Modeling the Contributions of Capacity and Control to Working Memory Development

Evan M. Russek[1], **Cameron Rouse Turner**[1], **Emma McEwen**[2], **Andreea M. Miscov**[2], **Amanda Seed**[2], **Thomas L. Griffiths**[1]

[1]Departments of Computer Science and Psychology, Princeton University, Princeton, NJ, USA

[2]School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, UK

Abstract

Adults are known to have superior working memory to children, but whether this improvement is driven primarily by differences in storage capacity or attentional control is debated. In particular, the understanding of how capacity and control influence the development of working memory is hampered by the fact that most theorizing about the effect of variation in either on behavior has been verbal. To address this, we extended a computational model of working memory to clearly separate the contributions of capacity and control, fitting the model to a recent developmental study. We find that the combined influence of capacity and control on working memory may be more complicated than previously appreciated. In particular, the general pattern of qualitative differences between children and adults could be produced by increasing either capacity or control alone. These results point to a need for additional experimental paradigms to clearly parse the differential impact of working memory components.

Keywords: Working memory; Retro-cue; Attention; Metareasoning; Reinforcement Learning; Development

Introduction

Working memory underpins the ability to retain and update task relevant information and perform mental operations in pursuit of a goal. While there are a number of theoretical models of working memory, there is consensus that it comprises multiple components, broadly involving the storage and processing of information. There is also agreement about several key distinct components of working memory performance (Oberauer et al., 2018). First, working memory is a constrained system of limited capacity, and variation in that capacity is relevant to performance (Cowan, 2012). Second, processes of selective attention can influence recall both at encoding (ignoring irrelevant information), and during retention (selectively retaining information as its relevance becomes clear; Shimi et al., 2014). Working memory undergoes marked developmental improvement throughout childhood. However, the degree to which this is attributable to increases in the different components of working memory, and specifically storage capacity and attentional control, is debated (Shimi & Scerif, 2017; Cowan et al., 2010; Cowan, 2016). Recently, marked differences in visual short-term memory have also been found between species of primates, with a clear phylogenetic signal. Our closest relatives, the apes, performed better than monkey species, which in turn outperformed lemurs (ManyPrimates et al., 2022). Again, the contributions of different aspects of memory to this difference (for example, limits in capacity and control) are unknown.

Empirical research to disentangle the effects of these multiple components of working memory on both developmental and phylogenetic differences must grapple with the complexity of the role played by limits in both capacity and control, or run the risk of misattributing the cause of individual differences in performance. Disentangling the influence of capacity and control on working memory has so far relied largely on intuitions about how variation in either causes changes in behavior. However, such intuitions might be misleading.

In this paper, we investigate the contributions of capacity and control to working memory performance by extending a model of control of working memory (Suchow & Griffiths, 2016) to explore the effects of parametric variation in storage capacity and attentional control. We use this model to simulate the effects of changing capacity and control on recall in a series of retro-cue experiments aimed at investigating how their interplay differs in children and adults (Shimi & Scerif, 2017). Through simulation, we find that the influence of these two components on recall in these experiments remains difficult to identify, offering a potential reinterpretation of how components of working memory change through development.

Background

Capacity and control across development

Working memory is known to be capacity limited (Oberauer & Kliegl, 2006; Cowan, 2012). The extent of such a limitation is typically indexed by requiring an individual to remember an array of visual stimuli and examining how performance decreases as the size of the array (load) increases (Zhang & Luck, 2008). Individuals can compensate for limited capacity by deploying attentional control to select which pieces of information are relevant to encode and represent, and which of these representations to prioritize and actively maintain (Raye et al., 2007; Oberauer, 2019). Attentional control has typically been measured through the use of cues that can convey to the participant which stimulus is more likely to be probed, and thus where they should devote resources. Whereas cues used prior to stimulus presentation can index attentional control in encoding stimuli, cues presented after stimulus presentation (retro-cues) have been used to ascertain the role of attentional control in active maintenance of working memory representations (e.g. Astle et al., 2012; Shimi et al., 2014). On trials in which a retro-cue correctly indicates the probe item, performance is better than on trials without such a cue, creating a memory benefit (Griffin & Nobre, 2003). Furthermore, testing uncued items from the array following a retro-cue incurs performance costs (Astle et al., 2012; van Moorselaar et

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

al., 2015). The size of these effects has been used to index individual variability in the efficacy of control (Shimi & Scerif, 2017; Cowan et al., 2010).

