Journal of Experimental Child Psychology 187 (2019) 104649

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Two-year-olds succeed at MIT: Multiple identity tracking in 20- and 25-month-old infants

Chen Cheng *, Zsuzsa Kaldy, Erik Blaser

Department of Psychology, University of Massachusetts Boston, Boston, MA 02125, USA

ARTICLE INFO

Article history:
Received 10 January 2019
Revised 6 June 2019

Keywords:
Object tracking
Multiple identity tracking
Multiple object tracking
Delayed match retrieval
Visual working memory
Infants

ABSTRACT

Infants' ability to remember objects and their locations emerges during the first year of life. However, not much is known about infants' ability to *track* objects' identities in a dynamic environment. Here, we tailored the delayed match retrieval eye-tracking paradigm to study infants' ability to track two object identities during occlusion—an infant version of multiple identity tracking (MIT). Delayed match retrieval uses virtual "cards" as stimuli that are first shown face up, exposing to-be-remembered information, and then turned face down, occluding it. Here, cards were subject to movement during the face-down occlusion period. We used complex non-nameable objects as card faces to discourage verbal rehearsal. In three experiments ($N = 110$), we compared infants' ability to track object identities when two previously exposed cards were static (Experiment 1), were moved into new positions along the same trajectory (Experiment 2), or were moved along different trajectories (Experiment 3) while face down. We found that 20-month-olds could remember two object identities when static; however, it was not until 25 months of age that infants could track when movement was introduced. Our results show that the ability to track multiple identities in visual working memory is present by 25 months.

* Corresponding author.
E-mail address: chen.cheng001@umb.edu (C. Cheng).

https://doi.org/10.1016/j.jecp.2019.06.002

Introduction

Multiple identity tracking

Keeping track of individual objects and agents—teammates on a field, passing cars on the road—is fundamental to maintaining an up-to-date representation of a dynamic visual world. Multiple object tracking (MOT) was introduced 30 years ago to study this skill experimentally (Pylyshyn & Storm, 1988). In this paradigm, the observer's task is to track a cued target subset of identical-looking, independently moving objects (classically circles) for a fixed amount of time. When the time is up, the objects stop and the observer needs to select the items that belonged to the target subset. This paradigm led to a number of insights into the mechanisms underlying visual attention and visual working memory (VWM) (Cavanagh & Alvarez, 2005; Meyerhoff, Papenmeier, & Huff, 2017; Pylyshyn, 2001; Yantis, 1992).

But how well can one track the *identities* of moving objects? To answer this question, variants of the MOT paradigm were developed. Pylyshyn (2004) placed numerals on the target subset before they started moving and found that although observers were able to subsequently pick out the members of the subset, they showed poor performance when they were asked to recall their numbers. Oksama and Hyönä (2004, 2008) introduced the term *multiple identity tracking* (MIT) for their variant of this paradigm, where instead of circles they used line drawings of everyday objects. In their version, object identities were visible throughout the movement phase, disappearing just before responses were required. Horowitz et al. (2007) used cartoon animals in their MIT task with simpler translational movements. In all cases, there was a consistent "content deficit" in that fewer items could be tracked when participants needed to report the location of targets with a particular identity (where did the *rabbit* go?) than when they needed to report just the locations of all targets in the set (where did the four animals go?).

Although we situated the current study in an MIT framework, it is also related to the literature on the development of VWM (for a recent review, see Fitch, Smith, Guillory, & Kaldy, 2016). VWM has been identified as a fundamental (but not sufficient) process underlying MOT (Drew, Horowitz, Wolfe, & Vogel, 2011; Fougnie & Marois, 2006; Meyerhoff et al., 2017), and in MIT memory demands are even higher because participants need to encode and dynamically maintain identity–location bindings. Robust interindividual correlations have been found between MIT performance and standardized measures of VWM (Oksama & Hyönä, 2004). In addition, Makovski and Jiang (2009) demonstrated that tracking performance in a MOT task was better when all objects were unique in color. (This at first seems to be at odds with the content deficit effect. However, in this paradigm, participants did not need to report the identities of targets, only their locations, after the tracking period. Indeed, when participants needed to hold a set of colors in their VWM in a dual-task design, this eliminated the advantage of tracking uniquely colored objects.)

Although MIT has an established literature in adults, little is known about its developmental trajectory. In the next section, we review what we do know about infants' emerging skills for remembering object identities (in static occlusion tasks) and keeping track of objects while hidden.[1]

The development of identity tracking

Object permanence emerges very early in development (Baillargeon, 1987; Baillargeon, Spelke, & Wasserman, 1985). By 5 months of age, infants can encode a hidden object's location in VWM and will look longer when it is retrieved from a different location. However, evidence for the ability to track an object's identity under 6 months of age is mixed (Newcombe, Huttenlocher, & Learmonth, 1999; Stavans & Baillargeon, 2018). At 6 months, infants can successfully encode two objects and will notice if one object is missing when occluders are removed, but they do not notice if one object has changed

[1] Kibbe (2015) drew an important distinction between object-based and feature-based representations and the two different classes of infant VWM paradigms that aimed at studying them. Here, we focused on findings from paradigms that tested object-based representations, where infants need to encode and maintain visual information over several seconds in naturalistic occlusion situations.

identity (Kibbe & Leslie, 2011). By 9 months of age, infants can remember not only the location of a hidden object but also its identity—that is, *what was where*—and will look longer if two objects were not in their original locations after taking away the occluders (Kaldy & Leslie, 2005). Later, at 12 months of age, infants can succeed at encoding three objects and their respective hiding locations (Kibbe & Leslie, 2013). Taken together, starting early in the first year of life, infants show a gradual increase in the ability to remember the identities of occluded objects.

Piaget's invisible displacement task (1954) was the first to test infants' ability to track an object changing location while occluded. In this task, infants were presented with multiple hiding locations and a target was placed in one of them and then transferred (while occluded) to another location, after which infants were allowed to search for it. Follow-up studies found that 18-month-olds could track an object's change in location (when experimenters moved the object while hidden in their hands), whereas 12- and 15-month-olds could not (Corrigan, 1981; Somerville & Haake, 1985). In addition, 20-month-olds demonstrated fewer perseverative errors (searching for the object in the original hidden place) than 15-month-olds (Sophian & Sage, 1983; Wiebe, Lukowski, & Bauer, 2010). Similarly, in a location-updating task, perseverative errors occurred when 23-month-olds encoded the initial location of an object through direct observation, testimony, or both, suggesting that information about the previously encoded location may have interfered with the information about the new location. It was only at 30 months of age that infants were able to resolve the conflicting information (Ganea & Harris, 2013). Taken together, evidence in invisible displacement tasks and verbal updating tasks suggests that the ability to track an occluded object gradually emerges between 18 and 30 months of age.

