

Developmental improvements in the resolution and capacity of visual working memory share a common source

Vanessa R. Simmering¹ · Hilary E. Miller¹

Published online: 21 June 2016
© The Psychonomic Society, Inc. 2016

Abstract The nature of visual working memory (VWM) representations is currently a source of debate between characterizations as slot-like versus a flexibly-divided pool of resources. Recently, a dynamic neural field model has been proposed as an alternative account that focuses more on the processes by which VWM representations are formed, maintained, and used in service of behavior. This dynamic model has explained developmental increases in VWM capacity and resolution through strengthening excitatory and inhibitory connections. Simulations of developmental improvements in VWM resolution suggest that one important change is the accuracy of comparisons between items held in memory and new inputs. Thus, the ability to detect changes is a critical component of developmental improvements in VWM performance across tasks, leading to the prediction that capacity and resolution should correlate during childhood. Comparing 5- to 8-year-old children's performance across color discrimination and change detection tasks revealed the predicted correlation

Significance The dominant explanations of visual working memory capacity limits have focused on the nature of representations as slot-versus resource-like. An alternative approach uses a dynamic neural field model to understand the processes that form, maintain, and use representations in service of behavior. Model simulations suggest that previous empirical work may have conflated the precision of representations and the resolution with which those representations can be used to generate behavior. We present results supporting the model prediction that resolution and capacity correlate during childhood, as both reflect the ability to detect changes, which improves developmentally through increasing memory stability.

✉ Vanessa R. Simmering
simmering@wisc.edu

¹ Department of Psychology, McPherson Eye Research Institute, and Waisman Center, University of Wisconsin, 1202 W. Johnson St., Madison, WI 53706, USA

between estimates of VWM capacity and resolution, supporting the hypothesis that increasing connectivity underlies improvements in VWM during childhood. These results demonstrate the importance of formalizing the processes that support the use of VWM, rather than focusing solely on the nature of representations. We conclude by considering our results in the broader context of VWM development.

Keywords Development · Memory: Visual working and short-term memory · Neural network modeling

Visual working memory (VWM)¹ plays a crucial role in everyday behavior by connecting our current experiences with long term memory of familiar objects, people, and places. Decades of research have demonstrated that VWM has a limited capacity, estimated to be only 3–5 simple items in adults (Cowan, 2010). However, the source of this limit remains hotly debated: Is capacity better conceptualized as arising from slot-like representations or a flexibly-allocated resource? We propose a dynamic systems account as an alternative approach to understanding capacity limits through considering the *processes* by which memory representations are formed, maintained, and used in service of behavior (Johnson, Simmering, & Buss, 2014; see also Johnson, Spencer, & Schöner, 2009) and how they change over development (Simmering, *in press*; Simmering, Miller, & Bohache, 2015; Simmering & Patterson, 2012). The current paper tests a prediction of this dynamic account of VWM, namely that developmental improvements in the capacity and resolution of VWM arise through a common underlying source.

¹ There has been inconsistent use of the terms “working” versus “short-term” memory to describe performance in the types of laboratory tasks used here. We use working memory in this paper to be more inclusive across tasks (see Simmering & Perone, 2013, for discussion).

The central argument we put forth here is that a more complete understanding of cognition and development can be gained from taking a process-based approach (which we achieve through computational modeling) rather than focusing primarily on the nature of representations. As an illustration of this approach, we show how a careful examination of the processes supporting the formation, maintenance, and use of representations in service of behavior lead to specific predictions for how performance relates across tasks estimating the capacity and resolution of VWM during early childhood. This predicted link between these characteristics of VWM does not discount prior theories of capacity limits, but rather suggests that these theories were incomplete in their accounts of how representations support behavior across tasks. By taking a process-based account, our work helps to reconcile previous results that appear inconsistent (e.g., Simmering, *in press*, described further in the [Discussion](#) section) and provides a more complete theory that synthesizes across different tasks and age groups. Importantly, our process-based approach shows that modifications to task structure can reveal different characteristics of the underlying memory system.

In the sections that follow, we first review theories of VWM capacity limits in adults, then discuss developmental changes in VWM and possible mechanisms proposed to account for these changes. Within these sections we introduce the dynamic neural field modeling framework that lays the foundation for the prediction tested here. Following our empirical results, we consider the implications of these findings for other theories and studies assessing precision and resolution through different behavioral tasks. Finally, we conclude by discussing the potential insights to be gained from implementing behavioral tasks within a process model.

Assessing visual working memory representations in adults

Decades of research have demonstrated that the amount of information that can be held in VWM is severely limited, but the nature of this limit is a continued source of debate (see Wolfe, 2014, for a recent special issue on the topic). Some theorists argue that VWM is limited by the *number of items* that can be held discretely in memory (often referred to as “slots” theories; e.g., Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2008), whereas others contend there is no limit on the number of items, but rather the *resolution of representations* is inversely related to the number of items to be remembered (often referred to as “resource” theories; e.g., Bays & Husain, 2008; Wilken & Ma, 2004). Most of these characterizations, as well as “hybrid” theories that consider both number and resolution to be limited (e.g., Alvarez & Cavanagh, 2004; Xu & Chun, 2006), focus almost exclusively on the nature of the representations within VWM without

much consideration for how such representations are formed, maintained, and used in service of behavior across different types of tasks. Here we argue that taking an alternative account, a process-based approach to understand how the dynamics of the VWM system lead to the limitations evident across behavioral tasks (Johnson et al., 2014), will provide a more complete account of cognition and development.

A classic paradigm for estimating VWM capacity is the change detection task. In this task, a memory array with a small number of simple objects (e.g., colored squares) is presented briefly (100–500 ms) on a computer screen, followed by a short delay (1 s), and presentation of a test array. The test array may include the same number of items as in memory array, or only a single item. Typically, the item(s) match the memory array on half the trials (no-change trials) and differ(s) by one item on the other half of trials (change trials). Participants respond “same” or “different” to end the trial. To estimate capacity (K) from this task, hit rates (correct responses on change trials) and false alarm rates (incorrect responses on no-change trials) are combined according to a formula that takes into account the set size, or number of items within the memory array. Adults’ performance declines (i.e., hits decrease and false alarms increase) as the set size increases, and calculations of K are used to estimate the number of items the participant held in memory across trials (see Rouder, Morey, Morey, & Cowan, 2011, for discussion).

Although this paradigm lends itself well to capacity estimates assuming slot-like representation of items, it was not designed to assess other characteristics of VWM, such as the precision of memory representations or the certainty of decisions. Wilken and Ma (2004) compared this classic paradigm (with the addition of confidence ratings) to a new paradigm in which participants were cued to recall the feature of one item at test. This type of estimation response (e.g., selecting the color in memory from a color wheel) allowed variance in the response that could not be assessed through the typically large changes used in the change detection paradigm (although the type of change can be manipulated to achieve a more fine-grained assessment of VWM in change detection; e.g., Awh, Barton, & Vogel, 2007). By fitting results from both change detection and recall paradigms with a number of signal detection models, Wilken and Ma showed that set-size dependent increases in representational noise (i.e., reduced precision) could account for behavioral results without positing a limit on the number of items held in memory.

More recently, van den Berg, Awh, and Ma (2014) fit behavioral results from ten studies with the delayed estimation paradigm with a set of 32 models that varied in assumptions regarding the number and precision of items held in memory, as well as whether colors may be mis-localized in space (accounting for response errors in which the wrong target value is recalled). They concluded that the best fit to the data arose from models with variable precision and a limited number of

representations, as well as potential (but rare) wrong-target responses (see van den Berg et al., 2014, for further details). These model comparisons highlight the need to consider multiple dimensions of VWM representations when explaining adults' performance on lab tasks: empirical and computational results do not support theories attributing performance limitations to only the number or only the precision of items in VWM. As such, a complete theory of VWM performance must account for limits in both the number and resolution of items held in memory.

As an alternative to theories focusing primarily on representations, Johnson et al. (2014) presented a computational model that can capture these characteristics through the same underlying processes by specifying how representations are formed, maintained, and used in the change detection task. Their dynamic neural field model architecture (referred to as “dynamic model” here) is composed of two excitatory fields coupled to a shared inhibitory field. Visual features (e.g., color or orientation) are represented as localized peaks of activation within these excitatory fields. Nodes within these fields are connected through local excitatory interactions that are mediated by lateral inhibitory interactions through the inhibitory field. The two excitatory fields correspond to the perceptual processing and memory/maintenance of visual information. Interactions among fields provide a mechanism for comparison between items in memory and perceptual inputs.

A sample simulation is shown in Fig. 1 to illustrate how the dynamic model performs a trial in the change detection task, with parameters tuned to adults' performance (Simmering, *in press*). Three inputs to the model, representing three colors in the memory array, formed three input-driven peaks in the perceptual field (Fig. 1a). During the presentation of the memory array, activation projected from the perceptual field to the working memory field, where the items were encoded as self-sustaining peaks (Fig. 1b). Activation projected from both excitatory fields into the shared inhibitory field, producing bumps of activation at the color values associated with the items held in working memory (see *Inhib* in Fig. 1b). The projection of inhibition back to the excitatory fields throughout the delay led to troughs of inhibition surrounding the peaks in the working memory field and at the corresponding values in the perceptual field (Fig. 1c). When the test array was presented (Fig. 1d), two inputs that matched items held in memory were suppressed in the perceptual field (see dashed circles), but a third novel input produces an input-driven peak (see solid circle). Decisions in the model are generated through a simple two-node system in which activation from the working memory field projects to a “same” decision node and activation from the perceptual field projects to a “different” decision node (not shown); these nodes compete in a winner-take-all fashion to generate a single response on each trial (see Johnson et al., 2014, for further details). In the sample trial in Fig. 1, this produced a correct “different”

decision driven by the peak in the perceptual field; note that projections to the decision system are tuned such that a single novel item can produce this response even when multiple items are held in working memory.

