Abstract

In the study of visual working memory, different tasks have been used to assess capacity across development: preferential looking in infancy, and change detection in children and adults. These different tasks have produced inconsistent results over development, showing higher capacity at 10 months than at 5 years. To reconcile these differences, we developed a new version of change detection to use with younger children (ages 3-5 years), and tested children in both types of tasks. Results indicate that preferential looking leads to higher estimates of capacity than change detection. In addition, capacity estimates from preferential looking were higher than the standard 3-4 items found in change detection with adults. This difference across tasks and development is explained by implementing both tasks in a single theoretical framework, the Dynamic Field Theory.

Keywords: development; visual working memory; capacity; dynamic field model

Visual Working Memory Capacity

The past decade has seen an explosion of interest in the processes that underlie visual working memory (VWM) and the capacity of this memory system in adults (for historical precedent, see Miller, 1956). Beginning with Luck & Vogel’s (1997) seminal paper, change detection has been the method of choice for testing the capacity of VWM. Figure 1 shows a standard change detection trial. First, a memory array is shown briefly (100-500 ms) to the participant, followed by a brief delay (250-900 ms). Then, a test array appears in which either all of the objects match the memory array, or one object has changed to a new feature value. The participant reports whether there was a change in the display.

Findings across a host of studies have shown that adults have a capacity for approximately 3-4 simple items (e.g., colored squares or oriented bars). Although these general findings have been replicated widely, they are not without controversy. For example, recent studies have shown that the details of the object features held in working memory can limit the capacity of this memory system, suggesting that working memory utilizes a pool of limited resources that can modify the number of “slots” available in a given task (Alvarez & Cavanagh, 2004).

Although the exact nature of the cognitive representations that underlie VWM is under debate, researchers agree that VWM is a capacity-limited system. A relatively new question for VWM researchers is the developmental origin of such limitations. This question has been addressed in the domain of verbal working memory, where children show improvement over development in tasks like digit span. Interestingly, such limited capacity might have adaptive consequences early in development. For instance, in language development, the “less-is-more” hypothesis (e.g., Newport, 1990) suggests that children’s limited memory of variable input actually helps them learn a regularized grammar. Recently, researchers have begun to explore whether there is a comparable developmental change in VWM capacity.

To explore developmental changes in VWM, researchers have taken two approaches. First, some studies have tested children in the standard change detection task, with some modification (Riggs, McTaggart, Simpson, & Freeman, 2006; Cowan, Naveh-Benjamin, Kilb, & Saults, 2006). These studies have shown that children have a lower memory capacity than adults. Moreover, capacity increases between 5 and 12 years of age. For instance, Riggs et al (2006) used the Pashler (1988) model and obtained capacity estimates of 1.52 items for 5-year-olds, 2.89 items for 7-year-olds, and 3.83 items for 10-year-olds. Cowan et al. (2006) also found improvements in performance in change detection between 8 and 10 years of age.¹

¹ Cowan et al. (2006) did not compute capacity directly, but did report developmental improvements in children’s overall percent correct, which corresponds closely to capacity estimates.
infant spends looking at each display is tabulated across each 20 sec trial. Change-preference scores are then calculated as the proportion of time on each trial that the infant spent looking at the changing screen.

The rationale behind this method is that if the number of the items on the screen is within the infant’s VWM capacity, the infant will remember the items over the delay, will notice that the items are changing on one screen, and will prefer to look at that screen. If the number of items is beyond the infant’s capacity, however, the infant will not be able to notice the difference across sides and should show no preference. Using this method, Ross-Sheehy et al (2003) found that by the age of 10 months, infants showed preferences when the set size (i.e., number of objects in the array) was as high as four items. By contrast, 6-month-old infants only showed reliable change preferences for displays with a set size of one item. These findings suggest that between the ages of 6 and 10 months, infants’ VWM capacity is rapidly increasing to adult-like levels.2

Dynamic Field Theory and VWM Capacity

The Dynamic Field Theory (DFT) is a process model that was originally developed to capture developmental changes in spatial cognition (Spencer, Simmering, Schutte, & Schöner, 2007), and has recently been extended to capture the processes underlying maintenance of non-spatial features in VWM (see Johnson, Spencer, & Schöner, in press; Perone, Spencer, & Schöner, 2007). Figure 3 shows the basic five-layer architecture of the DFT that has been applied to both infant looking (Perone et al., 2007) and change detection (Johnson et al., in press) for non-spatial features. The model consists of five layers of neurons “tuned” to particular continuous feature dimensions such as color: a perceptual field (PF; Figure 3B) and an associated long-term memory field (LTM PF; Figure 3C), a layer of inhibitory interneurons (Inhib; Figure 3D), and a visual working memory field (WM; Figure 3E) with an associated long-term memory field (LTM WM; Figure 3F). Arrows in Figure 3 represent excitatory (green) and inhibitory (dashed) projections.