These tasks, however, only tested infants' ability to track one object without testing their ability to concomitantly track its *identity*. Despite its importance in everyday tasks, there are only a handful of studies on infants' ability to track identity. Richardson and Kirkham (2004) were the first to test infants' ability to track objects' identities while hidden under moving occluders. In this influential study, they showed that after 6-month-olds learned an association between two sounds and two location placeholders, they were able to successfully track the placeholders as they changed locations, looking to the correct updated location upon hearing the associated sound. Follow-up studies (Kirkham, Richardson, Wu, & Johnson, 2012) described the developmental shift of using multiple cues across perceptual domains to track objects and locations from 3 to 10 months of age. It is important to note that in this paradigm infants first learned the (fixed) sound–object associations over 48 s of training; therefore, during test trials they could rely on long-term (recognition) memory for the initial identity–location bindings. Subsequently, VWM was required only for tracking the location of placeholders as they moved. In our task, however, even the specific identity–location association needed to be encoded and maintained in VWM on each trial ("*this time,* the swirl is on the face of the left card and the watermelon slices are on the face of the right card") based on a brief encoding interval (1–2 s) *and* these links needed to be tracked during occlusion and recalled during test. Trial by trial, infants needed to clear the contents of VWM so as to encode new bindings of identity and location (for a similar argument, see Reznick, 2008, and Kaldy & Leslie, 2005).

Overview of the current study

The current study aimed to examine infants' ability to track the identities of moving objects while occluded. To test this, we adapted the delayed match retrieval (DMR) eye-tracking paradigm (Kaldy, Guillory, & Blaser, 2016). In test trials, three face-down virtual "cards" were presented. Two of them (the Match and Non-match) were flipped over sequentially to expose the faces of the cards and then were flipped back face down. Next, the third card—the Sample—was flipped to reveal a match to one of the previously exposed (now face down) cards (i.e., the Match). Infants were to make an anticipatory saccade to the (face-down) Match card. Previously, we showed that both 10- and 13-month-olds could succeed in this task with both simple geometric shapes (Kaldy et al., 2016) and familiar objects (Cheng, Kaldy, & Blaser, 2019). In the current study, we introduced movement conditions to the face-down Match and Non-match cards before the Sample was revealed. To be successful here then, infants needed to track the identities of the cards as they shifted positions.

According to Piaget's classic description and recent studies, infants over 18 months of age can typically pass the invisible displacement task with one object (Diamond, Prevor, Callender, & Druin, 1997;

Piaget, 1954; Somerville & Haake, 1985; Sophian & Sage, 1983; Wiebe et al., 2010). Because our task requires tracking the identity of *multiple* objects, infants might not be able to succeed until a later age, so we tested 20- and 25-month-olds. By this age, infants are capable of silently labeling simple familiar objects (Mani & Plunkett, 2010); thus, they may attempt verbal rehearsal. To minimize this, similarly to Oksama and Hyönä (2008), we used complex non-nameable stimuli (see Fig. 1A).

In Experiment 1, we tested 20-month-olds in a standard "no movement" DMR task (Fig. 1B). Based on our previous findings, we hypothesized that by this age infants could remember two static identity–location bindings even with abstract non-nameable stimuli. Then, in Experiment 2, we introduced a relatively simple "translational" movement of the cards during a *retention phase* (Fig. 1C). In Experiment 3, we introduced a more complex "shuffle" movement, where the two cards simultaneously moved on orthogonal paths, with one of them crossing the midline (one of the face-down cards moved to a new location while the other card took the position that the first one had occupied) (Fig. 1D). We hypothesized that identity tracking emerges during the end of the second year of life, producing an age-related increase in performance between 20 and 25 months. Because we hypothesized that MIT is still developing over this age range, we further expected that participants would perform better with the relatively simple translational motion of Experiment 2 as compared with the complex shuffle movement of Experiment 3 (this experiment further tested the possibility that moving an object into a previously occupied location may interfere with encoding as compared with an object moved to an entirely new location).[2]

Experiment 1: DMR with no movement

Method

Participants

A total of 22 healthy full-term infants (11 girls) were recruited from the greater Boston area in the northeastern United States and tested at the University of Massachusetts Boston (mean age = 19.3 months, *SD* = 1.42, range = 17.70–22.76). This sample size was based on a power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) conducted for a one-sample *t* test to be able to detect a minimum *d* = 0.65 effect size (80% power, alpha = .05) and matched the sample size used in a previous DMR study from our lab (Cheng et al., 2019). To be included in the analyses, each participant needed to complete at least 3 trials, where they fixated each of the two to-be-tracked cards during the *encoding phase* and one of the (face-down) cards during the *response phase.* We used a minimum of 3 trials as an inclusion criterion to err on the side of inclusivity and to be consistent with the criteria used in our previous study (Cheng et al., 2019). An additional 5 infants were tested but excluded because they did not meet this criterion. Caregivers received a small gift and $20 compensation for their time and travel expenses. All caregivers gave informed consent before the experiment.

Apparatus and stimuli

Caregivers sat in chairs holding their infants in their laps in front of a Tobii T120 eye-tracker (Tobii Technology, Stockholm, Sweden) in a dimly lit testing room. Caregivers were asked to wear a visor to cover their eyes and not to interact with their infants during testing. The standard Tobii 5-point infant calibration was used. As described in the Introduction, experimental stimuli were virtual cards that could be shown face up, revealing an unfamiliar abstract object, or face down, obscuring it. We used a total of four different objects as card faces (Fig. 1A). Trial by trial, two different cards were chosen randomly as the Match and Non-match cards, whereas the Sample card had an identical image to the Match card. Cards subtended $5^\circ \times 5^\circ$ and were arranged symmetrically with their centers 5° from the center of the screen.

[2] Preliminary results from this study were presented in Cheng, Dhungana, Kaldy ,& Blaser (2018). An earlier pilot study with similar results was presented in O'Grady, Guillory, Blaser, & Kaldy (2015).