Similar to the model that provided the best fit in van den Berg et al.'s (2014) comparison, the dynamic model presented by Johnson et al. (2014) captures performance through variable number and resolution of items held in VWM.² Unlike theories of VWM that focus on the nature of representations (i.e., slots vs. resources), the dynamic model includes a process account of comparison between items in VWM and the test array in change detection, as well as the generation of a “same” or “different” response. Model simulations suggest that errors can arise through these processes as well, not just due to insufficient number or resolution of items in memory. An important question is how the precision metric used by van den Berg et al. relates to the resolution described by Johnson et al. Although these terms have been used with some overlap in the literature (e.g., Fukuda, Vogel, Mayr, & Awh, 2010), simulations of the dynamic model show that they may be dissociable when specifying how representations are used to generate behavior. The precision of a VWM representation (i.e., width of a peak along the color dimension) can be partially separated from the resolution with which that representation can be used (i.e., magnitude of change that can be reliably detected) due to the differences in how these characteristics relate to behavior. That is, through the continuous interactions within the fields, two peaks with similar width might have different inhibitory projections into the perceptual field, which would affect the reliability with which each peak could be used to detect small differences. Due to stochastic fluctuations in activation, it is likely that representation of the same color value across repeated trials will vary in the depth and width of the inhibitory trough, which will produce within-subject variability in responses to the same target across trials (cf. Schutte & Spencer, 2009; see Simmering & Patterson, 2012, for related behavioral results). Because this variability relates directly to the excitatory and inhibitory processes that support the representations and response generation, it emerges from the process-based approach of the dynamic model. The models tested by van den Berg et al., by contrast, were not used to explicitly simulate the generation of behavioral responses across trials, which leaves room for further explanations of the processes that operate upon memory representation in the context of different tasks.

In summary, model fits and simulations suggest that the dominant explanations of VWM representations as slot-

² One dimension addressed by van den Berg et al. (2014) but not Johnson et al. (2014) is the localization of targets in space. This has been addressed within the broader dynamic field theory framework (Schneegans, Spencer, & Schöner, 2015), and may be particularly important for understanding developmental change (Simmering, 2015; Simmering & Wood, 2016).

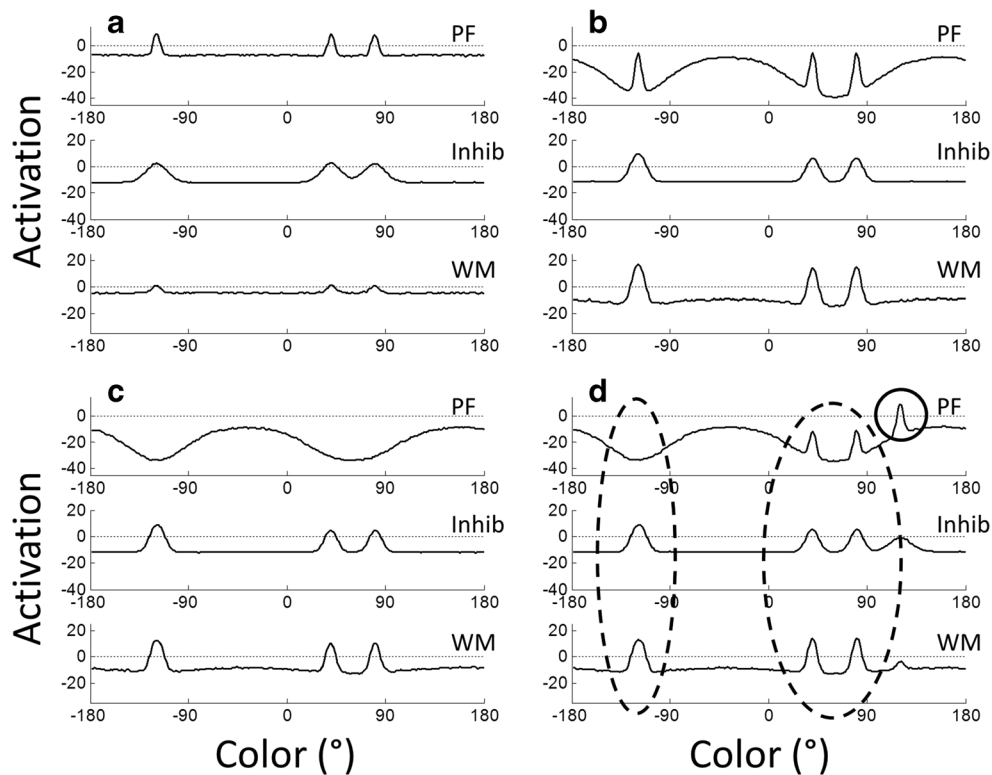


Fig. 1 Dynamic model simulation of a change detection trial with parameters tuned to adults' performance. Time-slices through the three fields show activation at: (a) the beginning of the memory array presentation; (b) the end of the memory array presentation; (c) the end of the

delay; (d) the generation of a response during the test array presentation. *PF* perceptual field, *Inhib* inhibitory field, *WM* working memory field. Horizontal dashed lines in each field indicate the activation threshold (0) for interactions among nodes

versus resource-like have not provided a complete account of performance, and that both the number and resolution of items vary in VWM. The process-based dynamic model suggests errors can arise during encoding, maintenance, comparison, and/or decision in the change detection task. Not only has this model quantitatively captured capacity estimates from change detection (Johnson et al., 2014; Simmering, *in press*), but it has also generated novel predictions regarding how items interact in memory (Johnson et al., 2014; Johnson, Spencer, Luck, & Schöner, 2009). In addition to accounting for adults' performance, this model includes a specific account for how VWM changes over development to account for improvements in both capacity and resolution.

Developmental changes in visual working memory

The majority of studies examining developmental change in VWM have used the change detection paradigm. Results have shown increasing capacity estimates from early childhood through adolescence (Buss, Fox, Boas, & Spencer, 2013; Cowan et al., 2005; Isbell, Fukuda, Neville, & Vogel, 2015; Kharitonova, Winter, & Sheridan, 2015; Riggs, McTaggart, Simpson, & Freeman, 2006; Riggs, Simpson, & Potts, 2011;

Simmering, 2012; Simmering et al., 2015; see Simmering, *in press*, for discussion of paradigms used to estimate capacity during infancy), although the magnitudes and trajectories of these estimates varies widely across paradigms (see Simmering & Perone, 2013, for related discussion). Most developmental studies have either implicitly or explicitly endorsed a slot-like view, suggesting that the number of items that can be represented increases without addressing whether there are corresponding changes in resolution (e.g., Cowan et al., 2005; Isbell et al., 2015; Kharitonova et al., 2015; Riggs et al., 2006, 2011).

Other studies have tested for developmental increases in precision and resolution. For example, one study using an orientation version of the delayed-estimation paradigm found developmental increase in precision between 8 and 13 years of age (Burnett Heyes, Zokaei, van der Staaij, Bays, & Husain, 2012). The authors attributed performance to developmental increase in precision as resources were divided among representations. However, results from a similar paradigm comparing performance between younger (7–9 years) versus older (10–12 years) children suggested no change in precision between groups, but rather a developmental decrease in incorrect-target responses (Sarigiannidis, Crickmore, & Astle, 2016). Together, these two studies indicate that

children's estimation responses improve during this developmental period, but it is unclear whether this change is attributable to changes in the nature of representations, the use of those representations, or both.

Simmering and Patterson (2012) used a color discrimination paradigm to test for developmental improvements in VWM resolution. In this task, children briefly viewed two colors separated by a short delay, then responded "same" or "different" regarding whether the colors matched exactly. To estimate thresholds for discriminating between colors, the difference in color between the two presentations increased or decreased in small increments across trials following a stair-casing procedure. By conducting multiple runs to each of two target colors, Simmering and Patterson calculated both a mean difference in colors necessary to respond "different," as well as variability across runs within each participant. Results showed improvements in both measures (mean difference and variability) between 4 and 6 years of age, indicating improved resolution in VWM over development. These results supported the prediction that the reliability with which children could use VWM representations to identify small changes improves during early childhood.

Together these empirical studies demonstrate developmental increases in the capacity, precision, and resolution of VWM. However, because these paradigms have focused on only one characteristic of memory at a time (i.e., capacity in change detection, precision in delayed estimation, or resolution in discrimination), it is unknown how these improvements relate to one another over development. In parallel to the theories of VWM in adults, the dominant theories that propose developmental increases in slots (e.g., Riggs et al., 2011) or improved resource distribution (e.g., Burnett Heyes et al., 2012) focus on the representations without specifying how VWM operates in the context of behavior. By contrast, the dynamic model framework describe above has been applied to both change detection and color discrimination tasks, with simulations suggesting that the same underlying mechanism could produce developmental increases in both VWM capacity and resolution.

To account for developmental improvements in VWM capacity, Simmering and colleagues (Simmering, *in press*; Simmering et al., 2015) proposed a variant of the spatial precision hypothesis that has previously explained developmental changes in spatial cognition (see Simmering & Schutte, 2015, for review). In particular, Simmering (*in press*) showed that strengthening connectivity (local excitation and lateral inhibition) within the dynamic model increased *real-time stability* in VWM over development, which produced increases in capacity among other changes. In contrast to long-term notions of developmental stability in which early behavior predicts later behavior, *real-time stability* refers to how robustly the VWM system functions in the moment of a behavioral task (see Simmering & Perone, 2013, for related

discussion). In the context of VWM, the term *real-time stability* refers to a collection of consequences that arise from strengthening connectivity: faster encoding speeds; more accurate representation of stimuli; decreased interference and decay during maintenance; increased capacity; and more reliable use of representation in service of different behaviors (see Simmering, *in press*, for further details). Importantly, all of these cognitive changes in *real-time stability* arise collectively through a common underlying change in the strength of excitatory and inhibitory connections.

Figure 2 shows a simulation to illustrate how the dynamic model performance of a trial in the change detection task differs with weaker parameters to capture early development (Simmering, *in press*). The trial followed the same progression as in Fig. 1, with three items from the memory array creating input-driven peaks in the perceptual field (Fig. 2a), then forming peaks in working memory (Fig. 2b) that are maintained throughout the delay (Fig. 2c). Compared to the simulation shown in Fig. 1, the simulation with the "child" parameters had weaker peaks in both the perceptual and working memory fields (reflecting weaker excitation) and shallower inhibition (reflecting both weaker inhibitory projections as well as lower activation in the Inhibitory Layer, which results from less activation projected from the other layers). Because inhibition plays a central role in comparison to detect new items (as described above), the "child" parameters result in more missed changes, as illustrated in Fig. 2d. In this trial, activation associated with the new item in the test array did not surpass threshold in the perceptual field (see circle in Fig. 2d) and therefore did not signal a change.

