Exogenous attention influences visual short-term memory in infants

Shannon Ross-Sheehy, Lisa M. Oakes and Steven J. Luck

1. Department of Psychology, University of Iowa, USA
2. Department of Psychology, University of California Davis, USA

Abstract

Two experiments examined the hypothesis that developing visual attentional mechanisms influence infants’ Visual Short-Term Memory (VSTM) in the context of multiple items. Five- and 10-month-old infants (N = 76) received a change detection task in which arrays of three differently colored squares appeared and disappeared. On each trial one square changed color and one square was cued; sometimes the cued item was the changing item, and sometimes the changing item was not the cued item. Ten-month-old infants exhibited enhanced memory for the cued item when the cue was a spatial pre-cue (Experiment 1) and 5-month-old infants exhibited enhanced memory for the cued item when the cue was relative motion (Experiment 2). These results demonstrate for the first time that infants younger than 6 months can encode information in VSTM about individual items in multiple-object arrays, and that attention-directing cues influence both perceptual and VSTM encoding of stimuli in infants as in adults.

Introduction

Short-term memory for visual information is necessary for infants to have a coherent experience of the visual world. During the first several months of life, fundamental visual abilities such as acuity, accommodation, oculomotor control, and attention undergo significant development (Atkinson, 1984; Colombo, 2001). These basic visual skills directly influence infants’ ability to integrate information over space and time. For example, the number and accuracy of saccades required to fixate a target changes over the first 6 months (Atkinson, Bradrick & Moar, 1977), and developing cortical and subcortical areas – such as the frontal eye fields and superior colliculus – cause changes in the proportion of automatic or reflexive saccades versus volitional or targeted saccades (M.H. Johnson, 1995a, 1995b; Richards & Hunter, 1998). Thus, infants are constantly adapting to their visual world, as they begin to integrate information that spans brief disruptions caused by eye blinks, saccades, and periods of occlusion.

Visual short-term memory (VSTM) is critical in these processes because it allows infants to form stable representations despite the fact that the visual input is punctuated by disruptions lasting tens or hundreds of milliseconds. In adults, for example, VSTM is critical for integrating information across eye movements (Irwin, 1991) and for perceiving visual stability despite frequent fixation errors (Hollingworth, Richard & Luck, 2008). In infants and adults, VSTM is often assessed using change-detection tasks (e.g. Luck & Vogel, 1997; Ross-Sheehy, Oakes & Luck, 2003). In the infant version of this task, a display is briefly presented (e.g. an array of three colored squares for 500 ms), followed by a brief retention period (300 ms), and then a new array is presented that may or may not contain a change (e.g. one of the squares is now a different color). Infants’ preference for stimulus streams that involve this type of change relative to equivalent stimulus streams in which no change occurs is taken as evidence that infants have encoded properties of the objects in VSTM (see for example: Ross-Sheehy et al., 2003).

Studies using this procedure show rapid increases between 6 and 8 months in infants’ ability to build representations in VSTM. Specifically, 4- to 6-month-old infants can remember the identity of only one object (Ross-Sheehy et al., 2003), and are unable to bind color to location (Oakes, Ross-Sheehy & Luck, 2006). Infants 7 months and older, in contrast, detect changes in arrays of multiple objects whether one or all of the objects changes, and even detect changes in the bindings of identity and location (Oakes et al., 2006). A similar developmental trajectory has been revealed in studies using very different tasks, time courses, and stimuli (but requiring that visual information be stored in working memory) (e.g. Káldy & Leslie, 2003, 2005).

The limitations on younger infants’ VSTM occur exclusively in the context of multiple objects (Káldy &...
Leslie, 2005; Ross-Sheehy et al., 2003). In change-detection tasks, 6-month-old infants do not prefer changing to non-changing stimuli for arrays containing two or three objects (Ross-Sheehy et al., 2003), even when all the items in each array change (Oakes, Messenger, Ross-Sheehy & Luck, 2009; Oakes et al., 2006). For example, Oakes et al. (2009) found that 6-month-old infants fail to prefer changing streams with arrays of three colored squares even though the colors of all three of the squares change color on each cycle (e.g. first blue, red, pink, then yellow, white, orange). Note that in this context detecting which stream is changing should be trivially easy because the changes can be detected even if the observer remembers the color and location of only a single square; thus 6-month-old infants appear not to remember information about any of the items in these three-item arrays. Seven-month-old infants, in contrast, show a robust preference for the change in this situation, indicating a fundamental change in VSTM ability over a period of weeks (Oakes et al., 2009).

