Delayed Match Retrieval: a novel anticipation-based visual working memory paradigm

Zsuzsa Kaldy, Sylvia B. Guillory and Erik Blaser

Department of Psychology, University of Massachusetts Boston, USA

Abstract

We tested 8- and 10-month-old infants' visual working memory (VWM) for object-location bindings – what is where – with a novel paradigm, Delayed Match Retrieval, that measured infants' anticipatory gaze responses (using a Tobii T120 eye tracker). In an inversion of Delayed-Match-to-Sample tasks and with inspiration from the game Memory, in test trials, three face-down virtual 'cards' were presented. Two flipped over sequentially (revealing, e.g. a swirl pattern and then a star), and then flipped back face-down. Next, the third card was flipped to reveal a match (e.g. a star) to one of the previously seen, now face-down cards. If infants looked to the location where the (now face-down) matching card had been shown, this was coded as a correct response. To encourage anticipatory looks, infants subsequently received a reward (a brief, engaging animation) presented at that location. Ten-month-old infants performed significantly above chance, showing that their VWM could hold object-location information for the two cards. Overall, 8-month-olds' performance was at chance, but they showed a robust learning trend. These results corroborate previous findings (Kaldy & Leslie, 2005; Oakes, Ross-Sheehy & Luck, 2006) and point to rapid development of VWM for object-location bindings. However, compared to previous paradigms that measure passive gaze responses to novelty, this paradigm presents a more challenging, ecologically relevant test of VWM, as it measures the ability to make online predictions and actively localize objects based on VWM. In addition, this paradigm can be readily scaled up to test toddlers or older children without significant modification.

Research highlights

• Here we introduce a novel paradigm to test infants’ visual working memory (VWM) using anticipatory gaze responses.
• Our paradigm improves on current approaches that are based on the novelty preference by encouraging participants (through rewards) to make online predictions and actively localize objects based on remembered information.
• Ten-month-olds were able to reliably maintain two object-location bindings over a 1.5-s delay. Eight-month-olds showed a robust learning trend.
• Our paradigm is well suited for longitudinal studies of VWM development as it is readily 'scaled-up' to test toddlers, older children or atypically developing, language-delayed populations without modification.

Introduction

Developmental psychologists have studied infants’ memory for over 50 years, often focusing on visual stimuli, such as faces or geometric patterns (e.g. Fagan, 1977; Rose, 1981; Rose, Feldman & Jankowski, 2001); showing, for example, that even a few-hour-old infant can remember a visual stimulus (Fantz, 1964; Slater, Earle, Morison & Rose, 1985). But beyond recognition of an object, or noting the relevance of a particular location, what is the developmental course of the integration of this information? After all, it is reasonable to think that the functional units of visual cognition are objects situated in space (Hollingworth & Rasmussen, 2010; Kahneman, Treisman & Gibbs, 1992; Leslie, Xu, Tremoulet & Scholl, 1998). Studies using classic methods, such as the Delayed Response task, with eye-gaze-based measures, have shown that 5-month-old infants can locate a hidden
object from multiple hiding locations over a 3-second delay (Hofstadter & Reznick, 1996; Reznick, Morrow, Goldman & Snyder, 2004). This ability gradually develops over the second half of the first year (Pelphrey, Reznick, Goldman, Sasson, Morrow et al., 2004). This has also been demonstrated using violation-of-expectation paradigms with 5- and 6-month-olds (Newcombe, Huttenlocher & Learmonth, 1999; Kaldy & Leslie, 2005) with even longer delays (7–10 s). On the other hand, there is ample evidence showing that infants at 4–6 months of age have trouble retaining featural/identity information about objects when more than one object is involved (Kaldy & Leslie, 2005; Kibbe & Leslie, 2011; Kwon, Luck & Oakes, 2013; Mareschal & Johnson, 2003; Simon, Hespos & Rochat, 1995). In a task that required remembering two object identity/location bindings, 9-month-olds (but not 6-month-olds) were able to succeed with a 7-second delay (Kaldy & Leslie, 2005). When the retention interval is minimized to 300 ms, there is evidence that infants as young as 7.5 months can remember object-location bindings for multiple objects (Oakes, Ross-Sheehy & Luck, 2006).