Although this simulation illustrates how weaker connectivity can lead to missed changes, this is not the sole mechanism by which performance changes over development in the dynamic model; the number of items that can be held in memory increases with stronger connectivity. Simmering and colleagues quantitatively fit change detection performance in 3- to 5-year-olds (Simmering, *in press*) and 7-year-olds (Simmering et al., 2015), demonstrating how strengthening connectivity can produce the types of developmental improvements seen in children's performance. Simmering (*in press*) reported that the number of items (peaks) that could be maintained in the model increased over development: in simulation of set size five, the parameters tuned for 3-, 4-, and 5-year-olds' performance maintained an average of 3.23, 3.93, and 4.38 peaks on these trials, respectively. Simmering et al. (2015) did not report the number of peaks from their simulations, but analysis of their simulation results show an average of 4.84 peaks in set size five when fitting 7-year-olds' performance. Johnson et al. (2014) simulated adults' performance in set size six, and showed that the "adult" parameters maintained an average of 5.79 peaks. Thus,

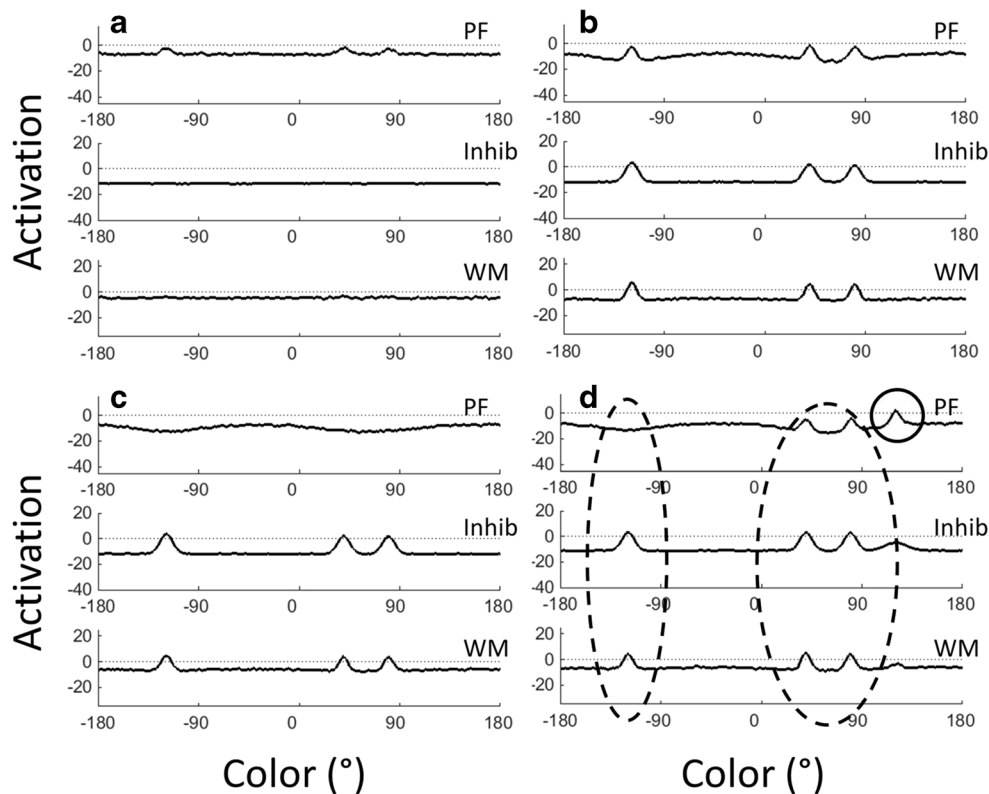


Fig. 2 Dynamic model simulation of a change detection trial with parameters tuned to children’s performance. Panels and lines are as in Fig. 1

strengthening connectivity over development quantitatively fit children’s and adults’ change detection performance through not only an increase in capacity, but also an improved ability to detect changes.

Using a similar model architecture and developmental change,³ Simmering and Patterson (2012) predicted developmental improvements in VWM resolution. In particular, they tested model’s detection of small changes in color after a brief delay in a color discrimination task (modeled after position discrimination tasks, e.g., Simmering & Spencer, 2008). As in the change detection task, detecting changes in this color discrimination task in the model depended on balance between excitation from the in-coming input and the spread of inhibition in the perceptual field. Their simulations showed that weaker connectivity in the “child” parameters led to shallow, inconsistent inhibitory troughs in the perceptual field. This led

³ The simulations presented by Simmering and Patterson (2012) used the architecture and parameters from Simmering (2008), which was a precursor to the model in Simmering (in press). Importantly, the slightly different model implementations between Simmering (2008) and Simmering (in press) were both able to quantitatively fit the same data set of children’s and adults’ performance in two VWM tasks. The ability to fit the same data set with variants of the model indicates that the precise parameter settings are not vital to the success of the model. Rather, the qualitative dynamics of the interactions that are produced by these similar architectures embody the central concepts of the theory that lead to the predictions generated by this theoretical perspective.

the model to require larger changes to detect a difference, and also produced variability in responding across repeated presentations of the same stimuli. In their behavioral tasks, 4- to 6-year-old children showed a similar developmental progression, with younger children requiring larger differences in color to reliably report differences, as well as more variability in young children’s response across runs within the stair-casing procedure.

The computational implementation of the theory allows for explorations of concepts like capacity, precision, and resolution in ways that are more difficult in solely conceptual theories. In particular, Simmering and Patterson’s (2012) model simulations indicated that improved detection of small changes over development arose not through more precise representations (i.e., narrower peaks), but rather through the related inhibitory component that provides the mechanism of comparison between memory and the perceptual input. Thus, improved resolution in how memory is used can produce a behavioral pattern that might be assumed to reflect memory precision (e.g., Machizawa & Driver, 2011). This computational implementation also suggests an explicit link between developmental increases in capacity and resolution of memory. In particular, as connectivity strengthens over development, this allows for more robust use of the VWM system to represent more items in memory at once (increased capacity) and use representations more reliably to detect small changes when a

single item is held in VWM (increased resolution). If these changes indeed share a common underlying source, then children's performance should correlate across tasks assessing capacity and resolution. We tested our prediction by comparing 5- to 8-year-old children's performance across tasks that had previously been used to measure these constructs in separate groups of children (Simmering, 2012; Simmering & Patterson, 2012).

Although this prediction arose from the account of change detection performance and developmental change put forth by Simmering and colleagues, it is not necessarily incompatible with other theoretical perspectives; rather, we contend that most alternative theories are not sufficiently specific to generate this prediction. As noted above, most theories of development have focused on only one characteristic of VWM representations (e.g., capacity, Cowan et al., 2005; Isbell et al., 2015; Kharitonova et al., 2015; Riggs et al., 2006, 2011; or precision, Burnett Heyes et al., 2012) and have not considered how changes in one characteristic might impact performance across behavioral tasks. For example, although Burnett Heyes et al. showed that precision improved developmentally when estimating the orientation of a bar held in memory alone or with two other items, they did not discuss how these changes in precision would affect performance in the change detection task used to estimate capacity. Through the computational specificity of the dynamic model, on the other hand, we can assess how the same change in underlying processes leads to different behavioral effects across tasks designed to measure capacity and resolution. The simulations we presented above indicate that the ability to detect changes is central to both types of tasks, although prior theories have posited separate processes for detecting small versus large changes (see Wilken & Ma, 2004). We test this prediction in our experiment by comparing resolution to capacity, as well as overall change detection performance, and hits versus correct rejections in change detection. As our results show, resolution was more strongly related to capacity than age in our sample, suggesting that these two characteristics of VWM improve together over development in a manner more specific than is captured by chronological age.

Empirical test of predictions

Method

Participants A total of 79 children participated; parents reported normal visual acuity and no history of colorblindness. Of these participants, eight were excluded from all analyses for the following reasons: incomplete data due to ending early (three) or experimenter error (one); not understanding/complying with instructions (two); equipment failure (one); or a capacity estimate at ceiling (one; described below). The

remaining 71 children were included for initial analyses, but six were identified as outliers on one or both measures (described further below), leading to a final sample of 65 children (M age = 6.56 years, SD = 0.72, range = 5.30–8.46 years; 29 girls) for the primary analyses. Children were recruited through a database of families interested in research participation maintained by a university-affiliated research center. Informed consent was obtained from parents before children participated.

Apparatus Tasks were presented on a 15.4-in widescreen Dell Latitude E6500 laptop computer in a room with dim overhead lighting. Stimulus presentation was controlled by Matlab using the Psychophysics Toolbox extension (version 3, Kleiner, Brainard, & Pelli, 2007). Color discrimination stimuli (mittens) appeared within 6.35 cm wide \times 7.62 cm tall white rectangles (approximately $6^\circ \times 7.27^\circ$ visual angle at a viewing distance of 60 cm) against a gray background (RGB values = 200, 200, 200); the interior colors of the mitten stimuli were determined according to a stair-casing procedure (described below). Mitten stimuli were centered vertically on the screen and presented on the left half of the screen for trials to Target 1 and the right half of the screen for trials to Target 2 (see Fig. 3). On each trial, the first mitten appeared in the left position and the second appeared on the right position, within the respective half of the monitor.

Change detection stimulus arrays consisted of one to five colored squares (2.54 cm or 2.43° on a side) presented at random locations on a gray background (RGB = 150, 150, 150). Colors were drawn randomly without replacement from the set of eight colors used in Simmering (2012), shown in Fig. 3b: black (RGB = 0, 0, 0), blue (0, 0, 255), cyan (0, 255, 255), green (0, 255, 0), red (255, 0, 0), violet (238, 130, 238), white (255, 255, 255), and yellow (255, 255, 0).