What is responsible for this change? Before 6 or 7 months, infants’ VSTM representations may simply lack the fidelity necessary for them to discriminate the color changes in this task. However, 6-month-old infants detect equivalent color changes in this task when there is only a single square in each stream (Ross-Sheehy et al., 2003). Infants also may be unable to encode two or three items in 500 ms. However, adults can encode a single item in VSTM in 50 ms (Vogel, Woodman & Luck, 2006), and Catherwood, Skoien, Green and Holt (1996) found that 4-month-old infants could learn two colors in a 250-ms exposure time. Thus, 6-month-old infants should be able to encode two or three colors in 500 ms. Six-month-old infants’ failure may also reflect immature perceptual processing. However, 6-month-old infants can detect changes in this task in arrays of three objects when no memory is required (Ross-Sheehy et al., 2003) or when the three objects are all the same color (Oakes et al., 2006), and they detect changes in numerosity in arrays that contain large numbers of items (Libertus & Brannon, in press). Moreover, when shown an array of three identically colored objects paired simultaneously with an array of three differently colored objects, 6-month-old infants prefer the heterogeneous displays (Oakes et al., 2006), indicating that they can perceive distinct colors in these multi-item arrays. Thus, developmental differences in VSTM capacity do not seem to reflect a lack of fidelity, speed of processing, or perceptual deficits.

It is more likely that this functional jump in capacity is driven, at least in part, by the development of attentional abilities. When presented with an array that contains more information than can be maintained in VSTM, one strategy is to selectively attend to some information and maintain attention to just those objects and their locations. Indeed, individual differences in adults’ VSTM capacity appear to be related to such processes – adults with low VSTM capacity seem to be unable to ignore distracting or irrelevant sources of competition (Fukuda & Vogel, 2009). As infants develop the ability to selectively orient attention in the context of multiple competing inputs, they may be able to select individual items from multiple-item arrays and store them in VSTM. At set size three, 6-month-old infants are presumably faced with arrays that exceed their VSTM capacity. They apparently do not select individual items, or attend to the binding between particular items in particular locations, in such arrays because they do not detect changes at set size three even when every item changes (Oakes et al., 2006; Oakes et al., 2009). Selectively attending to and encoding a single item – or remembering the identity of an item at one location – would have resulted in infants’ detection of the changes in those streams. This same strategy will not be sufficient to detect a change when one randomly chosen item changes from cycle to cycle, because selectively attending to only one item in such streams will result in seeing the changes on only every third cycle, on average.

Such selective attention is not without cost, however. Selective encoding and/or maintenance of only a few items will preclude the encoding of other potentially relevant targets. Indeed, this is the basic premise behind the phenomenon of inattentive amnesia, or adults’ apparent insensitivity to changes that occur in unattended locations in space (Wolfe, 1999). Presumably, although potential targets for attention are identified via preattentive processes (sufficient for guiding fixations), these preattentive percepts are not bound into objects until attention is deployed. Thus, a failure to successfully deploy attention to a particular item would result in the lack of a stable object representation, making the process of change detection impossible.

Despite the costs of competition and selectivity, they are critical for dealing with a capacity limited resource such as VSTM (Luck & Vogel, 1997). For example, if during a VSTM task attention is drawn to a particular location in space, it becomes easier to detect changes at that location and more difficult to detect changes at other locations (Schmidt, Vogel & Luck, 2002). Thus, drawing attention to the most relevant location in space can ensure that changes are detected that would otherwise likely be missed. Clearly, therefore, attention and VSTM together create a delicate balance between stability and flexibility: VSTM provides the stability necessary to support cognitive processes such as comparison despite disjointed or degraded inputs, and attention provides flexibility necessary to update VSTM to include only the most relevant information. The question is how does the development of selective attention contribute to infants’ developing ability to represent the items from multiple-item arrays into VSTM?

The present investigation sought to determine whether infants’ VSTM is influenced by attention, and how developmental differences observed for infants’ VSTM might be a function of developing selective attention abilities. Thus, this study goes beyond demonstrating developmental changes in VSTM capacity by exploring...
how VSTM and attention interact and how this interaction contributes to changes in VSTM capacity over the first year. This is a key issue in the adult literature on VSTM (e.g. Schmidt et al., 2002), particularly for understanding individual differences in adults’ VSTM (Cowan, 2001; Fukuda & Vogel, 2009).