Addressing limitations of VWM studies based on the novelty preference

In all of the above studies (except for the delayed-response task), the participant is a passive observer of a sequence of stimuli and events and the assumption is that he or she will spontaneously compare the contents of a memory representation with the currently perceived stimuli (for reviews, see Kavšek, 2013; Rose, Feldman & Jankowski, 2004). This same logic underlies the violation-of-expectation paradigms used to examine higher-level object cognition abilities (Baillargeon, Spelke & Wasserman, 1985) and classic studies using novelty preference procedures in which infants are shown a stimulus, and then their memory for that stimulus is inferred from their looking preference for a novel one (e.g. Visual Paired Comparison: Fagan, 1970; habituation/dishabituation: Pancratz & Cohen, 1970). Thus far, work on VWM for object-location bindings in infants has used the same general approach (Kaldy & Leslie, 2003, 2005; Mareschal & Johnson, 2003; Newcombe et al., 1999; Ross-Sheehy, Oakes & Luck, 2003). These VWM paradigms based on novelty preferences have moved our understanding forward, but have three limitations that our new paradigm seeks to address.

For one, the novelty preference requires both detection of a change and interest in that change. When an infant fails to react when, say, a red star goes behind an occluder and a pink star subsequently emerges, we do not know whether the infant failed to detect the change or was just not surprised by the change; the change may be noted but not notable. In our paradigm, we employ infant-tailed ‘rewards’ (e.g. engaging visual animations) for correct responses, similarly to Delayed-Match-to-Sample paradigms (since Weinstein, 1941). A compelling reward offers an incentive to remember and respond that is independent of inherent interest in task stimuli or events (a logic underlying training paradigms: e.g. Siqueland & Lipsitt, 1966; Rovee-Collier & Gekoski, 1979). (It is worth noting here too that the ‘mode’ of rewards is crucial. For instance, early results showed that infants failing at Delayed Non-matching to Sample (DNMS) until 21 months of age (Diamond, 1990; Overman, 1990). However, Diamond and her colleagues later showed that infants could succeed much earlier (at 9 months), if their response and the reward were clearly connected (e.g. immediate verbal praise or a reward physically attached to the sought-after novel object (Diamond, Churchland, Cruess & Kirkham, 1999)).

Secondly, the novelty preference can be ambiguous. When an infant does look longer when, say, a green star emerges where a red one had been, does the infant think that the the red star has undergone an unnaturally large change in appearance, or that a new object has usurped its location? Our paradigm sidesteps this issue altogether and asks a different question, simply, ‘Please show us where you saw this object before.’ This example also illustrates what we think is the third contribution of our paradigm. The novelty preference is passive; a reaction to a change in an object-location binding. There is much more to object cognition in dynamic real world settings than the expression of surprise. Many situations require the infant to recall and act upon what is where information. Our paradigm is specifically tailored to test the development of this ability.

Delayed Match Retrieval (DMR) requires the online application of object-location bindings

Eye-trackers have made the measurement of anticipatory gaze responses possible and enabled a ‘shift from studies of the macrostructure of looking behavior to the microstructure of patterns of fixation’ (Aslin, 2012, p.127). Predictive gaze responses have been used fruitfully in studies of infants’ social cognition (e.g. Falck-Ytter, Gredebäck & von Hofsten, 2006; Southgate, Senju & Csibra, 2007) and other domains, such as object tracking (Johnson, Amso & Slemmer, 2003), cognitive flexibility (Kovács & Mehler, 2009a, 2009b; Albareda-Castellot, Pons & Sebastián-Gallés, 2011) and categorization (McMurray & Aslin, 2004). The dependent variables are usually based on ‘first looks’ (both location and latency) and fixation durations. For instance,
Addyman and Marechal (2010) tested 4- and 8-month-old infants’ ability to learn the categories ‘same’ and ‘different’ in an anticipatory looking task. Infants were trained to search in one location if they categorized a pair of objects as ‘same’ and in another location if they categorized them as ‘different’. The authors measured infants’ binary preference based on first looks and proportion of looking time to the two alternatives. Both of these two measures showed rapid learning in the older group, but only the proportion of looking time did in the younger group. In the current study, we used both measures.

Our novel Delayed Match Retrieval (DMR) paradigm inverts the classic Delayed-Match-to-Sample task. In the classic task, a sample object is presented (say, a star) and removed, a delay period ensues, and then two objects are presented: a match (another star) and a non-match (say, a triangle). An adult would be asked, or an animal trained with rewards, to pick out the match. Importantly, DMR inverts this: potential matches/non-matches are presented first, then obscured. Only after that (and a delay) is the sample revealed; the participant then seeks out the (now hidden) match. This inversion means that high performance is achieved only when all object-location bindings are successfully maintained in VWM during the delay. Participants are encouraged to do this through visual rewards (i.e. ‘Here is a red star on the left and a green star on the right. Hold this what is where information in memory while they are hidden. Now, here is another red star. If you look back to where you saw the first one, you’ll see a fun reward there!’). Building on Addyman and Mareschal’s findings with 4- and 8-month-olds, we expected the 8–10-month-old infants in our task to be able to rapidly learn this rule. Our paradigm included training trials where (1) infants could familiarize themselves with the cards and images, (2) looking behavior could be recorded as the cards were flipping and moving, and (3) infants were shown that transitory, engaging animations occurred at the Match location (test trials also had this reward structure to serve as rule training).