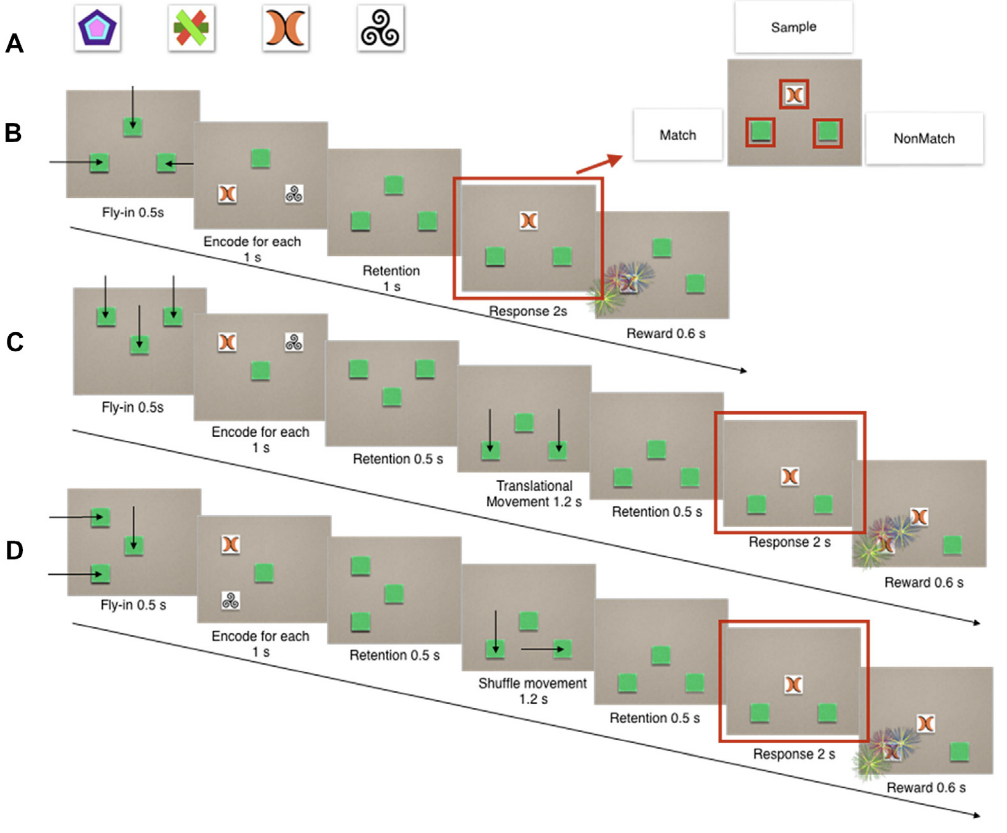

Fig. 1. Stimuli and sequence of test events in a typical trial. (A) Virtual cards, shown face up, exposing to-be-tracked objects. In each trial, two of the four cards were chosen randomly to be the Match and Non-match. (B–D) Sequences of test events in Experiment 1 (B), Experiment 2 (C), and Experiment 3 (D). Note that in Experiment 3, during the delay one of the two target cards moved to a location that had been previously occupied ("old" location) and the other card moved to a location that had not been previously occupied ("new" location). The Match card could be either one of the two cards. In half of the trials the Match ended up in an old location (as in Panel B), and in the other half the Match ended up in a new location. Anticipatory gaze responses were analyzed during the 2-s response phase (see red frame), where performance was measured as the percentage of trials where the infants' first look was to the Match card. Areas of interest for the Match, Non-match, and Sample cards are shown overlaid on the frame in the upper right corner.

Design and procedure

Infants were first presented with four brief familiarization sequences (10 s in total), during which two face-up matching cards entered the screen, approached each other, jiggled, and then exited together. This sequence was repeated once for each of the four object types.

In test trials, three cards (the Match, Non-match, and Sample) entered face down from the side of the screen and formed a triangular arrangement near the center. The Match and Non-match cards were always at the bottom corners, with the Sample card on top. During the encoding phase, the Match and Non-match cards were flipped face up (an animation that took ~500 ms) sequentially; first, one of the two cards, randomly chosen, was exposed, and then, after 1 s, the other card was flipped face up. The two cards stayed face up for 1 s, and then both were flipped face down simultaneously. After that, the Sample card, which had an identical image to the Match card, was flipped face up. The reveal of the Sample marked the beginning of the 2-s response phase, during which anticipatory gaze responses were recorded. After this phase, the Match card was immediately flipped face up, accompanied by a brief (~800 ms) reward animation (e.g., a colorful burst of fireworks) at its location. This reward was designed both to provide feedback about the location of the Match and to encourage participants to make an anticipatory saccade to the Match location so as not to miss the brief reward. The Match then moved to the Sample and touched it, and then all three cards flew off screen (Fig. 1B). To maintain infants' engagement, we added unique sounds to each of the card movements (e.g., flipping, moving, touching) and alternated among three different reward animations (fireworks, sparkles, and flashbulbs). (See the Experiment 1 demo in the online supplementary material for a demonstration of the test event sequence.)

A total of 12 test trials were presented. Card identity, the order of the cards being exposed, the side of the Match (left/right), and the reward animation type were counterbalanced across trials. The counterbalancing of the first three factors ensured that infants needed to keep track of the object identities in each trial. Each test trial was followed by an "attention-grabber" sequence (a cartoon sun rotating in the center of the screen for 4 s with a sound effect) to attract infants' gaze toward the screen.

Data analysis

Each card was bounded by a $7^\circ \times 7^\circ$ area of interest (AOI). (The area was slightly larger than the cards, which subtended $5^\circ \times 5^\circ$, to account for minor calibration errors.) During the response phase, participants should be motivated to look at the Match card as quickly and accurately as possible—and to linger there—in anticipation of the reward animation. We calculated VWM performance as the percentage of correct responses based on the dependent variable of their *first looks,*[3] that is, which of the two face-down cards, Match or Non-match, was fixated first during the response phase (i.e., after the Sample card was revealed). We also calculated the percentage of correct responses based on *longer looks,* that is, which of the two face-down cards accumulated longer looking time during the response phase. Because the response phase was only 2 s, on most trials the card that participants fixated first was also the one fixated longer, so these two variables are highly correlated. In analyzing both of these measures, we followed previously established practices (e.g., Addyman & Mareschal, 2010; Hochmann, Mody, & Carey, 2016; Kaldy et al., 2016). If an infant did not look at either of the two cards during the response phase, the trial could not be coded and was excluded from further analysis. Analyses of gaze data were done using custom MATLAB scripts.

Results

Overall performance

Participants contributed an average of 8.9 valid trials (SD = 2.5) out of 12 trials. Participants' average performance was 56% correct (SD = .12), which was significantly above chance (50%) according to a one-sample t test: $t(21) = 2.50$, $p = .021$, confidence interval (CI) = [0.51, 0.62], $d = 1.09$ (t tests were two tailed throughout all analyses) (Fig. 2). Infants' average performance based on which of the two cards (Match or Non-match) garnered a *longer look* showed similar results (M = 56%, SD = .12), $t(21) = 2.41$, $p = .025$, CI = [0.51, 0.62], $d = 1.05$. We tested the effects of the side of the Match card (left/right) and whether the Match was presented first or second using chi-square tests (treating each trial as an independent observation). Neither of these was significant (all χ^2s < 1.17, $p > .30$).

Learning effect: Analysis of trial-by-trial performance

It is also important to note that besides tracking two object identities, successful performance in our task required encoding the matching rule, namely that anticipatory looking toward the Match card during the response phase will result in being able to catch the short reward animation. Although this does not seem to be a performance-limiting factor in older infants (Kaldy et al., 2016), it is possible that it may interact with the increased demands of the memory task itself. Therefore, in each of our experiments, we also analyzed how performance changed over the block of trials. We performed a lin-

[3] Gaze position for each eye is collected at 60 Hz and averaged between the two eyes to reduce noise. Then, missing data are interpolated (if the gap is <100 ms). Following this, velocity peaks, within a sliding temporal window (of 5 frames @ 60 Hz, i.e., 83.8 ms), are identified. If a peak exceeds a set threshold value (0.42 pixels/ms), it is recorded as a new fixation (but if the distance between two candidate fixations is <35 pixels [<1°], they are merged). The duration of a fixation, then, is the elapsed time between peaks, and the position of the fixation is the median of the gaze coordinates during that interval (Tobii Technology, n.d.).