A main goal of this work was to establish whether, as has been found for adults, cuing infants’ attention increases the selectivity of what information is represented in VSTM. Although there have been demonstrations that attentional cues can influence infant eye movements (e.g. Hood & Atkinson, 1993; M.H. Johnson, 1995b; M.H. Johnson & Tucker, 1996; Posner, Rothbart & Thomas-Thrapp, 1997), no studies have yet shown that such cues enhance the binding of preattentive percepts into more stable object representations. We examined the effect of an attentional cue on infants’ preferences for changing stimulus streams. We reasoned that if such cuing operates in infancy as it does in adults, infants should prefer changing streams in which the cue directs attention to the item that is changing compared to changing streams in which the cue directs attention to a non-changing item, even if the infants might detect the change in other contexts. For example, 10-month-old infants presented with a changing array of three items may fail to detect the change if a cue directs attention to a non-changing item, even though they could detect the changes in the absence of a cue. Thus, the question here is not whether infants prefer changing over non-changing streams, but whether infants prefer streams in which a cue directs attention to a changing item over streams in which a cue directs attention to a non-changing item. It may seem obvious that such an effect must be observed; however, this result would show for the first time that cues not only produce overt orienting of gaze, but that they can actually influence the encoding of objects in VSTM.

A second goal was to determine whether developmental changes in infants’ visual attention might have contributed to previously observed developmental changes in VSTM, as assessed in a change detection task. Infants can only detect changes, and prefer the changing stimuli, if they have attended to and encoded the items in an array that change. When the number of items is within VSTM capacity, such selection is trivial – all items are encoded. However, when the number of items in an array exceeds VSTM capacity, encoding a subset of items requires that some items are attended to and selected. Young infants may be unable to engage in this selection. From birth to 6 months of life, infants undergo multiple dramatic changes in the development of attentional control (e.g. Atkinson, 1984; M.H. Johnson, 1995a; Posner & Petersen, 1990). Specifically, young infants have difficulty inhibiting express saccades, are relatively slow to orient to targets, and have more difficulty disengaging from a currently fixated target (Hood & Atkinson, 1993). Moreover, sensitivity to the particular features of some targets (i.e. motion or color) also changes over the first 6 months (Frick, Colombo & Saxon, 1999; Ross & Dannemiller, 1999). In this case, an exogenous attention cue may facilitate the selective encoding of a single item, scaffolding infants’ change detection. Thus, a related goal was to determine whether infants under 6 months can detect changes in briefly presented arrays containing multiple different colored items under any conditions.

However, the same immaturities in selective attention that make it difficult for infants to focus on just a subset of items in multiple-item arrays may also influence the effectiveness of particular cues at facilitating young infants’ attention to a single item in a multiple-item array. That is, the effectiveness of particular cues may depend on the broader context and how demanding the stimuli and task are on infants’ immature attentional processing. If the cues or other stimuli are themselves demanding of attention resources, infants may be unable to use them to select only one item from the array. We predict, therefore, that the effect of cuing on younger infants’ encoding of information on VSTM will depend on the characteristics of the cue. That is, younger infants will be able to take advantage of a cue only if the processing of the cue is not so demanding that it actually diverts resources away from the information being cued.

To explore these issues, we tested 5- and 10-month-old infants in a variation of the change preference procedure used by Ross-Sheehy et al., (2003). On each trial, we presented a single stimulus stream in which arrays of three objects briefly appeared (i.e. for 500 ms), briefly disappeared (i.e. for 300 to 600 ms), reappeared, disappeared, etc. over a 20-s period. At each reappearance, the color of one object had changed relative to the previous appearance of the array; therefore all the streams were changing streams. Recognizing a change in this context requires that infants rapidly (within 500 ms) build some type of memory representation of the objects in the array, retain that information during the 300-ms retention period, and compare that memory to the items in the now-visible display. Moreover, this task isolates short-term memory because the brief retention interval is sufficiently long to minimize contributions from iconic memory (Becker, Pashler & Anstis, 2000), the brief presentation of the arrays of colored squares requires rapid encoding (i.e. within 500 ms), and because the colors of the items are drawn from a small set, such that high within- and between-trial similarity would likely lead to substantial long-term memory interference (Reznick, Morrow, Goldman & Snyder, 2004).

We made two changes to the procedure. First, we incorporated an attention cue (a box surrounding the location of one square in Experiment 1, and the continual rotation of one square in Experiment 2) that either directed infants’ attention to a square that changed color (valid trials) or to a square that did not change color (invalid trials). Infants should look longer at a valid than an invalid trial only if such cues can facilitate their selective attention to it, helping them to encode and maintain it in VSTM. Comparing older and younger
infants will uncover developmental changes in this process.