Corresponding paradigms have fueled competing claims as to whether increases in capacity are the main contributor to working memory development (Gathercole et al., 2004; Cowan et al., 2010), or whether increases in attentional control are equally important (Astle et al., 2012; Shimi et al., 2014). Shimi & Scerif (2017) recently aimed to measure adults' and childrens' use of attentional control and how this interacts with capacity limitations. Across three experiments, 7-year-old children and adults were shown arrays of stimuli and later had to indicate whether a probe item had been in the original stimuli array. In some trials (cued trials), participants were shown an arrow pointing to the location of the stimuli to be probed. This arrow cue came either before (pre-cue) or after (retro-cue) presentation of the stimuli. In other trials, instead of the informative arrow, participants were shown an uninformative central square (neutral trials). The memory load (number of items in the stimuli array) and timing of the cue varied across experiments.

These experiments were interpreted as showing that children and adults differ in the efficiency of attentional control, in addition to capacity. However, these interpretations were made without the use of a computational model to specify how variation in capacity and control induce changes in behavior. We aimed to determine whether, when utilizing such a model, these interpretations are borne out. We review the details of the three studies in our results.

Computational models of working memory

A wide set of models have framed the limited capacity of working memory as reflecting a limited resource that needs to be divided across stimuli (Bays et al., 2022). Such models have differed in whether that resource is fundamentally a discrete quantity like a slot (Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2008), or a continuous quantity like precision (Van den Berg et al., 2012, 2014). While the model that we use assumes a discrete resource, quanta, this is not meant as a substantive claim, and this resource could be interpreted in terms of random samples, a view which may bridge these perspectives (Schneegans et al., 2020).

Within the framework of resource limitations, attentional control has been conceived of as a meta-decision about which stimuli resources should be invested in. Resource rationality postulates that this decision be made by balancing a cost of spending a limited resource with an increase in accuracy in responding to an eventual query (Lieder & Griffiths, 2020). Van den Berg & Ma (2018) demonstrated that when framed as a single decision over where to spend resources, this framework could explain effects of set size and probe probability. Suchow & Griffiths (2016) introduced a related model where this process was framed as part of a multi-step decision process where an agent makes repeated decisions about how to shift resources between stimuli over time. Because such a multi-step framing is necessary to model manipulations of delay-time and retro-cue, we build on this model for simulating the experiments in Shimi & Scerif (2017). We describe this model in more detail below.

We note that another line of work has used recurrent neural networks to model retro-cue tasks (Piwek et al., 2023; Wan et al., 2022). While such models have captured aspects of behavioral and neural data, specifying what exactly corresponds to varying control in such models is opaque, thus making them more challenging to adapt for our purposes.

Model

We extended the computational model of Suchow & Griffiths (2016) to determine whether previously reported differences in working memory between adults and humans (Shimi & Scerif, 2017), reflect differences in capacity and control. In a given trial, the model assumes an agent that is presented with S stimuli that it seeks to remember. The agent possesses N units of a discrete resource, quanta, that determine the success of recall for each stimulus. At stimulus presentation, quanta are randomly allocated across the items. Quanta change their item assignments over time, according to a Moran process (Moran, 1958), modified such that the agent has limited control over the dynamics. At each time step, noise randomly causes one of the quanta, q_E, to change its assignment to the current assignment of another, q_A, that is selected by the agent (Fig. 1A; adapted from Suchow et al., 2017). To foreshadow, we will relate capacity and control respectively to the total number of quanta the agent has available, and to its ability to optimally direct quanta to the appropriate stimuli.