Procedure Participants completed exact replications of Simmering and Patterson's (2012, Experiment 2) color discrimination task followed by Simmering's (2012; replication of Riggs et al., 2006) change detection task. The discrimination task was presented first to ensure that children would attend to small (within-category) changes in color, rather than only to the large (categorical) changes used in the change detection task. The total duration of participation was 20–30 min.

Experimenters explained the color discrimination task to children as a matching game in which they had to find mittens that were the exact same colors. Flashcards showing mittens that either matched exactly or had small differences in color supported the experimenters' explanation of the task as they presented four demonstration trials (two in which the colors matched, two in which the colors differed). Once children understood the task, the experimenter began the test trials on the laptop. On each trial, the mittens were shown for 500 ms

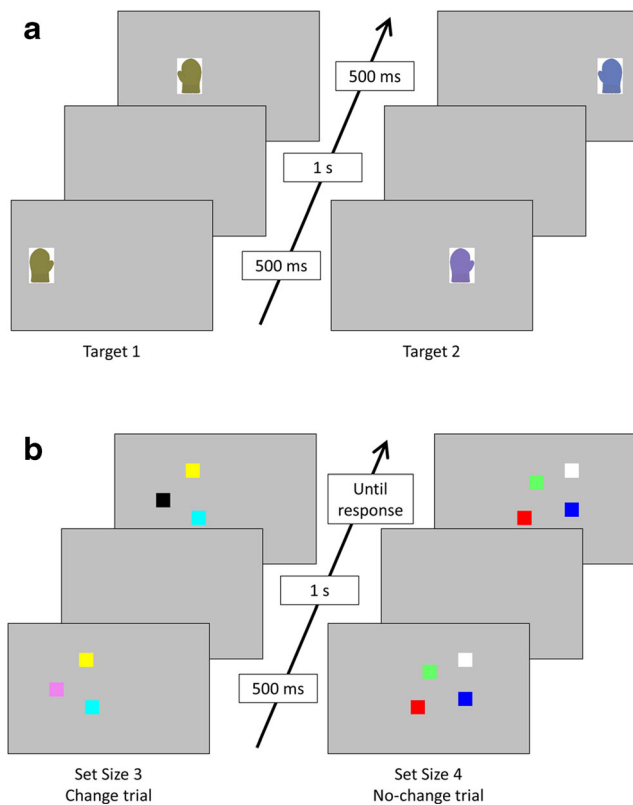


Fig. 3 Sample trials from (a) color discrimination and (b) change detection tasks. Target 1 in color discrimination shows no change in color; Target 2 shows a 16° change. Stimuli are not drawn precisely to scale

each, separated by a 1-s delay (see Fig. 3a). Children verbally responded “same” or “different” following presentation of the second mitten, and the experimenter entered the response on a keypad. Experimenters initiated each trial with a keypress and could repeat a trial if the child was not looking; this occurred at least once for 21 children (33 % of sample), and more than once for ten of those 21 children (15 % of total sample). Children did not receive feedback on the accuracy of their responses, although general positive encouragement was provided by the experimenter throughout the task.

Across trials, two target colors (136° [RGB = 136, 132, 20] and 356° [RGB = 131, 115, 189] in the 360° color space from Johnson, Spencer, Luck, et al., 2009; Simmering & Patterson, 2012, Experiment 2) were presented as the first mitten within a trial, and the color of the second mitten was selected according to a stair-casing procedure in alternating ascending and descending runs across trials. We chose a stair-casing design, rather than the method of constant stimuli, to reduce the total number of trial required from each participant, thus increasing the likelihood that children would be able to complete both tasks required for our current design (see Simmering & Patterson, 2012, for further discussion). Trial runs using Target 1 were presented on the left side

of the computer monitor, and trial runs using Target 2 were presented on the right side of the computer monitor; which color served as Target 1 versus 2 was counterbalanced across participants (Fig. 3 shows 136° as Target 1 and 356° as Target 2; note that colors may appear different in print or across different computer monitors).

Participants completed six runs per target, alternating between Targets 1 and 2. The first run to each target began with the mitten colors matching exactly; following each “same” response from the participant, the next trial would present a color one step (8°) in color space more different from the target color (ascending runs). Following a “different” response, ascending runs terminated, and the next run to that target presented a separate one step larger; the second run to a target would continue as an ascending run if the participant responded “same,” or as a descending run (presenting separations one step smaller) if the participant responded “different.” Runs continued in this fashion, alternating across target colors, until the participant completed twelve runs (six per target; see Appendix A for sample sessions illustrating the stair-casing procedure). The total number of trials completed by each participant varied according to their responses due to the stair-casing design; for participants who were not excluded as outliers in this task (as described below), the total number of trials ranged from 27 to 69 ($M = 38.31$, $SD = 9.49$).

Following a short break after completion of the color discrimination task, the experimenter explained the change detection task as a new matching game in which children would see multiple colors and have to identify whether the colors matched across sets. Children were shown a page with two sample trials in set size three, one in which one color changed and one in which the colors remained the same. The experimenter emphasized that if any of the colors changed between the first and second display, the child should respond “different.” Once children understood how to play the game, the experimenter began the computerized task. Children first completed a practice block with six trials, including one change trial and one no change trial each in set sizes one, two and three. For practice trials, the memory array was presented for 2 s, followed by a 1-s delay, then the test array, which remained visible until the child verbally responded “same” or “different” and the experimenter entered the response on a keypad. Following the practice block, children completed five test blocks corresponding to set sizes one through five (see Fig. 3B for examples of set sizes three and four) with 12 trials each for a total of 60 test trials (half change, half no-change). Set size order was randomized for each participant. For test trials, the duration of the memory array was decreased to 500 ms (Riggs et al., 2006; Simmering, 2012).

Throughout the change detection task, incorrect responses were followed by a buzzer sound.

Method of analysis For the color discrimination task, performance was analyzed following Simmering and Patterson (2012, Experiment 2) by taking the smallest color separation at which the participant responded “different” within each of the six runs per target, then computing the mean separation across runs as the participant’s threshold of “just noticeable difference” (JND); for ease of interpretation, we calculated these as difference scores, that is, negative values with scores closer to zero indicating better performance. JND estimates were averaged across targets to arrive at one value per participant. Four children were identified as outliers due to their JND being more than 2 SD below the group mean (shown in Fig. 4 and described further below).

For the change detection task, participants’ responses were classified as correct rejections (correct no-change trials), hits (correct change trials), misses (incorrect change trials), or false alarms (incorrect no-change trials). Performance was analyzed by estimating capacity (K) for each child using Pashler’s (1988) formula, $K = SS * (H - FA)/(1 - FA)$, based on the hit (H) and false alarm (FA) rates for each set size (SS). Because K cannot exceed the set size tested, this value may underestimate capacity when small set sizes are included. There have been a number of approaches to overcoming this limitation to arrive at a single capacity estimate per participant (see Appendix B for comparison of multiple methods). We chose to average multiple estimates per child, but exclude those likely to be limited by set size based on the each individual child’s performance. In particular, we ordered each child’s estimates by increasing set size (one through five; note that this was not the order in which set sizes were completed, as described above), then averaged all estimates beginning with first set size in which $K < SS$ for that child. For example, if a child had K estimates of 1.00, 2.00, 2.50, 2.40, and 2.50, respectively across set sizes one through five, only estimates from set sizes three through five (2.50, 2.40, and 2.50) would be averaged (resulting in $K = 2.47$), as $K = SS$ in set sizes one and two (see Appendix B for further details). We chose this method because it was more sensitive to individual differences in performance by including estimates from multiple set sizes, but excluded low estimates that may have been limited by set size. Using this method led to the exclusion of one child for whom $K = SS$ across all five set sizes (noted above as having capacity estimates at ceiling). Additionally, three children were identified as outliers due to their mean thresholds being more than 2 SD below the group mean (shown in Fig. 4A and described further below).

Results

We hypothesized that a shared mechanism of developmental change underlies both increases in resolution (JND) and capacity (K). To test this, we calculated Pearson’s correlations between K (including outliers $M = 2.03$, $SD = 0.83$; excluding outliers $M = 2.13$, $SD = 0.71$) and JND (including outliers $M = -18.73$, $SD = 10.96$; excluding outliers $M = -16.61$, $SD = 6.35$), as well as the correlations between each of these factors and age. We first computed these correlations without excluding outliers; Fig. 4A shows the full set of 71 participants’ data, which revealed a positive correlation between capacity and resolution, $r_{69} = .359$, $p = .001$ (all reported p -values are one-tailed). This figure clearly shows a few cases of poor performance on each measure, which may have influenced the overall correlation. This led us to exclude data points that were more than two standard deviations below the mean on either measure (i.e., K below 0.36, JND below -40.64); these outlier values are indicated by circles in Fig. 4A. After removing these six data points, the positive correlation between capacity and resolution remained significant (diamonds in Fig. 4A), $r_{63} = .356$, $p = .002$, supporting our prediction. Note that both capacity and resolution were not significantly correlated with age ($r_{63} = .150$, $p = .117$, and $r_{63} = .110$, $p = .191$ respectively), although there were small positive relations among these measures, as shown in Fig. 4B. Even though these correlations were not statistically significant, we followed up with a partial correlation controlling for age, to ensure that contributions from age were not driving our overall correlation: again, the positive correlation between capacity and resolution remained significant, $r_{62} = .345$, $p = .003$. Thus, the relation between children’s resolution and capacity does not depend only on the variance accounted for by chronological age. This is consistent with prior demonstrations showing relatively gradual developmental improvements in resolution and capacity in this age range (Simmering, 2012; Simmering & Patterson, 2012) and indicates that age is only a proxy for the maturity of VWM.