A simplified fixation system captures looking behavior in the model (Figure 3a, Perone et al., 2007). This system consists of three self-excitatory and mutually competitive nodes. When one of the nodes “wins” the mutual competition, the system “fixates” whatever is on the left screen, the right screen, or at an “away” location. These looking nodes are coupled to PF such that fixation to one display (e.g., activation of the left node) gates input (i.e., the visual contents of that display) to PF; in turn, above-threshold activation in PF supports activation of the currently active looking node. Thus, when the model “looks” to one display, input from that display comes into PF. Activation in PF feeds back to the active node, which maintains the current looking direction. This looking node will remain active and continue to fixate the same location as long as support from PF is strong enough and competition from the other nodes remains weak. If a competing node rises above threshold (e.g., due to stochastic fluctuations in activation), or if activation in PF drops, this will release fixation from the current node and allow the model to look elsewhere.

Input from the fixated display (e.g., the left display in Figure 2) projects strongly to PF and weakly to WM. This input leads to the formation of self-stabilized activation peaks in PF, corresponding to the feature values present in the fixated display (see Figure 3B). As activation builds in

---

2 Recent data suggests that this developmental transition occurs by about 7.5 months of age (Oakes, Messenger, Ross-Sheehy, & Luck, in press).
PF, it also builds in WM through the combination of weak input from the fixated display and stronger input from PF. Modulating this interaction between PF and WM, however, is a shared layer of inhibitory interneurons (see Inhib, Figure 3D). Because both layers receive inhibitory feedback from Inhib, peaks in PF and WM essentially compete with one another when coding for similar values. Note that activation in PF and WM is also supported by the LTM layers. As activation builds in PF or WM, traces of activation are laid down at associated sites in the LTM layers, leading to a form of Hebbian learning. Such traces feed back onto PF and WM, biasing the model to build peaks of activation at feature values that were activated on previous trials.

The following sections outline how the DFT explains the origin of capacity estimates from both types of tasks discussed above. This, in turn, led us to directly compare capacity estimates from the same children in the experiments reported here.

**Preferential Looking: Detecting ‘Same’**

To understand how the model would perform the preferential looking task, consider what is happening during a single trial with, for example, three colored squares on each screen. The squares are presented for 500 ms at a time, with 250 ms delays in between. On the “change” screen, one square will change color after each delay. On the “no-change” screen, all of the colors remain constant. If the model is “looking” at the no-change screen, this projects the input corresponding to the three colors to PF. These colors form three input-driven peaks in PF, sending additional activation to WM. Because the same colors are present after each delay, the activation sent to WM overlaps at each presentation, leading to the formation of self-sustaining working memory peaks (supported by LTM). Critically, forming peaks in WM takes some time, given that (a) the input from the display is turning “on” and “off” at a relatively rapid pace, and (b) the model has relatively weak neural interactions (which captures many characteristics of WM in early development, see Simmering, Schutte, & Spencer, 2008). Once peaks begin to form in WM, inhibition is projected back to PF via Inhib. This focused inhibition suppresses peaks in PF, removing support for the currently active fixation node. Consequently, the model looks away.

If the model then looks at the change screen, three different colors will now be input to PF. Because they are different than the colors that were presented on the no-change screen, the process of building peaks in WM starts again at these new feature values. After each delay, though, one of the colors will change. Because one component of the input is changing after each delay, peaks in WM will take longer to build, and fixation will be maintained for a longer period of time. As such, the model looks at the changing side longer, resulting in a change preference.

The model will show a change preference when it is able to successfully build WM peaks for items on the no-change screen; only then is the fixation system released to look at the other screen. This is where capacity limits emerge in the infant model: if, for example, neural interactions in WM allow only two peaks on the no change side to build simultaneously, then one of the colors in PF would not be suppressed because there was no corresponding WM peak. Consequently, the model would continue to fixate the no change side and fail to show a robust change preference score. Thus, change preferences in the preferential looking task arise from detection of sameness, that is, successful formation of a WM for items on the no change side.

**Change Detection: Detecting ‘Different’**

To capture performance in the change detection task, the excitatory layers of the model are connected to two self-excitatory, competitive decision nodes—one for “same”, one for “different”. Activation within PF is summed across the feature dimension and projected to the “different” node; activation within WM is summed and projected to the “same” node (Johnson et al., in press; see Simmering & Spencer, in press, and Simmering, Spencer, & Schöner, 2006, for a similar architecture). Note that, because there is only one display in this task, the fixation system does not play a critical role.