At a given time point, the agent will be probed to recall one of the S stimuli, with recall success being determined by the number of quanta assigned to that stimulus. Increasing the amount of quanta assigned to a probed stimulus, i, increases the probability of correct recall with diminishing returns, $P_{recall}(i) = .5 + \frac{1}{2\times(1+e^{-\beta_{recall}n_i})}$ (Fig. 1B). Here, n_i is the number of quanta assigned to a stimulus i at the time it is probed and β_{recall} determines the steepness of the function. If the agent selects actions randomly, the number of quanta assigned to each stimulus change gradually over time (Fig. 1C). When a stimulus loses all of its quanta it cannot acquire any more, so the stimulus is forgotten. After a long time has passed, this leads to a steady state where one stimulus has all quanta assigned to it, and the remaining stimuli have none. Given a time-series of quanta assignments, the recall function (Fig. 1B) specifies the probability that the agent would correctly answer a query about each stimulus at each time point (Fig. 1D).

The agent can improve its chances of correctly answering the query by selecting optimal actions. The model assumes that the agent does not know what time point it will be queried and thus tries to maximize its chances of correctly answering the query at every future time point. This is implemented by framing the agent as receiving a reward, r_t, at each time point, t, which is the probability it would correctly answer a query at that time point. This depends on both the quanta assignments

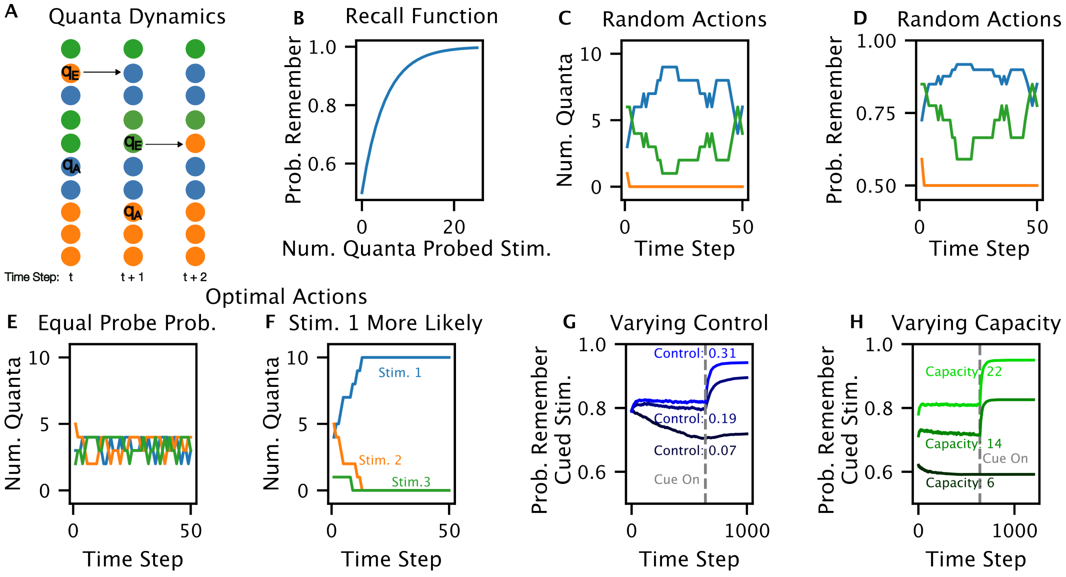

Figure 1: **Model of capacity and control in working memory.** A. MDP model of memory. Three time steps are displayed. Each circle denotes a quantum and color denotes stimulus assignment. q_E denotes environment selected quantum, whose assignment changes to that of q_A, selected by the agent. B. Recall function: the probability a probed stimulus will be remembered as a function of the number of quanta assigned to it ($\beta_{recall} = .2$). C, D. Example trial with random actions (no control). Each line denotes the number of quanta assigned to a stimulus (C) and probability correct response (D) if that stimulus were probed. E,F. Example trials with optimal actions selected (full control). E) Stimuli are equally likely to be probed. F) One stimulus is more likely to be probed. G,H) Varying control and capacity in a retro-cue experiment. Lines show mean probability of remembering the cued stimulus over 1000 trials for a given setting of control (G) or capacity (H). G: Capacity is set to 24. H: Control is set to $1 - \varepsilon = .45$.

at that time point and also on beliefs about which stimulus will be probed, $r_t = \sum_{i=1}^{S} P_{probe}(i) \times P_{recall}(n_{i,t})$. Here, $n_{i,t}$ is the number of quanta assigned to stimulus i at time point t, and $P_{probe}(i)$ is the agent's beliefs about the probability that stimulus i will be probed.