One potential limitation of these analyses is that the K formula assumes a slot-like representation of VWM. To avoid relying solely on the assumptions underlying this formula to support our results, we also computed the correlation between resolution and overall proportion correct in the change detection task ($M = .85$, $SD = .053$), which revealed a similar result, $r_{63} = .225$, $p = .036$. Furthermore, because our model simulations suggest that detection of changes is the central process that relates performance across tasks, we analyzed the correlation between resolution and overall hit rate in change detection ($M = .79$, $SD = .093$); this analysis revealed the predicted positive correlation, $r_{63} = .327$, $p = .004$. By contrast, resolution was not significantly correlated with overall correct

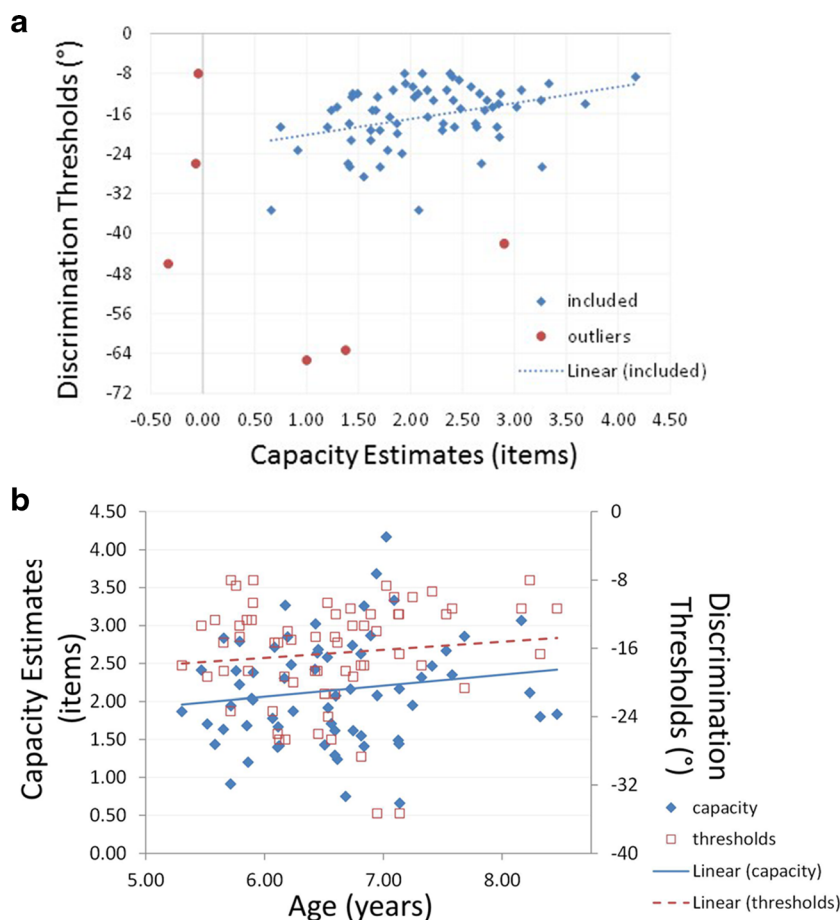


Fig. 4 Correlations between (a) capacity and resolution, (b) capacity and age on the left axis, and resolution and age on the right axis

rejections in change detection ($M = .91, SD = .073$),⁴ $r_{63} = .083, p = .256$. The different relations between resolution and hits versus correct rejections suggests our effect is not driven solely by children being generally good at memory tasks, but rather due to the shared underlying process of detecting changes, as predicted by model simulations.

Discussion

The goal of this paper was to test a dynamic account of VWM capacity limits, resolution, and development. Simmering and colleagues previously showed that this model could capture developmental improvements in both the resolution and capacity of VWM during early to middle childhood (Simmering, *in press*; Simmering et al., 2015; Simmering & Patterson, 2012). The current paper tested the prediction that behavioral estimates of

these characteristics should be correlated in children. Results supported this prediction with a significant correlation between resolution (as estimated from a color discrimination task) and capacity (as estimated in the change detection task), which was driven primarily by the detection of changes in both tasks. Through formalizing VWM performance in a process-based model, we have shown how VWM resolution and capacity estimates relate to the underlying process of comparing items held in memory to new inputs. These results provide further support for the dynamic account of VWM limits and changes in real-time stability as a mechanism of developmental improvements in VWM.

Although these results support the prediction derived from the dynamic model, a number of questions remain, which we consider in turn here. First, we compare other theoretical accounts of capacity limits and development to our proposed mechanism. Second, we discuss how the dynamic model would address tasks used to estimate the precision of VWM, and how precision and resolution might relate in the model and in visual cognition. Third, we consider how co-developing processes might contribute to memory task performance. Lastly, we conclude by arguing for increased

⁴ Eight children scored at ceiling on correct rejections across set sizes, which may have truncated variance in this measure; after excluding these children’s data ($M = .89, SD = .068$), the correlation remained non-significant, $r_{55} = .094, p = .243$.

specificity in theories to clarify the correspondence between representations, cognitive processes, and behavior.

Comparison with alternative theories of VWM capacity and development

The prediction derived from the developmental change put forth by Simmering and colleagues is not the only potential explanation for changes in resolution or capacity over development. It is not clear whether other theories could accommodate our results with some modification, but at the current level of specificity, no other theory would predict the correlation presented here. Most theories have addressed only increases in capacity (e.g., Cowan et al., 2005; Isbell et al., 2015; Kharitonova et al., 2015; Riggs et al., 2006, 2011) or precision (Burnett Heyes et al., 2012) as potential developmental changes, without considering whether or how the two relate developmentally. These theories are, at the least, under-specified to address our prediction, and at most directly inconsistent with our results. For example, theories of the source of capacity limits in adults have argued for limits only in the number (e.g., Zhang & Luck, 2008) or resolution (e.g., Wilken & Ma, 2004) of items in VWM, which would preclude predicting a systematic relationship between these two characteristics of VWM over development. More recent accounts that consider limits in both the number and resolution of items propose that representational resolution varies inversely with the number of items represented (with the possibility of random fluctuations across items and trials), without specifying the processes that lead to variation in resolution (e.g., van den Berg et al., 2014). These accounts may be able to explain how resolution relates to the number of items on a given trial (cf. Burnett Heyes et al., 2012), but would not predict that the resolution of a single item (as tested in our color discrimination task) relates systematically to the number of items represented in a different task (change detection).

In contrast to the dominant theories of VWM capacity and development, the dynamic model provides a specific mechanistic explanation of how local excitation and lateral inhibition contribute to limits on both the number and resolution of representations, as well as how strengthening connectivity over development leads to improvements in both capacity and resolution. Specifically, the role of inhibition as a mechanism for comparison between items held in memory and the perceptual input from the test array led to the prediction that the ability to detect small changes (resolution) corresponds to the ability to detect a change in one of multiple items (hit rate and capacity). Thus, at present, the dynamic model presented here provides the only specific account of how behavior should relate across these tasks over development.

The dynamic model architecture has not only quantitatively simulated children's and adults' behavior in the standard change detection task (Johnson et al., 2014; Simmering,

in press; Simmering et al., 2015), but also generated novel predictions regarding the effects of familiarity (Simmering et al., 2015) and similarity of items in memory (Johnson, 2008; Johnson et al., 2014; Johnson, Spencer, Luck, et al., 2009). Beyond the change detection task, this modeling framework has also been used to simulate infants' and children's performance in a looking task designed to assess VWM capacity (Perone, Simmering, & Spencer, 2011; Simmering, in press). This computational work bridging tasks and developmental periods provides an opportunity to examine how other behaviors relate to VWM processes. One of the primary advantages of implementing theories computationally is the ability to specify the manner in which processes interact in the context of specific tasks (Simmering & Patterson, 2012), which provides an opportunity to integrate findings across previously disconnected domains (Simmering, in press; Simmering & Spencer, 2008).

VWM precision

Further research is needed to test whether the dynamic model could also capture performance in tasks used to estimate VWM precision, and how behavioral estimates of precision and resolution relate in VWM over development. Studies with a focus on the precision of VWM typically use delayed estimation tasks with precision estimated from the variance in responses across trials. At present, the delayed estimation tasks have not been used with children younger than 7 years, which is one reason we chose not to include it in the current study. This is partially due the task requiring a relatively large number of trials (cf. 145 trials in Burnett Heyes et al., 2012; or 150 trials in Sarigiannidis et al., 2016), which is difficult for young children to complete, especially with other tasks as part of the same study. An important goal for future research is to incorporate performance on this task into a more comprehensive theory of VWM performance and development, which will require both empirical and theoretical work.

Would the dominant theoretical perspectives predict relations among precision (as estimated by the delay-estimation task), resolution (as estimated in discrimination tasks), and capacity (as estimated by change detection)? For the most part, slot and resource theories do not specifically address this question, partly because the notion of resolution we endorse here (using a representation to detect a change) is not measured using estimation tasks. Thus, it is possible that resource accounts would predict that resolution in detecting small changes would be directly related to precision of estimation responses, but theories thus far have not been specified to address this question. Furthermore, although resource accounts have specifically proposed that apparent capacity limits in change detection are due to load-related reductions in precision rather than a limited number of representations, it is not clear how strongly individual differences in single-item

precision should relate to load-related changes in precision (i.e., change detection performance) from this perspective.

Addressing the potential link between precision and resolution in the dynamic model would require further computational development to quantitatively fit performance in the delayed estimation task. Simulations thus far indicate that the precision of a representation (i.e., peak width) is not isomorphic with the resolution (i.e., size of change that can be reliably detected), suggesting a dissociation in behavioral estimates of these characteristics of memory. However, a primary lesson to be learned from the dynamic model implementation is that performance is multiply determined, and behavioral responses are not direct reflections of the underlying representations. Therefore, it is possible that the variability of estimation responses used as an index of VWM precision reflect more than simply the nature of the representation.