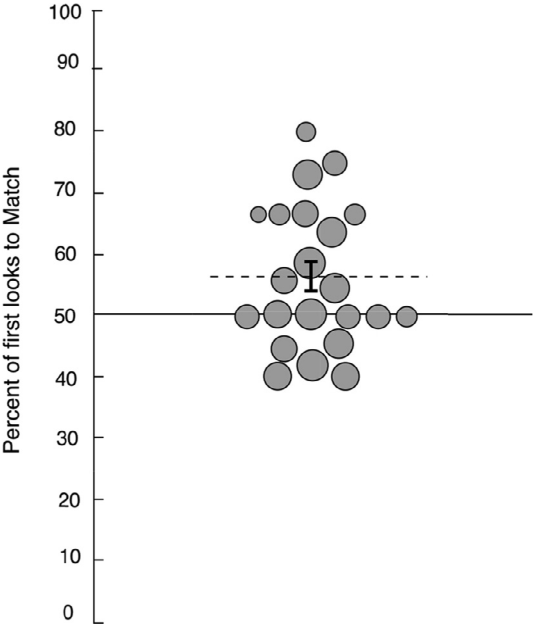

Fig. 2. Results of Experiment 1 (DMR with no movement). Individual and group average performances (percentage correct responses based on first looks) are shown. The size of each circle corresponds to the number of valid trials the infant contributed. Error bars reflect standard errors of the mean.

ear regression between average task performance in each trial and trial number (1–12). We did not find a significant effect of learning over trials (R^2 = .04, p = .54 based on first looks; R^2 = .01, p = .85 based on longer looks), suggesting that 20-month-olds' performance did not change linearly over the block of trials.

Experiment 2: DMR with translational movement

Experiment 1 showed above-chance performance in 20-month-old infants, suggesting that they could keep track of two (static) object identities during occlusion even when the identities were abstract unfamiliar shapes. In the next two experiments, we explored whether infants of the same age or older (25 months) could keep track of identity while the objects moved during occlusion. Here, during the retention phase, to-be-tracked cards underwent translational movement, from the top to the bottom of the screen, preserving their relative left/right positions.

Method

Participants

A total of 44 full-term healthy infants (17 girls) were recruited from the greater Boston area and tested at the University of Massachusetts Boston. The infants were recruited in the same way as in Experiment 1. We tested two age groups with 22 infants in each: 20-month-olds (M = 19.8 months, SD = 1.06, range = 18.13–21.73) and 25-month-olds (M = 24.8 months, SD = 1.36, range = 22.20–26.90). An additional 6 infants (4 in the 20-month-old group and 2 in the 25-month-old group) were tested but excluded due to an insufficient number of valid trials (i.e., <3, with trial inclusion criteria as defined in Experiment 1).

Apparatus and stimuli

The same apparatus and stimuli were used as in Experiment 1.

Design and procedure

Infants were first presented with the same familiarization trials as in Experiment 1. For test trials, the only substantive difference from Experiment 1 was that (face-down) cards moved during the retention phase. In each trial, three cards (the Match, Non-match, and Sample) entered from the side of the screen face down and formed a triangular arrangement near the upper side of the screen. The Match and Non-match were at the top corners of the triangle, and the Sample was always in the center of the screen. During the encoding phase of the trial, the Match and Non-match were flipped face up sequentially. The two cards stayed face up for 1 s, and then both were flipped face down at the same time. After that, the two face-down cards moved down to the bottom of the screen (1.2 s), and the Sample was exposed. The response phase (2 s) and the subsequent reward animations were the same as in Experiment 1. (See Fig. 1C and Experiment 2 demo in the supplementary material).

Data analysis

Data analysis was the same as in Experiment 1.

Results

Overall performance

Participants on average contributed 7.2 valid trials ($SD = 2.8$) in the 20-month-old group and 6.6 valid trials ($SD = 3.0$) in the 25-month-old group (out of 12 trials). There was no significant difference in the number of valid trials between the two age groups [two-sample t test: $t(42) = 0.68$, $p = .50$].

The 20-month-old participants' average performance based on first looks was 50% correct ($SD = .16$), which was not different from chance [one-sample t test: $t(21) = -0.11$, $p = .92$, CI = [0.42, 0.57], $d = 0.05$] (see Fig. 3). The 25-month-olds' performance, however, was significantly above chance at 61% correct ($SD = .18$) based on first looks to the Match card, $t(21) = 2.78$, $p = .011$, CI = [0.53, 0.69], $d = 1.21$. The same results were found based on longer looks to the Match card, with 20-month-olds again at chance ($M = 49\%$, $SD = .17$), $t(21) = -0.21$, $p = .84$, CI = [0.42, 0.57], $d = 0.09$, and 25-month-olds showing significantly above-chance performance ($M = 59\%$, $SD = .19$), $t(21) = 2.25$, $p = .035$, CI = [0.51, 0.67], $d = 1.21$.

We tested the effects of the side of the Match card (left/right) and whether the Match was presented first or second during encoding. Infants responded similarly when the Match card was on the left or right in both age groups (all χ^2s < 0.15, $ps > .69$). We did not find any differences in performance when the Match card was first versus second shown in either of the two age groups (all χ^2s < 0.26, $ps > .61$).

Learning effect: Analysis of trial-by-trial performance

To examine the potential effect of learning in 20-month-olds, we compared differences in task performance over trials. We performed a linear regression between average task performance (based on first looks) and trial number (1–12), and we found a moderate, marginally significant moderate positive trend indicating a learning effect over trials ($R^2 = .325$, $p = .052$). Task performance based on longer looks did not show this trend over trials ($R^2 = .13$, $p = .24$). For 25-month-olds, the result of the linear regression analysis between trial number and percentage of correct responses was not significant ($R^2 = .16$, $p = .19$ based on first looks; $R^2 = .02$, $p = .69$ based on longer looks).

Experiment 3: DMR with shuffle movement

In this experiment, during the retention phase, to-be-tracked cards underwent a more complex shuffle movement, shifting positions along orthogonal trajectories, with one of the two cards crossing the midline.

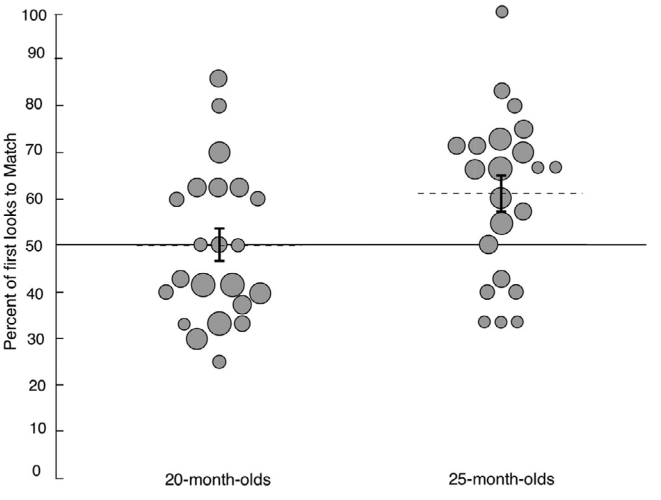

Fig. 3. Results of Experiment 2 (DMR with translational movement). Individual and group average performances (percentage correct responses based on first looks) are shown. The size of each circle corresponds to the number of valid trials the infant contributed. Error bars reflect standard errors of the mean.