Method

Participants

Two groups of healthy, full-term infants participated. There were 14 infants in the 8-month-old age group ($M_{\text{age}} = 243.1$ days, $SD = 13.2$, range: 224–261 days, six females) and 12 in the 10-month-old group ($M_{\text{age}} = 297.2$ days, $SD = 17.3$, range: 270–328 days, four females). Eleven additional infants were tested, but excluded (four due to fussiness, one due to inability to calibrate, three due to eye-tracker errors and three due to experimenter errors). Families were contacted using birth records, resided primarily in the Greater Boston area, and received a small gift for participation. None of our infant participants had first-degree relatives with color-blindness.

Stimuli

Stimuli were self-occluding, virtual ‘cards’ that could be presented face-up, revealing to-be-remembered objects, or, un informatively, face-down. This design is quasi-naturalistic (no sudden onset/offset of objects), and it has the advantage of not requiring additional objects to serve as occluders. One of four possible objects – star, swirl, face, or tree – could appear on a card (Figure 1a). Cards subtended $4 \times 6$ deg and were arranged symmetrically (top, left bottom, right bottom) with their centers 5 deg from the center of the screen.

**Figure 1**  (a) Experimental stimuli. Face-up, virtual ‘cards’ used in the roles of Sample/Match, and non-Match could contain one of four objects. (b) Areas of Interest (AOIs) for eye-tracking analyses were drawn around the Match and non-Match cards (shown face-down).


**Apparatus and procedure**

We used a Tobii T120 eye-tracker (Tobii Technology, Stockholm, Sweden) and Tobii Studio 2.1.8 software to measure eye movements. Participants sat on their caregivers’ lap, approximately 60 cm away from the eye-tracker’s display in a dimly lit, isolated testing area. Caregivers wore occluding glasses and were asked not to interact with their infants during testing. Before each block of trials, infants performed the default Tobii 5-point infant gaze calibration.

In *training trials*, there was no VWM task. Instead, infants were familiarized to the three-card configuration, objects, behaviors, sound effects, and the rule of ‘matching’ (Figure 2). First, infants saw three face-down cards fly in. After 1.5 s, the *Sample* card was revealed. After 2.5 s, the *Match* was revealed, accompanied by one of three brief, rewarding animations (e.g. the *Match* card jiggled and threw off pink sparks). The Match then moved next to the Sample and tilted toward it to make contact (imitating a ‘kiss’, accompanied by a kissing sound). The third card never flipped face-up. At the end of the trial, all cards flew off the screen. There were four such training trials, one for each of the four objects.

*Test trials* began when three face-down cards flew into view (Figure 3, please see also Supplementary Movie 1). Following this, the potential Match card and the potential *non-Match* card were exposed (order of exposure was counterbalanced). The first card was exposed for 2.5 s and remained face-up while the second card was exposed for 2.5 s. Both cards were then flipped face-down, starting a 1.5-second retention interval during which the participants had to maintain object-location information in VWM (i.e. the binding of location and identity of the cards). Then the third card, the Sample, flipped over, revealing a match to one of the previously seen, now face-down cards. This began the 4 s response interval, during which eye movements were monitored (see *Data analysis*).

1 The response interval is longer than is typical in anticipatory eye-movement paradigms. We had conducted a pilot study with a shorter (1.6-s) interval and found that many trials were lost due to the infants lingering on the just-exposed Sample and therefore not having sufficient time to fixate the Match and/or non-Match card before the response interval ended. We extended this interval to be at least 3 × the standard deviation of the mean latency of first fixation (to a valid AOI).
The stimulus pair (red/swirl or tree/face), the order of exposure of the Match and non-Match cards, which of the three cards was in which position (top, left bottom, and right bottom, see Figure 1b), and the identity of the Match were all randomized across trials in a block. The total length of a test trial was 19 seconds.

In each block of 16 trials, infants were first shown the four familiarization trials followed by 12 test trials with three attention-grabber sequences mixed in (where all four types of cards (Figure 1a) were shown jumping up and down). The total length of the block was 5 minutes. Infants ran two blocks of trials with a 3–5-minute break in between. Infants were re-calibrated before the second block. All events were accompanied by sound effects to maintain engagement.