When the sample array is presented, input again projects strongly to PF and weakly to WM. As peaks build in PF, activation is projected to WM, and the model forms self-sustaining peaks in WM. As the peaks sustain, the corresponding inhibition projects back to PF, which leads to troughs of inhibition at the remembered values. When the test array is presented, features that match items in the sample array will come into inhibited regions of PF. This will prevent strong peaks from building in PF, the WM peaks will sustain, and the “same” node will win the response competition. If, however, one of the inputs in the test array has changed, that input will come into PF at a relatively uninhibited value; this allows an input driven peak to form in PF, leading to a “different” response. In this task, then, the building of activation in PF at test—the neural event that captures detection of a change—supports the generation of a “different” response.

**Testing the DFT View of Capacity**

In the DFT, both preferential looking and change detection depend on the interaction between peaks in PF and WM via the shared Inhib layer. PF essentially serves to compare the information held in WM and new input. However, the different task structures lead to important differences in how WM representations are formed and how this comparison process influences behavior. In preferential looking, WM peaks can build over repeated presentations, and performance depends on recognizing a match between WM and input—a match leads the model to look away, producing a change preference. In change detection, on the other hand, WM representations must be formed quickly and maintained robustly for comparison with the test array—when a test item does not overlap with WM, the
model produces a different response. These differences lead to the prediction that the same system will show different capacity estimates across tasks. Specifically, because preferential looking depends on building WM peaks with support from input, capacity estimates from this task should be relatively high. In change detection, on the other hand, performance depends on maintaining stable peaks through a delay to detect changes during a single comparison at test. This requires more stable interactions in the model, and therefore should produce lower capacity estimates compared to preferential looking. Thus, based on the DFT, we predict that children will show higher capacity estimates in preferential looking than in change detection. Moreover, because children’s WM should be more stable than infants’ WM (as seen in the development of spatial working memory; see Simmering et al., 2008, for discussion), capacity estimates from preferential looking should be higher for children than the 4 items estimated for infants (Ross-Sheehy et al., 2003). Note that this suggests children’s capacity in preferential looking will be higher than the standard estimated capacity of adults’ WM based on change detection tasks. To test these predictions, we adapted both tasks for use with 3-, 4-, and 5-year-olds.

**Method**

**Participants** A total of 155 children participated in this study: 56 3-year-olds (M age = 3 years, 4.46 months; SD = 2.52 months), 44 4-year-olds (M age = 4 years, 2.87; SD = 1.93 months), and 55 5-year-olds (M age = 5 years, 0.81 months; SD = 1.74 months). All children participated in the preferential looking task, but only a subset of these children also participated in a color change detection task: 14 3-year-olds (M age = 3 years, 4.68 months; SD = 2.35 months), 14 4-year-olds (M age = 4 years, 2.56 months; SD = 1.77 months), 14 5-year-olds (M age = 5 years, 0.36 months; SD = 0.78 months). Note that a higher sample size was needed for the preferential looking task to achieve the necessary statistical power to detect significant preference scores. Preliminary analyses showed no difference in preferential looking performance across children who did or did not complete the change detection task.

**Apparatus** For the preferential looking task, we used the apparatus from the infant studies described above (Ross-Sheehy et al., 2003, Experiment 3). Looking was coded to the prediction that the same system will show different response. These differences lead to the prediction that the same system will show different capacity estimates across tasks. Specifically, because preferential looking depends on building WM peaks with support from input, capacity estimates from this task should be relatively high. In change detection, on the other hand, performance depends on maintaining stable peaks through a delay to detect changes during a single comparison at test. This requires more stable interactions in the model, and therefore should produce lower capacity estimates compared to preferential looking. Thus, based on the DFT, we predict that children will show higher capacity estimates in preferential looking than in change detection. Moreover, because children’s WM should be more stable than infants’ WM (as seen in the development of spatial working memory; see Simmering et al., 2008, for discussion), capacity estimates from preferential looking should be higher for children than the 4 items estimated for infants (Ross-Sheehy et al., 2003). Note that this suggests children’s capacity in preferential looking will be higher than the standard estimated capacity of adults’ WM based on change detection tasks. To test these predictions, we adapted both tasks for use with 3-, 4-, and 5-year-olds.