Method

Participants

A total of 44 full-term healthy infants (24 girls) were recruited from the greater Boston area and tested at the University of Massachusetts Boston. They were assigned to two age groups with 22 infants in each: 20-month-olds (M = 19.4 months, SD = 1.24, range = 17.83–21.77) and 25-month-olds (M = 25.0 months, SD = 1.80, range = 22.4–27.97). An additional 7 infants (3 in the 20-month-old group and 4 in the 25-month-old group) were tested but excluded due to an insufficient number of valid trials (i.e., <3, with trial inclusion criteria as defined in Experiment 1). The infants were recruited in the same way as in Experiment 1.

Apparatus and stimuli

The same apparatus and stimuli were used as in Experiment 1.

Design and procedure

Infants were first presented with same familiarization trials as in Experiment 1. For test trials, the only substantive difference compared with the test events of Experiment 1 was that the cards moved during the retention phase. In each test trial, the Match, Non-match, and Sample cards entered either from the left or right side of the screen face down and formed a triangular arrangement near either the left or right edge of the screen. The Match and Non-match were at the side corners of the triangle, and the Sample was always in the center of the screen. The timing of the encoding phase was identical to that in Experiment 1. The two cards stayed face up for 1 s, and then both were flipped face down at the same time. The Match and Non-match then moved clockwise (when cards started on the right side) or counterclockwise (when cards started on the left side) to the bottom of the screen to become horizontally aligned. In this way, the top card moved down to the bottom of the screen while the bottom card moved sideways to the opposite side of the screen in two simultaneous orthogonal movements. Movement unfolded over 1.2 s. (The timing of the events during retention was the same as in Experiment 2). After that, the Sample card was flipped face up. Events in the following response and reward phases were identical to those in Experiment 1. (See Fig. 1D and Experiment 3 demo in supplementary material).

Data analysis

Data analysis was the same as in Experiment 1.

Results

Overall performance

On average, 20-month-olds contributed 6.9 valid trials ($SD = 2.2$) and 25-month-olds contributed 8.1 valid trials ($SD = 2.9$) out of 12 trials. There was no significant difference in the number of valid trials between the two age groups [two-sample t test: $t(42) = 1.56$, $p = .13$].

The 20-month-olds' average performance was not different from chance based on the percentage of first looks (one-sample t test: $M = 48\%$, $SD = .20$), $t(21) = -0.50$, $p = .62$, CI = [0.39, 0.57], $d = 0.21$, or longer looks ($M = 47\%$, $SD = .19$), $t(21) = -0.72$, $p = .48$, CI = [0.39, 0.56], $d = 0.31$. The 25-month-olds' average performance over all 12 trials was 57% correct ($SD = .20$) based on first looks, which, although not significantly different from chance, $t(21) = 1.59$, $p = .13$, CI = [0.48, 0.65], showed a medium effect size ($d = 0.69$) (see Fig. 4). Infants' average performance based on which of the cards (Match or Non-match) garnered the longer look was significantly better than chance ($M = 58\%$, $SD = .19$), $t(21) = 2.114$, $p = .046$, CI = [0.50, 0.67], $d = 0.92$.

Learning effect: Trial-by-trial analysis

As in the previous two experiments, we conducted a linear regression between performance based on first looks in a particular trial and the trial number. We did not see a significant relationship in either of the two age groups in Experiment 3 (in 20-month-olds: $R^2 = .09$, $p = .35$ based on first looks, $R^2 = .13$, $p = .25$ based on longer looks; in 25-month-olds: $R^2 = .17$, $p = .19$ based on first looks, $R^2 = .06$, $p = .46$ based on longer looks). In sum, infants in Experiment 3 did not show a significant learning effect over trials.

Interference when updating information to a new location versus an old location

Updating remembered information at a previously encoded location may cause interference, a failure to update representations based on new information (Ganea & Harris, 2013). To explore this, we compared infants' average performance when the Match card moved to a "new" (previously unoccupied) location with their average performance when the Match card moved to an "old" (previously occupied) location. We did not find a significant difference in performance based on first looks in either of the two age groups [20-month-olds: $M_{Old\_Location} = 49\%$, $M_{New\_Location} = 50\%$, $t(20) = 0.08$, $p = .94$ (paired t test); 25-month-olds: $M_{Old\_Location} = 59\%$, $M_{New\_Location} = 55\%$, $t(21) = 0.49$, $p = .63$]. Performance based on longer looks showed the same pattern [20-month-olds: $M_{Old\_Location} = 51\%$, $M_{New\_Location} = 47\%$, $t(20) = 0.61$, $p = .55$ (paired t test); 25-month-olds: $M_{Old\_Location} = 60\%$, $M_{New\_Location} = 57\%$, $t(21) = 0.45$, $p = .66$].

We also tested the effects of the side of the Match card (left/right) and whether the Match was presented first or second. Neither of the effects was significant in either of the two age groups (all χ^2s < 3.28, $p > .07$).

The development of location updating performance

Given the similarities between the location updating tasks in Experiments 2 and 3, we tested the effect of task difficulty and age on performance (based on first looks). We performed a univariate 2×2 analysis of variance (ANOVA) with age group (20-month-olds or 25-month-olds) and experiment (Experiment 2 [translational movement] or Experiment 3 [shuffle movement]). This analysis showed no significant effect of experiment, $F(1, 86) = 0.293$, $p = .59$, $\eta^2 = .004$, but showed a significant main effect of age group, $F(1, 86) = 5.337$, $p = .023$, $\eta^2 = .06$. The interaction between age group and experiment was not significant, $F(1, 83) = 0.243$, $p = .623$, $\eta^2 = .003$.

Summary of results

See Table 1 and Fig. 5 for summary of results.

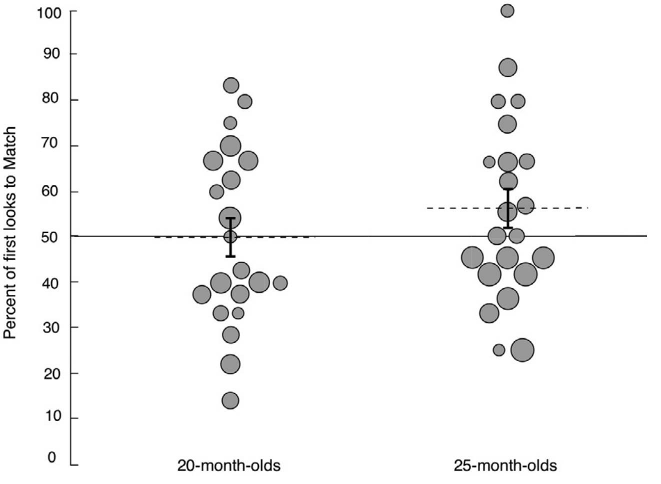

Fig. 4. Results of Experiment 3 (DMR with shuffle movement). Individual and group average performances (percentage correct responses based on first looks) are shown. The size of each circle corresponds to the number of valid trials the infant contributed. Error bars reflect standard error of the mean.