Combined with the dynamics, this reward function defines a Markov decision process (Sutton & Barto, 2018) for which optimal actions, which maximize the expected future-discounted reward, $a_t^* = \arg\max_{a_t} E[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | a_t]$, maximize the probability of correct responses to future queries. Because solving for such optimal actions using dynamic programming was intractable for problems with more than 3 stimuli or 10 quanta, we approximated optimal actions as those which maximized the next-step reward, $a_t^* = \arg\max_{a_t} E[r_{t+1} | a_t]$. This approximation works well for this class of reward function.

Optimal actions depend on beliefs about which stimulus will be probed, P_{probe}, which can be altered by cues. Without a cue, the agent assumes that all stimuli are equally likely, $P_{probe}(i) = \frac{1}{S}$. This leads to optimal actions that add quanta to stimuli with the lowest number (example trial in Fig. 1E). This is because when additional quanta have a diminishing marginal effect on recall, the average recall probability is maximized when quanta are shared equally across all stimuli. For experiments modeled here, cues always signal the probed item. Thus, following a cue, $P_{probe}(i) = 1$ if i is the cued stimulus and 0 otherwise. This leads to optimal actions that add quanta to the cued stimulus, which improves that chances that it will be recalled (Fig. 1F). To simulate retro-cue trials, where a cue is introduced during the maintenance period, these action-selection policies are combined, where before the cue actions are generated based on an assumption of equally probable probes (Fig. 1E), and after the cue on the assumption that the cued stimulus will be probed (Fig. 1F).

We formalize notions of capacity and control in the model by allowing two parameters to vary between agents. We define capacity as the number of quanta. To define control, we assume that, with probability ε, the optimal action fails and instead a randomly selected action is performed. The control parameter is defined as $1 - \varepsilon$. We illustrate the effects of varying capacity and control by simulating a 3-stimulus retro-cue experiment for agents defined with varying combinations of capacity and control (Fig. 1G, H). Increasing either capacity or control leads to improved recall. However, changes in either differently affect recall dynamics.

Methods

We evaluated what parameter combinations could produce recall performance consistent with either children or adult behavior in Shimi & Scerif (2017). Across three experiments that manipulated the type of cue, its timing, and the number of stimuli, Shimi & Scerif (2017) report mean Cowan's K values for 24 different conditions for adults and 20 different conditions for children. Cowan's K is defined as $N \times (h - f)$, where h is the hit rate and f is the false alarm rate. We performed a parameter sweep, varying capacity between 2 and 50 in increments of 1, and control between 0 and 1 in increments of .02. For these parameter combinations, we computed the mean squared error (MSE) between model-predicted and reported Cowan's K across each condition. For each condition, predicted Cowan's K was computed by simulating 1000 trials from that condition to compute a mean probability of correct response, p_{recall}, to the probed item at the time of the query (e.g. Fig. 1G,H). Predicted Cowan's K was computed by assuming no response bias, $h = p_{recall}$, and $f = 1 - p_{recall}$.

Simulating the model also required specifying the steepness of the recall function, β_{recall}, and the number of model time steps that occur per second, NT. To identify these hyperparameters, we performed a sweep, varying β_{recall} over .025, .05, .1, .2, .4 and varying NT over 25, 50, 100, 200, 400, 800. For each setting of β_{recall} and NT, we performed a coarse search over capacity and control, varying each over 20 values between 0 and 100 and 0 and 1 respectively, to estimate the minimum total MSE for that setting (under a model where children and adults differ in capacity and control parameters). Results presented below correspond to $NT = 800$ and $\beta_{recall} = .2$, which achieved overall minimum total MSE. We note that our results do depend on the setting of these hyperparameters and it is possible that future consideration of additional studies will lead to a more informed setting of these. We return to this point in the discussion.