In addition, we used a single-stream design rather than the dual-stream design used in previous studies (Oakes et al., 2009; Oakes et al., 2006; Ross-Sheehy et al., 2003). Given our hypothesis that young infants fail to detect change in set sizes larger than one (e.g. Ross-Sheehy et al., 2003) due in part to poor attentional control, we opted for the less demanding single-screen task design that reduces both the total eccentricity of the array and the need to make large across-array saccades.

**Experiment 1**

In Experiment 1, each array of colored squares was preceded by a cue, which was a black outlined box that appeared at a location that would subsequently contain one of the colored squares (see Figure 1). Such cues are effective at inducing overt orienting of gaze (M.H. Johnson, Posner & Rothbart, 1994; M.H. Johnson & Tucker, 1996; Richards, 2001).

**Method**

**Participants**

Twenty-six 5-month-old ($M = 25.68$ weeks, $SD = .95$, 13 boys and 13 girls) and 26 10-month-old infants ($M = 43.81$ weeks, $SD = 1.03$, 13 boys and 13 girls) were included in the final sample. All infants included in this investigation were healthy and full-term, with no history of birth complications or vision problems, and were not at significant risk for color blindness (e.g. we excluded male infants with maternal uncles or grandfathers who were colorblind). All of the infants’ mothers had graduated from high school, and 40 had completed at least a bachelor’s degree. Forty-four of the infants were White, and the other infants were mixed race. One infant was Hispanic.

An additional 12 infants (eight 5-month-old infants, four 19-month-old infants) were tested but excluded from the analysis due to fussiness or lack of interest. Infant names in this study and in all subsequently reported studies were obtained from county birth records, and all parents were contacted by letter and received a follow-up phone call to schedule their appointment. Families were not paid for their participation, but infants received a small toy and parents’ parking expenses were reimbursed.

**Stimuli and apparatus**

A Macintosh G4 computer was used to present the stimuli on a single, 17-in ViewSonic monitor with a viewable surface of $18.26\,"(w)\times 13.5\,"(h)$ at a distance of 100 cm. As illustrated in Figure 1, the stimulus streams consisted of sequences of arrays that blinked on and off. Each array contained three colored squares measuring $3.6\,cm\times 3.6\,cm$, subtending $2.06\,\times \,2.06\,\beta$ per square at a viewing distance of 100 cm. In addition, an attention cue (an unfilled, black rectangle measuring $4.4\,cm\times 4.4\,cm$ or $2.52\,\beta$ square with a line thickness of .2 cm or .11") appeared 300 ms before the presentation of the array of squares in a single, randomly chosen location that would be occupied by a square.

Prior to each trial, an attention-getting stimulus was used to orient infant gaze to the monitor. Within a given trial, the stimulus streams contained the following sequence of events: (1) the cue appeared on the blank screen and remained visible for 300 ms; (2) the array of three colored squares appeared and remained visible for 500 ms, with one of the squares located within the cue and the other two squares at random locations; and (3) the cue and array disappeared simultaneously and the screen remained blank for a 300-ms delay period. This cycle repeated continuously for 20 s, and we measured look duration while the infant scanned each array.

On each reappearance of the array, the color of one of the squares changed, whereas the color of the other two

![Figure 1](image-url)  
*Schematic depiction of stimuli used in Experiment 1. The cue-array-delay sequence repeated for a total of 20 seconds. Infants saw one stream at a time, valid or invalid, in random order for a total of six trials.*

© 2010 Blackwell Publishing Ltd.
squares remained the same (initial colors were drawn randomly without replacement from a pool consisting of red, green, blue, yellow, orange, brown, purple or pink). For valid streams, the square inside the cue changed color. For the invalid streams, one of the uncued squares changed color. Thus, across any 20-s trial, there were 18 color changes in both valid and invalid streams. The critical difference between the two types of streams was whether the square that changed color was cued (valid trials) or not (invalid trials). The location of the cue and the object that changed color were constant across cycles within a given trial, but varied randomly from trial to trial. Note that the valid and invalid streams were identical in many ways: they both contained a cue; one of the objects appeared within the cue; one of the three objects changed color on each reappearance; and the location of the changing object was constant across the 20-s stream. The only difference was whether the changing object or one of the unchanging objects was cued.