Table 1
Summary of results across three experiments.

Experiment	*N* or *n*	Age in months (*SD*)	Valid trials, out of 12 (*SD*)	Performance based on first looks [CI]	Performance based on longer looks [CI]
Experiment 1: No movement	22	19.3 (1.42)	8.9 (2.5)	56% (*) [0.51, 0.62]	56% (*) [0.51, 0.62]
Experiment 2: Translational (20-month-olds)	22	19.8 (1.06)	7.2 (2.8)	50% (n.s.) [0.42, 0.57]	49% (n.s.) [0.42, 0.57]
Experiment 2: Translational (25-month-olds)	22	24.8 (1.36)	6.6 (3.0)	61% (*) [0.53, 0.69]	59% (*) [0.51, 0.67]
Experiment 3: Shuffle (20-month-olds)	22	19.4 (1.24)	6.9 (2.2)	48% (n.s.) [0.39, 0.56]	47% (n.s.) [0.39, 0.57]
Experiment 3: Shuffle (25-month-olds)	22	25.0 (1.80)	8.1 (2.9)	57% (n.s.) [0.48, 0.65]	58% (*) [0.50, 0.67]

Note. CI, confidence interval.
(*) $p < .05$.
(n.s.) not significant.

General discussion

Summary of results

The current study examined infants' MIT abilities. We used a modified DMR task, where virtual cards were shown face up with to-be-tracked objects on their faces and then turned face down. After a retention phase, a "sample" card was revealed that matched one of the two previously seen cards, and infants were expected to find (i.e., make an anticipatory saccade to) the matching card. In Experiment 1, we replicated previous results using DMR (Cheng et al., 2019; Kaldy et al., 2016), showing that infants were able to remember the identities of two objects without movement. We then examined infants' MIT ability as cards moved during the retention phase. Cards either had translational movement (Experiment 2), when perceptual grouping could help tracking, or (more challengingly) shuffled to new locations (Experiment 3), which involved orthogonal trajectories.

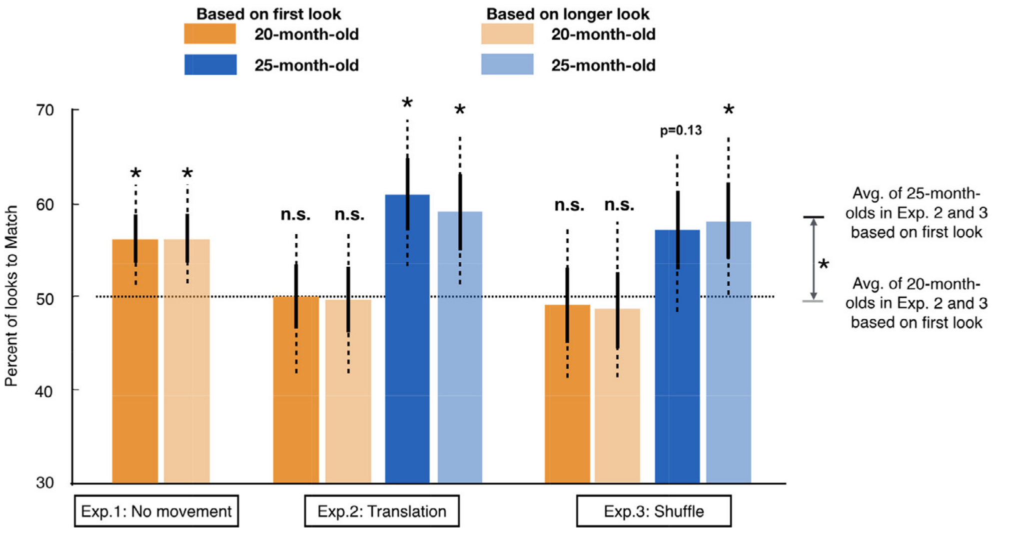

Fig. 5. Summary of average VWM performance across three experiments. Percentage correct performances based on first looks (darker bars) and percentage correct performances based on longer looks (lighter bars) in Experiment 1, 2, and 3 are shown. Error bars indicate standard errors of the mean, and whiskers indicate 95% confidence intervals. $^{*}p < .05$; n.s., not significant.

Overall, 20-month-olds showed at-chance performance on the task when movement was involved, whereas 25-month-olds performed significantly better than chance even with translational or shuffle movement.[4] Although the overall ANOVA did not show a significant effect of movement type, the effect size in the 25-month-olds was larger with translational movement (Experiment 2: *d*s = 1.21 and 1.21 for results based on first looks and longer looks, respectively) than with shuffle movement (Experiment 3: *d*s = 0.69 and 0.92, respectively). We also analyzed how infants' performance changed over trials. We did not expect a learning effect in the conditions that showed above-chance performance overall (20-month-olds in Experiment 1 and 25-month-olds in Experiments 2 and 3). However, with the 20-month-olds in Experiment 2, there was some indication that success was not entirely out of reach with the simpler translational movement given that these younger infants showed a learning trend ($R^2 = .325$) over the block of 12 trials.[5]

The cognitive demands of multiple identity tracking

Keeping the identities of two (static) objects in VWM is within the abilities of infants by 20 months of age. Beyond the current findings in Experiment 1, this has been shown in infants under 1 year of age in paradigms with widely different task demands by Feigenson and Carey (2003), Kaldy and Leslie (2003), and Ross-Sheehy, Oakes, and Luck (2003). It is clear, however, that *tracking* those identities once movement is introduced increases cognitive demands. Several adult studies have measured

[4] In general, performance on this sort of task is limited by lapses in task understanding, engagement, and response execution common to infant populations. Previous studies with DMR have shown similar performance levels (e.g., 62% correct in 10-month-olds [Kaldy et al., 2016]; 58% correct in 13-month-olds [Cheng et al., 2019]; 59% in 14-month-olds [Hochmann et al., 2016]). This is a general limitation of infant looking time paradigms because performance in two-alternative forced-choice tasks is typically under 65% (e.g., McMurray & Aslin, 2004; Kwon, Luck, & Oakes, 2014; Oakes, Baumgartner, Barrett, Messenger, & Luck, 2013, Vlach & Johnson, 2013).