Initial evidence supporting this intuition can be found in simulations of children's and adults' performance in spatial recall tasks, which have a structure comparable to single-item color and orientation estimation tasks. Spencer and colleagues have quantitatively simulated children's (e.g., Schutte & Spencer, 2009) and adults' (e.g., Lipinski, Simmering, Johnson, & Spencer, 2010) spatial recall performance in a related dynamic model. On each trial in these tasks, a single spatial position was presented as a target within a homogenous display (e.g., a sandbox or computer screen), then following a 10-s delay the participant pointed to the remembered target location. Results have shown variability in responses to the same target location across repeated trials within individuals (Schutte & Spencer, 2009). To simulate these tasks, Spencer and colleagues generated "pointing" responses from the model by reading out the position of the peak along the spatial dimension at the end of the delay. Critically, model simulations showed variability in responses from trial to trial that correspond to the variability shown by children and adults. In the model, this variability across trials resulted from peaks "drifting" along the spatial dimension by differing amounts. Furthermore, developmental simulations showed that strengthening connectivity produced a reduction in variability, which could be interpreted as an increase in precision over development (cf. Burnett Heyes et al., 2012). Comparison of peaks in the working memory field of the spatial simulations showed some narrowing of these representations over development,⁵ but not enough to account for the behavioral

differences. Rather, reduction in behavioral variance also depended on less drift in peaks across trials, suggesting that the typical index used to estimate precision (i.e., variability in responding) might not solely reflect the precision of individual memory representations.

To generalize these simulation results to the color or orientation recall tasks, the position of the peak along the color or orientation dimension in the model could be read out to generate the recall response (e.g., Johnson, 2008; Johnson, Ambrose, Dineva, & Spencer, 2016). Variability in responses across trials (as seen in the spatial recall simulations) is interpreted as reflecting the precision of the underlying color or orientation in the delayed-estimation tasks (Bays & Husain, 2008). However, simulations of spatial recall suggest that such variability in behavior could arise through shifts in the position of peaks along relevant dimension alone, not necessarily the width (i.e., precision) of the peak. Further model simulations will be needed to verify that such an effect can be captured quantitatively (see Johnson & Simmering, 2015, for further discussion), but conceptually the two behavioral assessments of the precision or resolution of memory representations—estimation responses or "same"/"different" judgments—could correspond in this architecture without reflecting solely the width of peaks held in the working memory field. The central argument arising from this review of model simulations and the current behavioral data is that theories focusing exclusively on representations provide an incomplete account of behavior, which can result in potential misattribution of behavioral effects to characteristics of memory.

Related cognitive processes

An important contrast between the dynamic model and classic information processing explanations of memory is the continuity of processes within a task. Early descriptions of memory systems posited separation of encoding, maintenance, and retrieval processes (e.g., Atkinson & Shiffrin, 1968), but in the dynamic model these processes are identical: the same excitatory and inhibitory connections support these "stages" in the task, and only differences in the presence or absence of input from the task structure differentiate encoding from maintenance of items (see Simmering, *in press*, for evidence that task structure can modulate capacity). This raises the question of how other cognitive characteristics that have classically been considered separable from memory storage, such as attention or processing speed, could emerge within the dynamic model architecture. To preview, behavioral effects attributed to attention and processing speed could arise through the continuous processes implemented in the dynamic model, and the apparent dissociation from memory could reflect changes in the structure of the tasks rather than reliance on different cognitive abilities.

⁵ Simulations of spatial recall also include peaks corresponding to perceptual reference frames in the perceptual field; over development, such reference peaks shift from an excitatory influence to an inhibitory influence as connectivity strengthens. Reference-related inhibition contributes to delay-dependent drift, and also serves to narrow peaks as they are maintained in the working memory field (see Simmering & Schutte, 2015, for further details). No analogous reference is present in color memory tasks, which would eliminate this source of narrowing from simulations of color estimation tasks.

We first consider the relations among cognitive processes in the context of infant visual recognition. Rose, Feldman, and Jankowski (2002) developed a “processing speed” task in which infants view a series of paired images, one that remains the same across trials and one that changes. Trials continued until the infant showed a robust novelty preference, and the number of trials to criterion can be used as an individual metric of processing speed, or the amount of accumulated experience needed to form a robust memory representation of the familiar item. This task also affords a second measure, purported to measure attention: shift rate, which is the frequency with which infants shift gaze between the two images on each trial. Across a series of studies, Rose and colleagues have demonstrated both individual and developmental differences in processing speed and shift rate (see Rose, Feldman, & Jankowski, 2007, for review).

A variant of the dynamic model presented here has been used to simulate infants’ performance in the processing speed task (Perone & Spencer, 2013, 2014). These simulations

suggest that individual and developmental differences could be realized through the same type of mechanism. They presented a series of long-term learning simulations in which “infant” models gained experience in a simulated visual world; the model was equipped with a simple Hebbian learning process by which above-threshold activation in the perceptual and working memory fields left excitatory traces at the corresponding feature values. These excitatory traces built up slowly with experience, which allowed for distribution across the feature dimension through varying experiences. Following this autonomous development, Perone and Spencer (2013) tested these individual “infants” (i.e., models with individual developmental histories) in the processing speed task. These simulations showed that this accumulation of experience provided the typical developmental improvements: “Older” models with more experience encoded items in to memory more quickly and showed faster shift rates (cf. parameter tuning by Perone & Spencer, 2014). Comparing within a given “age” group also revealed robust individual differences, with slower encoding at the simulated equivalent of 7 months predicting slower encoding in the same models at the equivalent of 12 months. These simulation results provide compelling evidence that individual and developmental differences in measures of processing speed and attention (shift rate) could arise through accumulated experience within an integrated memory system.

Results from the Perone and Spencer (2013, 2014) simulations call into question classic notions of separability between cognitive processes, and instead highlight the need to understand how behavior arises from the underlying processes within specific task contexts. A similar demonstration can be found in Simmering’s (*in press*) comparison of 3- to 5-year-old children’s performance between the change detection task and an infant change-preference task designed to assess capacity (Ross-Sheehy, Oakes, & Luck, 2003). Behavioral results showed a developmental increase in shift rate in the change-preference task that paralleled increases in capacity estimates from the change detection task. Model simulations indicated that these behaviors related through their reliance on the processes supporting the formation, maintenance, and comparison of items within the tasks. In particular, fixation in the model is linked to perceptual processing, such that activation within the perceptual field supports continued fixation. As memory representations build in the working memory field, activation in the perceptual field is suppressed (as described above and shown in Figs. 1 and 2, this is the recognition mechanism in the dynamic model), and fixation is released. Thus, in the model, an increased shift rate reflects faster encoding and recognition of items in memory, not just attention. Simmering’s behavioral results supported this link, showing that individual differences in shift rate in the looking

Table 1 Trial sequence in the color discrimination task for participant 1

Trial no.	Run/Target no.	Run direction	Mitten 1 color	Mitten 2 color	Response
1	R1/T1	Ascending	136	136	Same
2	R1/T1	Ascending	136	144	Same
3	R1/T1	Ascending	136	152	Different
4	R1/T2	Ascending	356	356	Same
5	R1/T2	Ascending	356	4	Same
6	R1/T2	Ascending	356	12	Same
7	R1/T2	Ascending	356	20	Different
8	R2/T1	Descending	136	160	Different
9	R2/T1	Descending	136	152	Different
10	R2/T1	Descending	136	144	Same
11	R2/T2	Descending	356	28	Different
12	R2/T2	Descending	356	20	Different
13	R2/T2	Descending	356	12	Same
14	R3/T1	Ascending	136	136	Same
15	R3/T1	Ascending	136	144	Different
16	R3/T2	Descending	356	4	Different
17	R3/T2	Descending	356	356	Same
18	R4/T1	Descending	136	152	Different
19	R4/T1	Descending	136	144	Same
20	R4/T2	Ascending	356	356	Same
21	R4/T2	Ascending	356	4	Same
22	R4/T2	Ascending	356	12	Same
23	R4/T2	Ascending	356	20	Different
24	R5/T1	Ascending	136	136	Same
25	R5/T1	Ascending	136	144	Different
26	R5/T2	Ascending	356	28	Same
27	R5/T2	Ascending	356	36	Different
28	R6/T1	Descending	136	152	Different
29	R6/T1	Descending	136	144	Different
30	R6/T1	Descending	136	136	Same
31	R6/T2	Descending	356	44	Different
32	R6/T2	Descending	356	36	Different
33	R6/T2	Descending	356	28	Same

Note. Mitten colors are reported in degrees. The first trial in each run is marked in bold for ease of reading

Table 2 Trial sequence in the color discrimination task for participant 2

Trial no.	Run/ Target no.	Run direction	Mitten 1 Color	Mitten 2 Color	Response
1	R1/T1	Ascending	356	356	Same
2	R1/T1	Ascending	356	4	Same
3	R1/T1	Ascending	356	12	Different
4	R1/T2	Ascending	136	136	Same
5	R1/T2	Ascending	136	144	Different
6	R2/T1	Descending	356	20	Different
7	R2/T1	Descending	356	12	Different
8	R2/T1	Descending	356	4	Different
9	R2/T1	Descending	356	356	Same
10	R2/T2	Descending	136	152	Different
11	R2/T2	Descending	136	144	Same
12	R3/T1	Descending	356	356	Different
13	R3/T1	Descending	356	356	Same
14	R3/T2	Ascending	136	136	Same
15	R3/T2	Ascending	136	144	Same
16	R3/T2	Ascending	136	152	Different
17	R4/T1	Ascending	356	356	Repeat (Didn't See)
18	R4/T1	Ascending	356	356	Same
19	R4/T1	Ascending	356	4	Same
20	R4/T1	Ascending	356	12	Same
21	R4/T1	Ascending	356	20	Different
22	R4/T2	Descending	136	160	Different
23	R4/T2	Descending	136	152	Different
24	R4/T2	Descending	136	144	Same
25	R5/T1	Descending	356	28	Different
26	R5/T1	Descending	356	20	Different
27	R5/T1	Descending	356	12	Different
28	R5/T1	Descending	356	4	Different
29	R5/T1	Descending	356	356	Different
30	R5/T1	Descending	356	356	Repeat (Didn't See)
31	R5/T1	Descending	356	356	Same
32	R5/T2	Ascending	136	136	Same
33	R5/T2	Ascending	136	144	Same
34	R5/T2	Ascending	136	152	Same
35	R5/T2	Ascending	136	160	Different
36	R6/T1	Descending	356	356	Different
37	R6/T1	Descending	356	356	Same
38	R6/T2	Descending	136	168	Different
39	R6/T2	Descending	136	160	Different
40	R6/T2	Descending	136	152	Different
41	R6/T2	Descending	136	144	Different
42	R6/T2	Descending	136	136	Same

Note. Mitten colors are reported in degrees. The first trial in each run is marked in bold for ease of reading

task correlated with capacity estimates from the change detection task (after controlling for age).