Data analysis

To measure gaze behavior, we defined two identical rectangular Areas of Interest (AOIs, size: 7 × 9 deg) around the Match and non-Match cards (Figure 1b). Using these data, we determined infants’ choice between the (correct) Match and the (incorrect) non-Match card on each trial, during the 4-second response interval while both were face-down. If the infant did not fixate either of the two AOIs during the response interval, the trial was excluded from analysis.

Participants should be motivated to use remembered information to maximize their chances of fixating the reward animation, i.e. to look at the Match card as quickly and accurately as possible – and to linger there – in anticipation of the reward. This ability, then, should be captured by seeing card they look at first (i.e. percent correct performance, the percent of trials where the infant only fixated the Match, or fixated the Match before the non-Match), and which card they look at longest (i.e. percent looking time, the relative proportion of time spent fixating the Match card: time_on_Match/ (time_on_Match + time_on_non-Match)); these form our two dependent variables. We also analyzed the average-first look latency² to the Match versus the non-Match, to capture the time it takes for an infant to make a decision.

Results

In all, 294 and 244 trials were presented to the younger and the older group, respectively. The number of ‘valid’

² Trial-by-trial paired comparison of fixation latency to the Match vs. the non-Match cannot be conducted in the majority of trials, since infants typically only looked at one of the two target cards.

Main results based on first looks

To assess percent correct performance, we used a univariate ANCOVA with Age group (8- vs. 10-month-olds) and Gender (m/f) as fixed factors (both centered), Number of valid trials as a covariate (centered), and percent correct based on first looks as the dependent variable. This full model showed that the main effect of gender was not significant (p = .422; η_p^2 = 0.031) so this factor was dropped from further analyses. The resultant model shows a significant main effect of Age group (F_1 = 8.09, p = .009, η_p^2 = 0.269). The main effect of Number of valid trials and the interaction between the two factors was not significant (F_1 = 3.80, p = .064, η_p^2 = 0.147; F_1 = 3.21, p = .087, η_p^2 = 0.127, respectively).

Planned comparisons of percent correct performance showed that 10-month-olds significantly outperformed 8-month-olds (M_10 months = 61.66%, SD = 15.2, M_8 months = 45.45%, SD = 11.4, d = 1.207, t_24 = 3.10, p < .005, two-tailed) and that performance was significantly above chance in the older, but not in the younger age group (d = 0.767, t_11 = 2.659, p = .022; d < 0.01, t_13 = -1.495, p = .159, respectively, see Figure 4a).

Average first look latency to the Match versus the non-Match was not different in either of the age groups (8-month-olds, M_Match = 1.55 s, SD = 0.47, M_non-Match = 1.42 s; SD = 0.35, d = 0.301, t_13 = 0.834, p = ns; 10-month-olds, M_Match = 1.44 s, SD = 0.40, M_non-Match = 1.39 s, SD = 0.63, d = 0.080, t_11 = 0.17, p = ns); infants took no longer to make an incorrect decision than a correct one.

Main results based on percent looking time

Mirroring our main results, planned comparisons showed that percent looking time (to the Match) was significantly above chance in the older, but not in the younger, age group (M_10 months = 60.47%, SD = 14.7, d = 0.712, t_11 = 2.472, p = .031; M_8 months = 51.69%, SD = 12.6, d = 0.134, t_13 = 0.501, p = ns). The difference between the two groups did not reach significance (d = 0.642, t_24 = 1.641, p = .114, see Figure 4b). Given the length of our response interval (4 s, which included
exposition of the Sample), infants often did not have time to look at both target cards, therefore it is not surprising that measures based on first looks and looking time are tightly coupled (r = 0.844, p = .0001).

The effect of learning
In order to examine learning trends, we compared percent correct performance in the first vs. the second block of trials. (Three infants in each age group did not have data in their second block, leaving nine infants in the older and 11 infants in the younger group for this within-subject comparison.) A univariate ANOVA with Age group (8- vs. 10-month-olds) and Block (first/second) as fixed factors was conducted. The main effects were not significant (F_1 = 2.618, p = .114, η^2_p = 0.068; F_1 = 0.813, p = ns, η^2_p = 0.022, respectively), but there was a significant interaction between the two factors (F_1 = 4.862, p = .034, η^2_p = 0.119). Planned comparisons showed that younger infants’ performance significantly increased over time (M_8 months, Block1 = 39.63%, SD = 15.8, M_8 months, Block2 = 61.55%, SD = 21.8, d = 1.151, t_{10} = 2.491, p = .032), while older infants’ did not (M_{10 months}, Block1 = 66.60%, SD = 18.2, M_{10 months}, Block2 = 57.41%, SD = 31.3, d = 0.359, t_{8} = −0.966, p = ns; see Figure 5a). The same planned comparisons were conducted with percent looking time, yielding essentially identical results (see Figure 5b).