[5] There is another potential measure of learning—the change in average fixation latency to the Match card (when the first look is to the Match) during the response period over trials. However, this measure did not show any systematic trends; throughout the three experiments, correlations between the rank number of trials and fixation latency to the Match were not significant in any of the groups (Experiment 1: $R^2 = .003$, $p = .86$; Experiment 2: $R^2 = .004$, $p = .85$ in 20-month-olds, $R^2 = .004$, $p = .85$ in 25-month-olds; Experiment 3: $R^2 = .21$, $p = .13$ in 20-month-olds, $R^2 = .08$, $p = .37$ in 25-month-olds).

how much more challenging MIT tasks are compared with MOT tasks, that is, the content deficit (Horowitz et al., 2007). Hollingworth and Rasmussen (2010) presented a surprising finding with adults in a paradigm that was similar to the current study. They contrasted VWM performance in two conditions when two colored squares swapped positions during a delay. For example, participants saw a red square in a 5 o'clock position and a green square in an 11 o'clock position. Then the colors disappeared and the square outlines moved along a circular path such that each ended up taking the other square's position. Participants were then asked to identify the colors in the updated versus original positions. They found that, contrary to the predictions of theories of MOT, binding of color to the *original* locations of the objects was stronger (adults reported the objects' color faster and more accurately) than binding to the *updated* locations despite clear visual evidence that the objects had moved. This shows that encoding of identity–location bindings at the original position was more robust and that location updating reduced the robustness of memory for identity in adults. This effect may contribute to the lower performance of 20-month-olds versus 25-month-olds in Experiments 2 and 3 of the current study.

In terms of the complexity of the required location updating, we contrasted two types of movements. Infants viewing the translational motion of Experiment 2 could potentially exploit perceptual grouping to reduce cognitive demands just like adults (Woodman, Vecera, & Luck, 2003; Yantis, 1992). Previous studies have also demonstrated that infants during the second year of life can use cues (e.g., perceptual/conceptual similarity) to help them track more items (Feigenson & Halberda, 2008; Rosenberg & Feigenson, 2013). In contrast, the shuffle movement of Experiment 3 likely required infants to track two trajectories separately. Adult MOT studies found that more changes in target trajectories impaired tracking performance (Ericson & Beck, 2013). Although we did not see a main effect of movement type in the ANOVA, there were hints that tracking of the shuffle movement was less robust. Results in that experiment were mixed, with above-chance performance found only in the total look duration measure but not in the first look measure. Further work is required to investigate this finding, but we speculate that occasionally participants who were less certain about the location of updated bindings made an initial (likely random) first look (thereby diminishing first look performance), followed by a corrective look to the other location.

The brain mechanisms underlying identity tracking

Neuroimaging studies with children in an MIT task have not yet been conducted. However, several neuroimaging studies have investigated the cortical mechanisms underlying MIT performance in adults. Despite differences in the behavioral paradigms (in one object identities were constantly changing during tracking, and in the other objects rotated behind occluders), two independent functional magnetic resonance imaging studies concluded that activation in frontal regions (in the inferior precentral sulcus) and posterior parietal areas (intraparietal sulcus and superior parietal lobule) were responsible for MIT (Lyu, Hu, Wei, Zhang, & Talhelm, 2015; Takahama, Miyauchi, & Saiki, 2010). Another recent study directly contrasted brain activity patterns underlying MOT and MIT tasks (Nummenmaa, Oksama, Glerean, & Hyönä, 2017). There, although both tasks activated the same extended frontoparietal circuits identified above, in the MIT tasks there also was an additional load-dependent activity increase in the lateral prefrontal cortex and ventral temporal areas known to subserve object recognition and VWM. Given this, we would hypothesize that it is the maturation of the network described by Nummenmaa and colleagues, over the course of the second year of life, that accounts for older infants' better performance in our task.

Limitations and future directions

Our study is the first, to our knowledge, to study MIT in infants. In that context, it is important to note limitations of our study and places for follow-up work. First, by adding the movement trajectory in Experiments 2 and 3, we also lengthened the retention interval relative to Experiment 1 (from 1 to 2.2 s). A few prior studies have looked at the effect of the length of delay on VWM performance in infants and children, although not with 1.5- to 2-year-old infants. Kaldy and Leslie (2005) found that 6-month-old infants' VWM performance was not affected when the delay was increased from 4 to 7 s (with one object) or, at the same age, from 3 to 5 s (with two locations) (O'Gilmore & Johnson, 1995); similarly, 7-year-old children's visual short-term memory performance was not affected by a change from 1.5 to 2.3 s (with up to four objects) (Shimi & Scerif, 2017). Taken together, while longer delays may ultimately have a negative impact on performance, it is unlikely that the amount of the difference here (1 to 2.2 s) had a substantial impact on performance. Furthermore, no results have suggested an interaction effect with age, so any negative effect of an increased delay is unlikely to have created the *relatively* poor performance of the 20-month-olds versus the 25-month-olds in Experiments 2 and 3.

Second, we evaluated the development of MIT by contrasting tasks where tracking was required (the translational and shuffle movements of Experiments 2 and 3, respectively) with one task where objects remained static (Experiment 1). Here, we interpreted older infants' facility with MIT and younger infants' relatively poor performance as evidence that the ability to track identity is developing in this age range. However, it is possible that younger infants' lower performance in Experiments 2 and 3 was not due to identity-tracking challenges per se but instead was due to some distraction caused by the movement itself. There is no prior work that we know of that directly addresses this question, but in general infants' representations of unfamiliar objects can be fragile (Horst & Samuelson, 2008). Future work should investigate this possibility by testing a condition where the cards move inconsequentially (e.g., jiggling in place) or move and return to their original locations.

Finally, it is possible that infants may use a "process of elimination" cognitive strategy to reduce the need to remember both objects. Through disjunctive reasoning (the principle of mutual exclusivity), instead of remembering both to-be-remembered objects (A and B), infants could just remember one (say, A) and then, if the Sample is B, use a "find NOT-A" strategy (Halberda, 2003; Markman, Wasow, & Hansen, 2003) to identify the correct, matching card. A recent study suggested that infants as young as 12 months are able to reason this way (Cesana-Arlotti et al., 2018); however, infants did not spontaneously do it in a DMR task at 14 months of age (Hochmann et al., 2016). Our task was not designed to isolate the use of this specific strategy, and future studies (e.g., with more to-be-tracked objects) should test at what age can children spontaneously apply this strategy in this type of VWM task.

Closing remarks

In addition to the empirical contributions discussed above, our study also makes some important methodological contributions. As demonstrated here, the DMR task can be tailored to different ages in early development (8- and 10-month-olds [Kaldy et al., 2016]; 13-month-olds [Cheng et al., 2019]; and 20- and 25-month-olds [current study]), making it possible to study visual attention and VWM for objects across a wide age range from infancy to early childhood. The task does not require receptive language skills, making it ideal for preverbal populations or children with language delays or deficits (e.g., children with autism spectrum disorder). Furthermore, this task can be easily modified to parametrically study the multiple cognitive components of VWM (i.e., processing speed, sustained attention, inhibitory control) by manipulating task parameters (e.g., encoding time, number of objects, retention duration).

Acknowledgments

This research was supported by National Institutes of Health Grant R15HD086658 awarded to Z.K. and E.B. and by a Dissertation Grant from the University of Massachusetts Boston awarded to C.C. Preliminary results from this study were presented at the annual meeting of the International Congress on Infant Studies in May 2018 (Cheng, Kaldy, Dhungana, & Blaser, 2018). We thank the caregivers and their children who participated in our experiments. We also thank Shaun O'Grady for his contributions to a pilot study for this project (O'Grady, Guillory, Blaser, & Kaldy, 2015), and Sangya Dhungana and other members of the UMass Boston Baby Lab for their help with data collection.