The position of the “cards” on each trial alternated from left to right. This helped reduce interference across trials. Stimuli were 1” x 1” colored squares; colors included red, green, blue, yellow, cyan, violet, black, and white (see Vogel, Woodman, & Luck, 2001, for precise color values). Colors were selected randomly on each trial with the constraint that a color could not appear twice within the same display. Stimuli could appear at one of five equally-spaced positions in a 2” diameter circle around the center of the frame. For each SS, the positions were chosen randomly for the first trial, but remained constant across the 12 trials to that SS.

**Procedure** After the parent completed the consent form, the child and parent were taken into the main laboratory room and completed the first task—the preferential looking task—which used the same procedure as in previous studies (see Ross-Sheehy et al., 2003, for full details), with two exceptions. First, most children completed this task alone, seated in front of the two video monitors, with the coder behind a curtain. If the child wanted the parent to accompany them, though, the parent wore darkened glasses and sat with the child in his/her lap. Second, children were told that they were going to watch six short videos, and that the experimenter would ask them questions about the videos at the end. No further instructions were provided, as we did not want to bias children’s looking. Children completed six 20-s trials in random order: two each showing SS2, SS4, and SS6. The screen showing the changing display was counter-balanced within each SS. Note that we tested children in the preferential looking task first to ensure that their experience in the change detection task would not influence their looking behavior.

After completing the looking task, children took a short break before starting the change detection task, to ensure that fatigue would not impair performance in this task. Once they were seated in the change detection room, the task was described as a matching game in which they had to help a teddy bear find cards that matched. The experimenter then demonstrated the task using flashcards for training. Flashcards were placed on either the right or left side (alternating across trials) of a large sheet of cardstock in the experimenter’s lap. The first card was shown for approximately 2 s, and the child was instructed to “Look at the picture and remember the colors”. The card was then removed and after a brief delay, the second card for the trial was shown in the same location. The experimenter then asked the child if the two cards matched. After the child responded, the experimenter placed both cards side-by-side and praised or corrected the child as needed. For all children, the flashcard trials were presented in the same order: SS1 no-change, SS1 change, SS2 change, SS2 no-change, SS3 change, SS3 no-change.

Once the child understood the task, the experimenter began the computerized version of the task. The first block of trials were practice, with the sample array presented for 2 s, a delay of 500 ms. The test array remained visible until
the response was entered. The practice block included eight trials in random order: four trials to SS1 and four trials to SS2 (half of the trials were change trials, half were no change trials). Between each block of trials, the child was offered a short break. The order of test blocks was semi-randomly assigned (see Riggs et al., 2006), such that the first two blocks were either SS1 or SS2; for 3- and 4-year-olds, the third block was always SS3; for the 5-year-olds, SS3-5 were randomly ordered across blocks 3-5. A different order was used for younger children because they often chose not to complete SS4 and 5; these high-SS blocks were difficult for all children, and young children tended to become discouraged as their performance declined. Each block of test trials included six change and six no-change trials, in random order. For 5-year-olds, the duration of the sample array on test trials was shortened to 500 ms, to be comparable with previous studies (Riggs et al., 2006).3

Results

Preferential Looking Change-preference scores were computed as the time (in s) children spent looking to the change display divided by the total looking time to both displays (i.e., change / [change + no-change]) for each trial. We then averaged across the two trials to the same SS. Mean change-preference scores for each SS in each age group, shown in Figure 4, were compared to chance (0.5) using two-tailed t-tests. These analyses revealed that each of these change-preference scores was significantly above chance (3 years: $t_{35} = 7.29$ [SS2], 2.75 [SS4], 2.80 [SS6], $p < .01$; 4 years: $t_{43} = 3.46$ [SS2], 2.64 [SS4], 2.57 [SS6], $p < .05$; 5 years: $t_{64} = 5.13$ [SS2], 2.57 [SS4], 2.77 [SS6], $p < .05$). These results suggest that 3-, 4-, and 5-year-old children have a VWM capacity of at least 6 items, although this estimate may have been higher had we included trials with higher SS. The estimates of capacity from this task show an increase relative to the performance of 10-month-olds, but also suggest a higher capacity than that shown by adults in change detection.

Change Detection For each SS, we computed three measures of performance: percent correct, percent hits (i.e., correct responses on change trials), and capacity. Note that, although all 5-year-olds completed SS4 and SS5, most 3- and 4-year-old did not; therefore, we analyzed only percent correct and percent hits across SS1-3. For each child, however, we included all completed SS to compute capacity estimates.