Discussion
We developed a novel, non-verbal, anticipation-based visual working memory paradigm to measure memory for object-location bindings: Delayed Match Retrieval (DMR; an inversion of Delayed Match-to-Sample, with inspiration from the game Memory). Success in DMR requires that participants maintain and use what is where information in VWM. We found that 10-month-old infants remembered the object-location bindings for two objects. Overall, 8-month-olds’ performance was at chance, but showed a robust learning trend. The results of the younger infants are very intriguing; they may also have this ability, but may require more training to acquire the ‘matching’ rule. Further studies are planned to investigate this.

Our pattern of results is consistent with previous novelty-based studies of infants’ VWM capacity for object-location bindings (Oakes et al., 2006; Kaldy & Leslie, 2005). DMR provides novel insight into VWM as it overcomes some of the limitations of novelty-based tests. Compared to violation-of-expectation paradigms that measure passive gaze responses to novelty, this paradigm presents a more challenging, arguably more ecologically relevant test, as it requires online predictions and active localization based on remembered information. (And this information need not be strictly visual. DMR could be used to investigate how auditory

We would like to note that there is a potential strategy to aid performance that does not require memorizing all object-location bindings. With two cards, the participant could memorize the object-location binding for just one of them and then when the Sample is exposed, use process-of-elimination or disjunctive syllogism. We believe that this strategy is more cognitively challenging than remembering the locations of both objects. In any case, it is likely not available at 10 months of age. The mutual exclusivity bias in word learning (Markman, 1990; Merriman & Bowman, 1989) has similar reasoning requirements, and in a looking time paradigm, only 17-, but not 14-month-olds, were able to apply this strategy (Halberda, 2003).
information is integrated with object representations. For example, would verbal labels help infants retain object-location bindings in our task? This would be a particularly interesting question to investigate in younger, 8-month-old infants.)

Most tests of visual short-term and working memory are either tailored for infants (4–12-month-olds, change detection with two streams (e.g. Ross-Sheehy et al., 2001), violation-of-expectation (e.g. Kaldy & Leslie, 2005)), for young toddlers (12–20-month-olds, reaching and cracker choice tasks; e.g. Feigenson & Carey, 2005) or for preschool-age children (3–5-year-olds, change detection task; e.g. Simmering, 2012). Novelty-based paradigms for infants are notoriously hard to implement in toddlers, creating a ‘toddler gap’. DMR is readily scaled-up for toddlers and older children, for instance by increasing the number of cards or the similarity of objects (for this, we have developed methods for ‘calibrating’ stimuli along different feature dimensions: Kaldy, Blaser & Leslie, 2006; Kaldy & Blaser, 2009, 2013). Indeed, an ongoing study using DMR with 21-month-olds is currently under way in our laboratory (O’Grady, Guillory, Blaser & Kaldy, 2015) and preliminary findings show that the task is sufficiently engaging for toddlers. Bridging the toddler gap is also important for future longitudinal studies. Such studies hold a particularly important promise since individual differences in VWM show remarkable stability in early development (Bell & Wolfe, 2007) and VWM capacity highly correlates with non-verbal IQ in children (Cowan, Fristoe, Elliott, Brunner & Saults, 2006). We believe that DMR is particularly well suited for tracking individual differences in infancy and beyond.

Acknowledgements

This research was supported by National Institutes of Health Grant 2R15EY017985 awarded to EB and ZK. Preliminary results from this study were presented at the Biennial Meeting of the International Society on Infant Studies in Minneapolis, MN, in June 2012 (Kaldy, Blaser & Biondi, 2012). We would like to thank Marisa Biondi, Shaun O’Grady, and other members of the UMass Boston Baby Lab for their help with data collection.

References

Diamond, A., Churchland, A., Crueiss, L., & Kirkham, N.Z. (1999). Early developments in the ability to understand the
relation between stimulus and reward. Developmental Psychology, 35, 1507–1517.


A novel visual working memory paradigm

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Supplementary Movie 1 Three sample DMR test trials showing the eye trace recording of an infant in the 10-month-old group. See Figure 3 for a description of test trial events. (The file is in MPEG-4 format.)