Appendix A. Supplementary material

Please see: https://osf.io/7kx2f/.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2019.06.002.

References

Addyman, C., & Mareschal, D. (2010). The perceptual origins of the abstract same/different concept in human infants. *Animal Cognition, 13*, 817–833.

Baillargeon, R. (1987). Object permanence in 3½- and 4½-month-old infants. *Developmental Psychology, 23*, 655–664.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition, 20*, 191–208.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences, 9*, 349–354.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science, 359*, 1263–1266.

Cheng, C., Kaldy, Z., & Blaser, E. (2019). Focused attention predicts visual working memory performance in 13-month-old infants: A pupillometric study. *Developmental Cognitive Neuroscience, 36* 100616.

Cheng, C., Kaldy, Z., Dhungana, S., & Blaser, E. (2018, June–July). Successful updating of spatial information in visual working memory in 20- and 25-month-olds. Poster presented at the annual meeting of the International Conference on Infant Studies, Philadelphia.

Corrigan, R. (1981). The effects of task and practice on search for invisibly displaced objects. *Developmental Review, 1*(1), 1–17.

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development, 62*. 4, Serial No. 252.

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the neural signatures of tracking spatial position and working memory during attentive tracking. *Journal of Neuroscience, 31*, 659–668.

Ericson, J. M., & Beck, M. R. (2013). Changing target trajectories influences tracking performance. *Psychonomic Bulletin & Review, 20*, 951–956.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science, 6*, 568–584.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 9926–9930.

Fitch, A., Smith, H., Guillory, S. B., & Kaldy, Z. (2016). Off to a good start: The early development of the neural substrates underlying visual working memory. *Frontiers in Systems Neuroscience, 10*. https://doi.org/10.3389/fnsys.2016.00068.

Fougnie, D., & Marois, R. (2006). Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. *Psychological Science, 17*, 526–534.

Ganea, P. A., & Harris, P. L. (2013). Early limits on the verbal updating of an object's location. *Journal of Experimental Child Psychology, 114*, 89–101.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*, B23–B34.

Hochmann, J.-R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match- and non-match-to-sample. *Cognitive Psychology, 86*, 87–111.

Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 543–564.

Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception & Psychophysics, 69*, 172–184.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*, 128–157.

Kaldy, Z., Guillory, S. B., & Blaser, E. (2016). Delayed match retrieval: A novel anticipation-based visual working memory paradigm. *Developmental Science, 19*, 892–900.

Kaldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: Integrating "what" and "where" information. *Developmental Science, 6*, 360–373.

Kaldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition, 97*, 153–177.

Kibbe, M. M. (2015). Varieties of visual working memory representation in infancy and beyond. *Current Directions in Psychological Science, 24*, 433–439.

Kibbe, M. M., & Leslie, A. M. (2011). What do infants remember when they forget? Location and identity in 6-month-olds' memory for objects. *Psychological Science, 22*, 1500–1505.

Kibbe, M. M., & Leslie, A. M. (2013). What's the object of object working memory in infancy? Unraveling "what" and "how many". *Cognitive Psychology, 66*, 380–404.

Kirkham, N. Z., Richardson, D. C., Wu, R., & Johnson, S. P. (2012). The importance of "what": Infants use featural information to index events. *Journal of Experimental Child Psychology, 113*, 430–439.

Kwon, M.-K., Luck, S. J., & Oakes, L. M. (2014). Visual short-term memory for complex objects in 6- and 8-month-old infants. *Child Development, 85*, 564–577.

Lyu, C., Hu, S., Wei, L., Zhang, X., & Talhelm, T. (2015). Brain activation of identity switching in multiple identity tracking task. *PLoS One, 10*(12) e145489.

Makovski, T., & Jiang, Y. V. (2009). The role of visual working memory in attentive tracking of unique objects. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1687–1697.

Mani, N., & Plunkett, K. (2010). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science, 21*, 908–913.

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*, 241–275.

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy, 6*, 203–229.
Meyerhoff, H. S., Papenmeier, F., & Huff, M. (2017). Studying visual attention using the multiple object tracking paradigm: A tutorial review. *Attention, Perception, & Psychophysics, 79*, 1255–1274.
Newcombe, N., Huttenlocher, J., & Learmonth, A. (1999). Infants' coding of location in continuous space. *Infant Behavior and Development, 22*, 483–510.
Nummenmaa, L., Oksama, L., Glerean, E., & Hyönä, J. (2017). Cortical circuit for binding object identity and location during multiple-object tracking. *Cerebral Cortex, 27*, 162–172.
Oakes, L. M., Baumgartner, H. A., Barrett, F. S., Messenger, I. M., & Luck, S. J. (2013). Developmental changes in visual short-term memory in infancy: Evidence from eye-tracking. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00697.
O'Gilmore, R., & Johnson, M. H. (1995). Working memory in infancy: Six-month-olds' performance on two versions of the oculomotor delayed response task. *Journal of Experimental Child Psychology, 59*, 397–418.
O'Grady, S, Guillory, S. B., Blaser, E., & Kaldy, Z. (2015, March). 21-Month-old toddlers pass an anticipatory version of the invisible displacement task. Poster presented at the biennial meeting of the Society for Research in Child Development, Philadelphia.
Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition, 11*, 631–671.
Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive Psychology, 56*, 237–283.
Piaget, J. (1954). Language and thought from a genetic perspective. *Acta Psychologica, 10*, 51–60.
Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*, 127–158.
Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition, 11*, 801–822.
Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179–197.
Reznick, J. S. (2008). Working memory in infants and toddlers. In M. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 355–378). New York: Psychology Press.
Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General, 133*, 46–62.
Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science, 16*, 610–621.
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.
Shimi, A., & Scerif, G. (2017). Towards an integrative model of visual short-term memory maintenance: Evidence from the effects of attentional control, load, decay, and their interactions in childhood. *Cognition, 169*, 61–83.
Somerville, S. C., & Haake, R. J. (1985). The logical search skills of infants and young children. In H. M. Wellman (Ed.), *Children's searching: The development of search skill and spatial representation* (pp. 73–104). Hillsdale, NJ: Lawrence Erlbaum.
Sophian, C., & Sage, S. (1983). Developments in infants' search for displaced objects. *Journal of Experimental Child Psychology, 35*, 143–160.
Stavans, M., & Baillargeon, R. (2018). Four-month-old infants individuate and track simple tools following functional demonstrations. *Developmental Science, 21* e12500.
Takahama, S., Miyauchi, S., & Saiki, J. (2010). Neural basis for dynamic updating of object representation in visual working memory. *NeuroImage, 49*, 3394–3403.
Tobii Technology. (n.d.). Tobii Studio user's manual (Version 3.4.5). Retrieved from https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf.
Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127*, 375–382.
Wiebe, S. A., Lukowski, A. F., & Bauer, P. J. (2010). Sequence imitation and reaching measures of executive control: A longitudinal examination in the second year of life. *Developmental Neuropsychology, 35*, 522–538.
Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review, 10*, 80–87.
Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*, 295–340.