Figure 5 shows children’s percent correct across SS, separated by age. An ANOVA with SS as a within-subjects factor and Age as a between-subjects factor revealed significant main effects of SS ($F_{2, 76} = 36.60, p < .001$) and Age ($F_{2, 39} = 4.44, p < .05$). Follow-up Tukey HSD tests ($p < .05$) showed that performance decreased significantly from SS1 ($M = 93.16$) to SS2 ($M = 89.09$) and SS3 ($M = 76.64$), as well as from SS2 to SS3; in addition, performance was higher for 5-year-olds ($M = 90.38$) than 3-year-olds ($M = 80.77$), but neither differed from 4-year-olds ($M = 87.74$).

Figure 6 shows children’s percent hits (i.e., correct responses on change trials) across SS, separated by age. An ANOVA with SS as a within-subjects factor and Age as a between-subjects factor revealed significant main effects of SS ($F_{2, 78} = 22.54, p < .001$) and Age ($F_{2, 39} = 3.52, p < .05$). Follow-up Tukey HSD tests ($p < .05$) showed that performance decreased significantly from SS1 ($M = 93.43$) to SS3 ($M = 73.23$), as well as from SS2 ($M = 88.38$) to SS3, while SS1 and SS2 did not differ; in addition, performance was higher for 5-year-olds ($M = 90.75$) than 3-year-olds ($M = 77.90$), but neither differed from 4-year-olds ($M = 85.00$).

Figure 7 shows capacity estimates across age groups.4 A one-way ANOVA revealed a significant main effect of Age ($F_{2, 42} = 8.58, p < .01$). Follow-up Tukey HSD tests ($p < .05$) showed that capacity was higher for 5-year-olds than 3- and 4-year-olds.

---

3 We also collected data from a group of 5-year-olds using an exact replication of the Riggs et al. (2006) design, for comparison with our modified “card” task. Analyses revealed no differences across tasks, suggesting that, although our modifications made the task easier for younger children to understand, it did not improve performance relative to the design used by Riggs et al. In addition, this suggests that completing the preferential looking task first did not affect children’s performance.

4 Note that these capacity estimates are slightly larger than those previously reported in the development literature. This is due to slight differences in the use of Pashler’s (1988) formula for capacity. Specifically, Riggs et al (2006) averaged across SS blocks to arrive at a single capacity estimate for each child. As an alternative, we chose to use each child’s highest estimate across SS blocks; this prevents under-estimation due to small estimates from lower SS blocks (see Olsson & Poom, 2005).
for both 3- and 4-year-olds, and capacity did not differ significantly from 3 to 4 years.

Cross-Task Comparisons Results from the two tasks separately indicate much higher capacity estimates in preferential looking (at least 6 items) than in change detection (2-3 items). However, because both tasks rely on the same underlying memory system, we would expect a relationship in performance across these tasks. To test this, we compared performance across tasks directly for the children that participated in both tasks, by correlating capacity with change-preference scores across age groups, shown in Figure 8. As this figure shows, capacity estimates were positively correlated with SS2 change-preference score (Pearson’s R = .328, p < .05). Thus, although these two tasks yield different estimates of capacity over development, performance across these tasks is related.

Discussion

Based on the Dynamic Field account of the processes that underlie preferential looking and change detection, we predicted that the looking task would yield higher capacity estimates than change detection in 3- to 5-year-old children, and that estimates of children’s capacity from the looking task would also be higher than the standard 3-4 items estimated in preferential looking with infants and change detection with adults. Results supported these predictions, showing capacity estimates of at least 6 items for all age groups in the looking task, but capacities of only about 2-3 items in change detection. This difference across tasks reflects the different demands placed on the perceptual and working memory system.

Conclusions

Although both preferential looking and change detection measure aspects of VWM, the two tasks are not equivalent and bear distinct relationships to the processes that underlie VWM. This answers an important developmental question: why did previous studies show a discontinuity in memory capacity from 10 months to 5 years of age? To address this question, it was necessary to understand the real-time dynamics at work in each of these tasks, as well as testing both tasks in the same age groups.

Our results raise an important question about the characterization of VWM as having a specific “capacity” that may be directly measured, independent of the behavior required in the task. Similar questions have been raised by researchers in the adult literature (e.g., Alvarez & Cavanagh, 2004). By implementing these tasks in a real-time process model of VWM, we have begun to understand the nature of the representations that underlie VWM as well as some of the magic that leads to developmental changes in working memory capacity.

Acknowledgments

Thanks to Jeff Johnson, Sammy Perone, and Shannon Ross-Sheehy for help with the background and design of this study. Thanks also to Valerie Vorderstrasse for help with piloting and data collection. This research was funded by NIMH MH62480 and NSF HSD0527698 awarded to JPS.

References