Results

Fig. 2 shows MSE between model-predicted and reported Cowan's K values across all three experiments for children (A) and adults (B) as a function of control and capacity parameters. Contour lines designating parameter values with lowest 5% MSE (white lines) demonstrate that increasing capacity and control trade-off in explaining the data. This trade-off is especially apparent in children, whose recall can potentially be explained both by (1) moderate capacity paired with low control, or (2) low capacity paired with moderate control. We formally considered three possibilities for how children and adults differ in control and capacity. First, adults have higher capacity but equal control relative to children. Second, adults have higher control, but equal capacity relative to children. Third, adults have both higher capacity and higher control. To separate these possibilities, we defined four models, which permit comparison of the degree to which each of these possibilities can explain differences in behavior between children and adults. *Capacity Shared* allowed control to vary between adults and children but constrained capacity to be the same. *Control Shared* allowed capacity to vary between adults and children but constrained control to be the same. *Neither Shared* allowed both control and capacity to vary between adults and children. For completeness, *Both Shared* required both capacity and control to be the same for adults and children. Fig. 2A-C shows the parameter values selected by these models. Notably, parameters selected by either Control Shared or Capacity Shared models lie in the lowest 5% MSE range, suggesting that either can achieve relatively low error (Fig. 2C). To formally compare these models, we computed the AIC of each model's fit, treating MSE as a measure of (scaled) negative log likelihood (thus assuming a Gaussian error distribution). Control Shared provided the most parsimonious fit to the data, marginally outperforming Capacity Shared ($\Delta AIC = 1.1$) and Both Shared ($\Delta AIC = 1.9$), and greatly outperforming Neither Shared ($\Delta AIC = 28.6$).

These results suggest that the data reported in Shimi & Scerif (2017) weakly favor developmental differences in capacity alone. To more strongly determine the advantage of this explanation over alternatives, we next examined each model's ability to account for qualitative differences in behavior between children and adults (Fig. 3). This was done by plotting each model's predictions for each condition in each experiment, at the best parameter values for adults and children. Surprisingly, we find that models with either shared capacity or shared control can both capture the overall qualitative patterns in the data. Rather, the only model which can be confidently ruled out is Both Shared, for which neither component differs between adults and children.

Experiment 1 compared the effect of memory load (number of stimuli) on the effect of different types of cues on memory between adults and children (Fig. 3A). The experimental conditions varied were the number stimuli (Load: 2 or 4), whether a cue indicated which item would be probed (Cued or Neutral), and whether the cue was presented prior to (300 ms; Pre) or following (300 ms; Retro) stimuli presentation. Empirically, both children and adults benefited from cues more for four stimuli compared to two stimuli. Additionally, compared to adults, children made less use of retro-cues compared to pre-cues. All models except for Both Shared were able to capture both of these effects qualitatively.

Experiment 2 compared the effect of time between stimulus presentation and cue presentation on the effect of cues on memory (Fig. 3B). Conditions varied whether a (retro-)cue was presented 200 ms (IM) or 1000 ms (VSTM) following stimulus presentation, and also whether a retro-cue indicated which item would be probed (Cued or Neutral). Empirically, compared to adults, children's performance on both cued and uncued trials decayed more over time. For children, cues still provided a benefit in both IM and VSTM conditions, however less so than for adults. Additionally, for children, the effect of cue was less in the VSTM condition than in the IM condition. As with Experiment 1, all models except for Both Shared could capture these qualitative effects. However, we note that the predicted extent of decrease in neutral condition over time is less than was observed.

Experiment 3 manipulated both load and decay, in order to look at how their interaction affected recall in children versus adults. Load was varied between 3 and 6 stimuli and delay was varied between the IM and VSTM conditions, as well