Beyond these types of looking tasks, the dynamic model architecture has been used to simulate two other notions of attention over development. First, in a larger dynamic neural field architecture, Buss and Spencer (2014) simulated children’s performance in the dimensional change card sort task, which requires children to first sort a set of simple colored shapes by one dimension (e.g., by shape) then switch to sort by the other dimension (e.g., color). Buss and Spencer implemented the sorting rule as a boost to the relevant dimension (i.e., increasing the resting level of the shape or color field in the model). To simulate both individual and developmental differences in this task, they varied the amount of this boost. A similar mechanism could be incorporated into the architecture presented here, and future simulations could test whether a volitional modulation of resting levels could capture attentional control in this task context. A follow-up study showed that attention to the relevant dimension could be supported in young children by providing additional experience along that dimension (i.e., playing a memory game with colors before sorting by color; Perone, Molitar, Buss, Spencer, & Samuelson, 2015). These empirical and computational results converge with the infant simulations to illustrate the tight link between memory and attentional processes.

Additionally, using the spatial memory architecture described in the previous section, Johnson and Spencer (2016) simulated behavioral results that were previously proposed to support the notion that spatial attention serves as a rehearsal mechanism for spatial memory (Awh, Jonides, & Reuter-Lorenz, 1998). In this task, participants were required to make a color discrimination judgment during the memory delay of a spatial recall trial. Awh et al. (1998) showed that spatial recall was impaired when such judgments were required, and interpreted the effect as a withdrawal of attention from the remembered location. Using a dynamic model architecture, Johnson and Spencer showed that spatial memory was not merely worse following an intervening task, but that the position in which the color stimulus appeared interacted with the item held in memory. In particular, recall errors were systematically biased toward the location of the color stimulus, through the continuous maintenance processes supporting memory during the delay. This effect was achieved in a model that had no specific “attention” component, but rather emerged as a consequence of the processes involving perceiving, remembering, and responding to stimuli in the context of the specific task.

As this series of empirical and simulation results illustrates, implementing behavioral tasks into a process-based model allows for careful analysis of the underlying processes necessary to produce specific patterns of behavior. This provides the opportunity to discover cases in which

the behavioral signatures attributed to one process (e.g., shift rate as a measure of attention) arise through other processes required by the task (e.g., encoding and comparison), and can explain why and how behavior relates across different tasks. Together with our empirical results indicating a shared source underlying developmental improvements in VWM resolution and capacity, this theoretical perspective suggests a potential synthesis of prior hypotheses to arrive at a more general explanation for developmental change (see Simmering, *in press*, for further discussion).

Conclusion

The results we present here provide evidence for a shared source contributing to developmental improvements in the resolution and capacity of VWM, two characteristics that have often been presented as alternative accounts of VWM limits. The predicted correlation was derived from model simulations showing that the same underlying developmental change—strengthening connectivity to increase real-time stability of VWM—produced increases in the number of items that could be maintained, as well as a more accurate use of memory representations in service of behavior. Implementing our theory in a process-based computational model allows for dissociation between the nature of representations and how representations are used in service of behavior. This highlights the utility of computational models to formalize theoretical constructs, beyond “slots” and “resources”, and to test how the same representations might show different signatures across behavioral tasks (see also Johnson et al., 2014; Simmering, *in press*). This level of specificity produces testable hypotheses that can provide novel insights into cognitive processes, allowing for further theory refinement and synthesis of prior findings over a broad range of tasks and age groups to provide a unifying theory of the functioning and development of VWM.

Acknowledgments The authors thank the families who participated and the research assistants who aided in data collection. Participant recruitment and programming was funded by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health (R03-HD067481) and the Waisman Intellectual and Developmental Disabilities Research Center grant (P30-HD03352). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Thanks to Brad Postle for input on a previous version of this manuscript. A subset of the data included here was presented at the 70th biennial meeting of the Society for Research in Child Development in Seattle, WA, USA.

Appendix A

This section includes further information on the data produced from the color discrimination task, including data from two

sessions with children to illustrate the stair-casing procedure, as well as behavioral estimates from ten lists of randomly-generated responses. Trial sequences for the child participant are listed in Tables 1 and 2, respectively, and described in the next section. Estimates from the random responses are discussed in the second section of this appendix.

Data from child participants

Trial 1 for Participant 1 presented color 136° for the first and second mittens (see Figs. 1 and A1), and the participant correctly responded “same.” The next trial in this run presented 136° for the first mitten, and 144° for the second mitten (an ascending step of 8°) and the participant again responded “same.” The next trial presented 136° for the first mitten, and 152° for the second mitten (another ascending step of 8°) and the participant responded “different.” This ended Run 1 to Target 1, and the next trial (Trial 4) began Run 1 to Target 2, presenting 356° for the first and second mittens. The participant responded “same,” so Trial 5 presented 356° as the first mitten and 4° as the second mitten (an ascending step of 8°; 360° wraps to 0° along the color dimension). The participant again responded “same”, so the separation was increased by another ascending step to 12°; following another “same” response, Trial 6 presented the 20° for the second mitten. At this point, the participant responded “different,” terminating Run 1 to Target 2, and proceeding to Run 2 to Target 1.

To begin Run 2 to Target 1, Trial 8 presented 136° as the first mitten, but now the second mitten was 160°, a color one step larger than the color ending the previous run to this target (which was 152° in this case). When the participant responded “different,” this initiated a descending run, so the next trial presented a separation one step smaller (152°). The participant responded “same,” which terminated this run and began Run 2 to Target 2. Runs continued in this manner, alternating across targets until six runs had been completed for each target. Note that Run 1 to each target was always ascending; for all subsequent runs, the direction of change depended on whether the participant responded “same” (ascending) or “different” (descending) on the first trial within that run.

Data from randomly-generated responses

To ensure that children’s performance did not reflect random responding, we generated ten lists of “same” or “different” responses chosen randomly for each trial, then entered these responses into our experimental program. The first notable difference between our random

responses and the child participants is the number of ascending versus descending runs. The stair-casing procedure was set up to proceed with alternating directions of runs, beginning with an ascending run that terminated with a “different” response, then switching to descending run that began with a separation one step larger than the final separation tested in the prior run (see examples in Tables 1 and 2). To account for the possibility that a child may respond “different” prematurely in an ascending run, however, we added the contingency that a “same” response at the beginning of a descending run would switch the direction and continue to ascend instead. Concretely, this is shown in Table 1 at Trial 26 (Run 5 to Target 2). This feature of the stair-casing procedure could then produce different numbers of ascending versus descending runs within the six runs completed per target.

Most child participants had the expected three ascending and three descending runs to one or both targets (22 and 38 participants, respectively, for 92 % of the final sample). The remaining five participants who did not have three ascending runs to either target had two runs in one direction and four runs in the other direction (i.e., only one run per target that did not conform to the expected direction). In contrast to the real participants, our randomly-generated responses resulted in only one run in a given direction for six of these simulated participants and zero for one simulated participant. This high rate of violating the expected direction in the simulated participants did not match the real participants’ performance, which suggests that children were not responding randomly, and that unexpected responses (i.e., “same” at a larger separation than a prior “different” response) were relatively rare.

Next we calculated JND for each of the ten sessions in the same manner as for participants. This yielded a mean JND of -29.40 with a standard deviation of 16.14 , which was notably worse than the behavioral mean, even with outliers included ($M = -18.73$, $SD = 10.96$). Thus, the pattern of performance that would arise from random responding is not consistent with children’s behavior in our task.

Appendix B

Here we describe the various methods we considered for deriving a single estimate of capacity, from the five K estimates calculated from set sizes one through five, per participant. Note that all descriptions of resulting estimates include data from the children who were ultimately excluded as outliers as well as the child who was excluded due to $K = SS$ estimates in all set sizes (thus, $n = 72$ for these calculations).

The first option we considered was to take the maximum estimate across set sizes for each child (Olsson & Poom, 2005; Todd & Marois, 2005); Simmering (2012) showed that this

method of estimation at the group level (i.e., 3- vs. 5- vs. 7-year-olds) yielded similar statistical results to comparing proportion correct across age groups. However, we felt this provided only a limited index of individual differences in performance, as two children with the same maximum estimate may have different estimates from other set sizes. For example, two children in our sample had K estimates of 5.00 in set size five; in set size four, however, one child’s estimate was 4.00 whereas the other child’s estimate was 3.33. By including only the set size five estimate, the difference in performance on set size four would be masked.

Next we considered averaging across all set sizes to maximize the influence of individual differences; this method is generally avoided because it underestimates capacity by including low estimates (e.g., Vogel, Woodman, & Luck, 2001). Indeed, 63 participants (88 % of the sample) had K estimates of 1.00 in set size one, and 53 participants (74 %) had K estimates of 2.00 in set size two. Although this method includes all of the data from each participant, we felt these maxed-out estimates at low set sizes may have been artificially truncating the variance in children’s performance. This led us to consider averaging only set sizes higher than one or two; although this overcame the limitation of the low estimates, it may have excluded meaningful data from children whose estimates were not equal to the set size (i.e., $K < 1$ in set size one, $K < 2$ in set size two). The large age range we included provided ample variance in children’s memory abilities, indicating the need to include as much meaningful data per participant as possible.

Finally, we decided to use a method that included data from multiple set sizes, but excluded those likely to be below the individual participant’s capacity. We calculated this by ordering participants’ K estimates by increasing set size, then progressively eliminating estimates if $K = SS$, and averaging all K estimates from set sizes above the lowest set size at which $K < SS$. Beginning with set size one, there were ten participants (14 % of the full sample of 73) whose K estimate was less than 1; these ten participants’ K values included all five K estimates from set sizes one through five (66 trials total). Of the remaining 62 participants, 14 (19 % of the full sample) had K estimates below 2 in set size two, leading to the averaging of four K estimates (derived from 50 trials total) in their K value. In set size three, 32 of the remaining 48 (44 % of the full sample) had K estimates below 3, resulting in K values that included three K estimates (derived from 38 trials). Only one of the remaining 16 children had a K estimate of 4 in set size four—and was excluded from all analyses due to her K estimate equaling 5 in set size five—allowing the use of K estimates from set sizes four and five (24 trials total) to be averaged for the final 15 children (21 % of the full sample).

