DuBrow and Davachi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