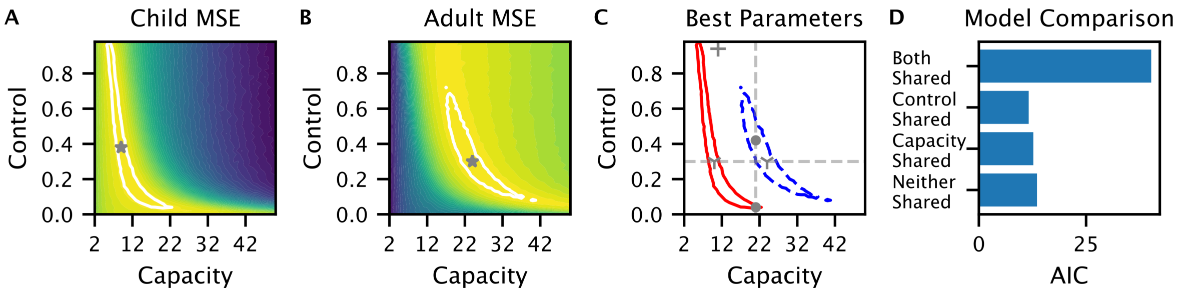

Figure 2: **Model fit to three experiments in Shimi and Scerif (2017).** A, B. Mean squared error (MSE) between model-generated predictions and data from children (A) and adults (B). Brighter image intensity depicts better fit to data (lower MSE). White line designates contour encircling parameters with lowest 5% MSE. Stars designate parameter setting with lowest MSE if children and adults vary in both capacity and control (Neither Shared Model). C. Models where capacity varies keep parameters in lowest MSE range. Colored lines denote contour encircling parameters with lowest 5% MSE for children (red) and adults (blue). Gray symbols denote best fit parameters for models for which either the capacity parameter is shared between adults and children (Capacity Shared; circle), the control parameter is shared between adults and children (Control Shared; downward tri-arrow), or both a parameters are shared (Both Shared; plus). D. A model with shared control and varying capacity parameters marginally provides the most parsimonious fit to data. Akaike Information Criterion (AIC) for models which vary depending on whether either, both or neither control and capacity parameters are shared between adults and children.