Note that our results do not depend critically on the method we chose for calculating K . For comparison, we also calculated correlations between JND and K_{mean} (i.e., including K from

all set sizes) for all participants except the one with $K = SS$ ($n = 71$), and found similar results: including outliers, $r_{69} = .367$, $p < .001$ excluding outliers, $r_{63} = .332$, $p = .003$. Furthermore, Pearson's correlation between K and K_{mean} was very high ($r = .978$). Given the similar conclusions regarding individual differences that can be drawn from these methods of calculation, we felt that K provided a more transparent estimate of capacity, as it was not limited by low estimates.

References

- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory Volume 2* (pp. 89–195). New York: Academic Press.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628.
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 780–790.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*, 851–854.
- Burnett Heyes, S., Zokaei, N., van der Staaij, I., Bays, P. M., & Husain, M. (2012). Development of visual working memory precision in childhood. *Developmental Science*, *15*(4), 528–539.
- Buss, A. T., Fox, N., Boas, D. A., & Spencer, J. P. (2013). Probing the early development of visual working memory capacity with functional near-infrared spectroscopy. *Neuroimage*. doi:10.1016/j.neuroimage.2013.05.034
- Buss, A. T., & Spencer, J. P. (2014). The emergent executive: A dynamic field theory of the development of executive function. *Monographs of the Society for Research in Child Development*, *79*(2).
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.
- Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*, 51–57.
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42–100. doi:10.1016/j.cogpsych.2004.12.001
- Fukuda, K., Vogel, E. K., Mayr, U., & Awh, E. (2010). Quantity not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin and Review*, *17*(5), 673–679.
- Isbell, E., Fukuda, K., Neville, H. J., & Vogel, E. K. (2015). Visual working memory continues to develop through adolescence. *Frontiers in Psychology*, *6*, 696.
- Johnson, J. S. (2008). *A dynamic neural field model of visual working memory and change detection*. University of Iowa.
- Johnson, J. S., Ambrose, J. P., Dineva, E., & Spencer, J. P. (2016). Neural interactions in working memory cause variable precision and similarity-based feature repulsion. *Manuscript in Preparation*.
- Johnson, J. S., & Simmering, V. R. (2015). Integrating perception and working memory in a three-layer dynamic field architecture. In G. Schöner, J. P. Spencer, & the DFT Research Group (Eds.), *Dynamic thinking: A primer on dynamic field theory* (pp. 151–168). New York: Oxford University Press.
- Johnson, J. S., Simmering, V. R., & Buss, A. T. (2014). Beyond slots and resources: Grounding cognitive concepts in neural dynamics. *Attention, Perception, & Psychophysics*, *76*(6), 1630–1654. doi:10.3758/s13414-013-0596-9
- Johnson, J. S., & Spencer, J. P. (2016). Testing a dynamic-field account of interactions between spatial attention and spatial working memory. *Attention, Perception, & Psychophysics*. doi:10.3758/s13414-015-1058-3
- Johnson, J. S., Spencer, J. P., Luck, S. J., & Schöner, G. (2009). A dynamic neural field model of visual working memory and change detection. *Psychological Science*, *20*, 568–577.
- Johnson, J. S., Spencer, J. P., & Schöner, G. (2009). A layered neural architecture for the consolidation, maintenance, and updating of representations in visual working memory. *Brain Research*, *1299*, 17–32.
- Kharitonova, M., Winter, W., & Sheridan, M. A. (2015). As working memory grows: A developmental account of neural bases of working memory capacity in 5- to 8-year old children and adults. *Journal of Cognitive Neuroscience*, *27*(9), 1775–1788.
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? Perception. 36 ECVF Abstract Supplement 14. Available: <http://www.perceptionweb.com/abstract.cgi?id=v070821>. Accessed 2013 Dec 16.
- Lipinski, J., Simmering, V. R., Johnson, J. S., & Spencer, J. P. (2010). The role of experience in location estimation: Target distributions shift location memory biases. *Cognition*, *115*, 147–153.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281. doi:10.1038/36846
- Machizawa, M. G., & Driver, J. (2011). Principal component analysis of behavioural individual differences suggests that particular aspects of visual working memory may relate to specific aspects of attention. *Neuropsychologia*, *49*(6), 1518–1526.
- Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Science U.S.A.*, *102*(24), 8776–8780. doi:10.1073/pnas.0500810102
- Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*, 369–378. doi:10.3758/BF03210419
- Perone, S., Molitor, S., Buss, A. T., Spencer, J. P., & Samuelson, L. K. (2015). Enhancing the executive functions of 3-year-old children performing the dimensional change card sort task. *Child Development*, *86*(3), 812–827. doi:10.1111/cdev.12330
- Perone, S., Simmering, V. R., & Spencer, J. P. (2011). Stronger neural dynamics capture changes in infants' visual working memory capacity over development. *Developmental Science*, *14*, 1379–1392. doi:10.1111/j.1467-7687.2011.01083.x
- Perone, S., & Spencer, J. P. (2013). Autonomous visual exploration creates developmental change in familiarity and novelty seeking behaviors. *Frontiers in Cognitive Science*, *4*, 648. doi:10.3389/fpsyg.2013.00648
- Perone, S., & Spencer, J. P. (2014). The co-development of looking dynamics and discrimination performance. *Developmental Psychology*, *50*(3), 837–852. doi:10.1037/a0034137
- Riggs, K. J., McTaggart, J., Simpson, A., & Freeman, R. P. J. (2006). Changes in the capacity of visual working memory in 5- to 10-year-olds. *Journal of Experimental Child Psychology*, *95*, 18–26. doi:10.1016/j.jecp.2006.03.009
- Riggs, K. J., Simpson, A., & Potts, T. (2011). The development of visual short-term memory for multifeature items during middle childhood.

- Journal of Experimental Child Psychology*, 108, 802–809. doi:10.1016/j.jecp.2010.11.006
- Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2007). Developmental aspects of visual recognition memory in infancy. In *Short- and long-term memory in infancy and early childhood: Taking the first steps toward remembering* (pp. 153–178). New York: Oxford University Press.
- Rose, S. A., Feldman, J. F., & Janowsky, J. J. (2002). Processing speed in the 1st year of life: A longitudinal study of preterm and full-term infants. *Developmental Psychology*, 38, 895–902. doi:10.1037/0012-1649.38.6.895
- Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, 74, 1807–1822.
- Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin and Review*, 18, 324–330.
- Sarigiannidis, I., Crickmore, G., & Astle, D. E. (2016). Developmental and individual differences in the precision of visuospatial memory. *Cognitive Development*, 39, 1–12. doi:10.1016/j.cogdev.2016.02.004
- Schneegans, S., Spencer, J. P., & Schöner, G. (2015). Integrating “what” and “where”: Visual working memory for objects in a scene. In G. Schöner, J. P. Spencer, & the DFT Research Group (Eds.), *Dynamic thinking: A primer on dynamic field theory* (pp. 197–226). New York: Oxford University Press.
- Schutte, A. R., & Spencer, J. P. (2009). Tests of the dynamic field theory and the spatial precision hypothesis: Capturing a qualitative developmental transition in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1698–1725.
- Simmering, V. R. (2008). *Developing a magic number: The dynamic field theory reveals why visual working memory capacity estimates differ across tasks and development* (Unpublished doctoral thesis). University of Iowa, Iowa City.
- Simmering, V. R. (2012). The development of visual working memory capacity in early childhood. *Journal of Experimental Child Psychology*, 111, 695–707. doi:10.1016/j.jecp.2011.10.007
- Simmering, V. R. (2015). *Feature binding in visual working memory is impaired during early childhood and can be improved through adding spatial structure*. Toronto: Presented at the Annual Meeting of the Jean Piaget Society.
- Simmering, V. R. (in press). Working memory capacity in context: Modeling dynamic processes of behavior, memory, and development. *Monographs of the Society for Research in Child Development*
- Simmering, V. R., Miller, H. E., & Bohache, K. (2015). Different developmental trajectories across feature types support a dynamic field model of visual working memory development. *Attention, Perception, & Psychophysics*, 77(4), 1170–1188. doi:10.3758/s13414-015-0832-6
- Simmering, V. R., & Patterson, R. (2012). Models provide specificity: Testing a proposed mechanism of visual working memory capacity development. *Cognitive Development*, 27(4), 419–439. doi:10.1016/j.cogdev.2012.08.001
- Simmering, V. R., & Perone, S. (2013). Working memory capacity as a dynamic process. *Frontiers in Developmental Psychology*, 3, 567. doi:10.3389/fpsyg.2012.00567
- Simmering, V. R., & Schutte, A. R. (2015). Developmental dynamics: The spatial precision hypothesis. In G. Schöner, J. P. Spencer, & the DFT Research Group (Eds.), *Dynamic thinking: A primer on dynamic field theory* (pp. 251–270). New York: Oxford University Press.
- Simmering, V. R., & Spencer, J. P. (2008). Generality with specificity: The dynamic field theory generalizes across tasks and time scales. *Developmental Science*, 11(4), 541–555. doi:10.1111/j.1467-7687.2008.00700.x
- Simmering, V. R., & Wood, C. M. (2016). The importance of real-time stability for visual working memory performance: Young children’s feature binding can be improved through perceptual structure. *Manuscript Submitted for Publication*.
- Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective, & Behavioral Neuroscience*, 5(2), 144–155. doi:10.3758/CABN.5.2.144
- van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, 121(1), 124–149. doi:10.1037/a0035234
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 92–114.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4(12), 1120–1135.
- Wolfe, J. M. (2014). Introduction to the special issue on visual working memory. *Attention, Perception, & Psychophysics*, 76(7), 1861–1870. doi:10.3758/s13414-014-0783-3
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440, 91–95.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233–235.