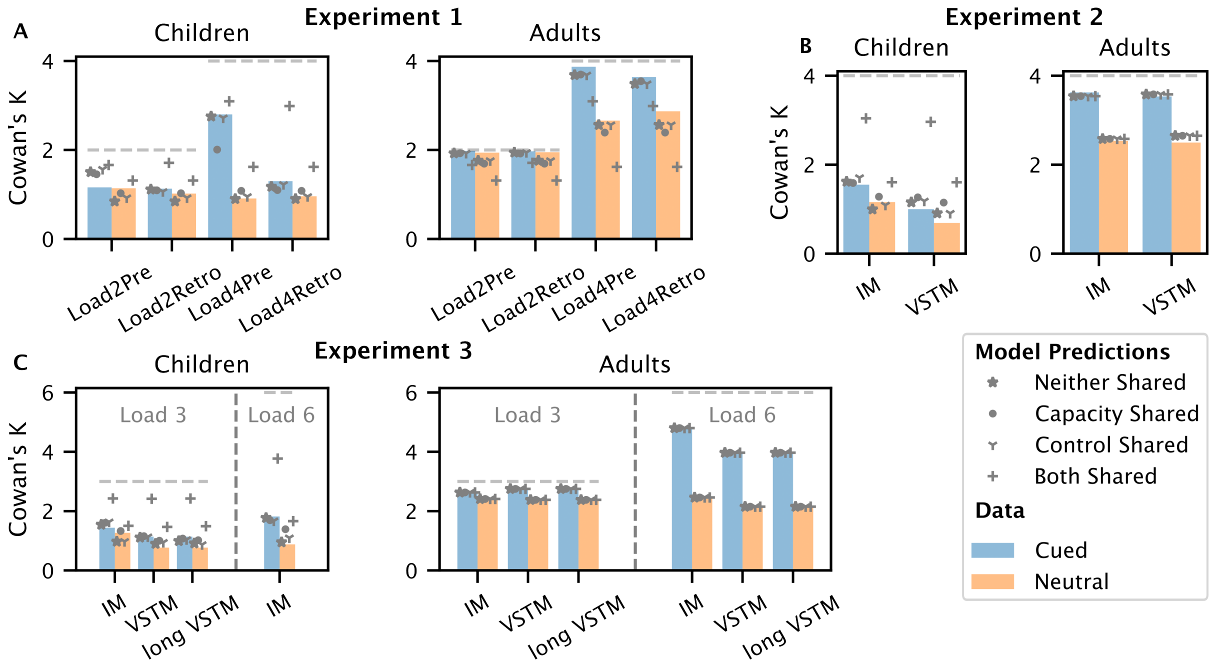

Figure 3: **Differences in either capacity or control can explain qualitative differences between children and adults across three experiments in Shimi and Scerif (2017).** A. Experiment 1. Conditions varied on number of stimuli, either 2 (Load 2) or 4 (Load 4), whether a cue was informative or not (Cued or Neutral), and whether the cue was presented before or after stimulus presentation (Pre or Retro 300ms before or after presentation). B. Experiment 2. Conditions varied whether a (retro-)cue was informative (Cued or Neutral) as well as whether the (retro-)cue was presented 200 ms (IM) or 1000 ms (VSTM) following stimulus presentation. C. Experiment 3. Conditions varied whether a (retro-)cue was informative (Cued or Neutral), whether the retro-(cue) was presented 200 ms (IM) 1000 ms (VSTM) or 1800 ms (long VSTM) following stimulus presentation, and also whether 3 (Load 3) or 6 (Load 6) stimuli were presented. For Load 6, children only completed the IM condition. A-C) Bars reflect Cowan's K values reported in Shimi and Scerif (2017). Symbols reflect model predictions at best-fit parameter values. Horizontal lines denote perfect memory.

as a third longer condition, longer VSTM, which increased the time between stimulus presentation and cue to 1800 ms. In the Load 6 condition, children only completed IM. Empirically, children's memory decay plateaued after the IM period. Adults had larger effects of cues, and additionally, had slower overall decay. As with the above experiments, we found that all models except for Both Shared were able to capture these effects qualitatively.

Overall, we find that models with either shared capacity or shared control can produce qualitative patterns observed in the data. This suggests that differences between adults and children could be in capacity, control, or both processes, without giving strong support to one scenario over the others. While we find that models with varying capacity provide a marginally better quantitative fit than models with varying control, potentially due to better prediction of the effect of cues on children's performance (Fig. 3A), we discuss caveats to this interpretation below.

Discussion

While numerous studies have documented differences in working memory between children and adults, whether these differences reflect underlying differences in capacity or control is still debated. Here, we moved towards resolving this debate by extending a computational model of working memory so that capacity and control could be captured via model parameters. This allowed examination of possible explanations for differences in recall between children and adults reported in a recent study (Shimi & Scerif, 2017). Informed intuition provides the expectation that adults will have superior working memory because they have both greater capacity and control. Surprisingly, and in contrast to the interpretation offered in Shimi & Scerif (2017), we find that qualitative differences in recall between children and adults can be produced either by variation in capacity or control alone. Together, our findings motivate further examination into what changes during the development of working memory.

Overall, our findings suggest alternative interpretations of aspects of the experiments in Shimi & Scerif (2017). First, it is often assumed that the extent of an effect of a pre-cue or retro-cue indicates efficacy of attentional control. However, as shown in Fig. 1 (E,F), changes in capacity can also drive the extent of a retro-cue effect, provided some control is present. Second, Shimi & Scerif (2017) as well as prior studies (Cowan et al., 2010) interpret the finding that retro-cue effects on recall increase with load on working memory as evidence that control is strategically deployed more when load is greater. However, our model does not vary control dependent on load, and can recapitulate these findings. Finally, Shimi & Scerif (2017) interpret differences in the retro-cue effect between IM and VSTM manipulations to suggest that these memory systems provide fundamentally different representational systems over which retro-cues can operate. Interestingly, our model does not assume distinct memory systems that are differently affected by cueing, yet can still capture the qualitative findings in Experiments 2 and 3.

We note that the interpretation of the results of Shimi & Scerif (2017) offered by the model may depend on the setting of hyperparameters NT and β_{recall} which respectively set the rate of model time steps and the recall function. In additional simulations, not reported due to space limitations, we found that the setting of these hyperparameters affects the ability of the Capacity Shared model to explain differences in adult's and children's performance by affecting the rate at which predicted performance converges following a cue. Because the influence of control on performance requires time (Fig. 1G), speeding up convergence causes control to act more like capacity by allowing its effects to occur more instantaneously. In line with this, we found that speeding up convergence by increasing NT increases the ability of variation in control alone to explain developmental differences. In contrast, decreasing either NT or β_{recall} decreases the ability of control to explain differences, however also worsens the fit of all models. Whereas here we set NT and β_{recall} to values which minimized MSE of the Neither Shared Model, it is likely that using additional studies to further constrain these parameters will improve model inferences.

Because our model's quanta dynamics were based on the Moran process, results from evolutionary biology may themselves help with understanding how control and capacity interact. The Moran process has been used to examine how evolution is influenced by stochastic events (genetic drift) versus deterministic selection (advantageous genes propagate) (Otto & Day, 2007). A canonical finding is that selection strongly influences which genes are retained in large populations. This is because in small populations stochastic events can idiosyncratically remove a substantial proportion of advantageous genes, before selection allows them to accrue. By analogy, a larger capacity reduces the influence of stochastic events that cause quanta to be assigned in suboptimal ways across items (e.g. for items to lose all quanta and be forgotten). Consequently, control (analogue of selection) has more chance to drive quanta to a desirable distribution when capacity is high (e.g. allocating quanta to a cued item). That is, increased capacity boosts the efficacy of control, so that substantial increases in control are not needed to produce the better performance in recall. This may explain the characteristic capacity-control trade-off found by our model in fitting recall data (Fig 2).

In future work, we intend to understand how increasing control pays off for an agent as a function of capacity, and vice versa. This will enable application of our models to evolutionary modeling, which in turn should generate principled predictions for cross-species differences in control and capacity.

Acknowledgements

This work was supported by a grant from the Templeton World Charity Foundation under the Diverse Intelligence Initiative (TWCF Number: 20648).

References

Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception, & Psychophysics*, *74*, 146–162.

Bays, P., Schneegans, S., Ma, W. J., & Brady, T. (2022). Representation and computation in working memory.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114.

Cowan, N. (2012). *Working memory capacity*. Psychology Press.

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*(2), 239–264.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, *13*(1), 120–133.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177-190.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

ManyPrimates, Aguenounon, G., Allritz, M., Altschul, D., Ballesta, S., Ballesta, A., … others (2022). The evolution of primate short-term memory. *Animal Behavior and Cognition*, *9*(4), 428–516.

Moran, P. A. P. (1958). Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society* (Vol. 54, pp. 60–71).

Oberauer, K. (2019). Working memory and attention–a conceptual analysis and review. *Journal of Cognition*, *2*(1).

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, *55*(4), 601–626.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., … others (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885.

Otto, S. P., & Day, T. (2007). *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press.

Piwek, E. P., Stokes, M. G., & Summerfield, C. (2023). A recurrent neural network model of prefrontal brain activity during a working memory task. *PLOS Computational Biology*, *19*(10), e1011555.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex*, *43*(1), 135–145.

Schneegans, S., Taylor, R., & Bays, P. M. (2020). Stochastic sampling provides a unifying account of visual working memory limits. *Proceedings of the National Academy of Sciences*, *117*(34), 20959–20968.

Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child Development*, *85*(2), 578–592.

Shimi, A., & Scerif, G. (2017). Towards an integrative model of visual short-term memory maintenance: Evidence from the effects of attentional control, load, decay, and their interactions in childhood. *Cognition*, *169*, 61–83.

Suchow, J. W., Bourgin, D. D., & Griffiths, T. L. (2017). Evolution in mind: Evolutionary dynamics, cognitive processes, and bayesian inference. *Trends in Cognitive Sciences*, *21*(7), 522–530.

Suchow, J. W., & Griffiths, T. (2016). Deciding to remember: Memory maintenance as a markov decision process. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, *121*(1), 124-149.

Van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. *ELife*, *7*, e34963.

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785.

van Moorselaar, D., Olivers, C. N., Theeuwes, J., Lamme, V. A., & Sligte, I. G. (2015). Forgotten but not gone: Retro-cue costs and benefits in a double-cueing paradigm suggest multiple states in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1755.

Wan, Q., Menendez, J. A., & Postle, B. R. (2022). Priority-based transformations of stimulus representation in visual working memory. *PLOS Computational Biology*, *18*(6), e1009062.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235.