Design and procedure

This experiment incorporated a $2 \times 2 \times 2$ design with factors of Trial type (valid or invalid), Age (5 or 10 months) and Block (one or two). Each infant saw two blocks of six trials, for a total of 12 trials. Each block contained three valid trials and three invalid trials, pseudo-randomized with the criterion that no three consecutive trials were the same type (valid or invalid). If the cue helps infants to selectively encode the square at the attended location, then infants should prefer the valid trials to the invalid ones. This pattern would indicate that: (1) the cue directed the infants’ attention to a specific location; (2) infants encoded the item presented at that location into VSTM; and (3) infants detected when that item had changed.

We presented infants with two blocks of trials because pilot data using this single-stream procedure revealed that many infants responded to the novelty of the dynamic displays with indiscriminate looking during the first several trials, potentially obscuring any differences between valid and invalid trials. For example, during the first trials infants may look equivalently to blinking, multiple-object arrays despite the somewhat more subtle differences in the validity of the cue. Indeed, some studies have shown that preferences between two stimuli are not apparent in early trials but emerge over trials (Bahrick, Moss & Fadil, 1996; Moore, Benenson, Reznick, Peterson & Kagan, 1987). We therefore designed the experiment to contain two blocks of six trials, each with three valid trials and three invalid trials, and analyzed the two blocks separately.

Note that 10-month-old infants do detect changes in multiple-object arrays even when the change is not cued (Oakes, Ross-Sheehy & Luck, 2007; Ross-Sheehy et al., 2003), suggesting that they can spontaneously selectively attend to some items in the array, or store all the items in arrays of three items. Thus, these infants should not need the cue to selectively encode one item. However, the attention cue may have an obligatory influence on what subset of information infants represent. Even though they could, in principle, represent more information in the display, 10-month-old infants may nevertheless selectively encode only the item in the cued location, ignoring the items at the other locations, and prefer the valid to the invalid stream.

Infants were seated on a parent’s lap approximately 100 cm in front of a large black curtain that hung ceiling to floor obscuring their view of the experimental apparatus. Openings in the curtain revealed the computer monitor, a small grey speaker, and a video camera focused on the infant’s face. In another room, a trained observer viewed the infant via live video and recorded the duration of looking by pressing and holding computer keys. In addition, 25% of all of the data reported here were recoded offline by a second trained observer. Average between-observer reliabilities were very good: mean inter-observer correlation for the duration of looking on each trial was high, $r = .98$, and the mean absolute difference between observers for the duration of looking was low, $M = .55$ s.

Before each trial, the screen flashed from grey to white at approx 2 Hz paired with a police whistle. This flashing oriented the infant’s gaze toward the monitor and ensured trials only started when the infant was actually looking at the monitor (note: we used a shapeless attention-getting stimulus to reduce the possibility that encoding of the objects in the attention-getter would interfere with infants’ encoding of the objects in the relevant stimulus stream). When the infant looked at the flashing monitor, the experimenter pressed a computer key that simultaneously ended the attention-getting stimulus and began the trial. The trial durations were infant-controlled; the stimulus stream was presented for 20 s, or until the infant had looked away for 1 consecutive second following a minimum look of at least 1.5 s. If the infant did not look within the first 10 seconds of the trial, the trial was ended, the attention-getting stimulus was initiated, and the trial was repeated. Infant looks were measured and recorded using software developed for this purpose (Cohen, Atkinson & Chaput, 2004).

Results

Mean infant looking times were calculated for each trial type during each block of six trials (i.e. trials 1 through 6 were block 1, and trials 7 through 12 were block 2; there were three valid and three invalid trials in each block). An inspection of the means for the valid and invalid trials suggested that, as expected, looking time decreased from block 1 ($M = 6.95$ s, $SD = 3.18$) to block 2 ($M = 4.85$ s, $SD = 2.61$). Indeed, a mixed-model Analysis of Variance (ANOVA) with Trial type (valid, invalid) and Block (one, two) as within-subjects variables, and Age (5 or 10 months) as the between-subjects variable conducted on the mean looking times revealed a significant
main effect of Block, $F(1, 50) = 28.20$, $p < .001$, $\eta^2_p = .36$, confirming that, overall, infants looked significantly longer during the first block than during the second block of trials. None of the other effects or interactions was significant.

However, Figure 2 shows that 5-month-old and 10-month-old infants responded differently. Although infants of both ages decreased their looking overall in the second block, the younger infants looked for similar amounts of time during both block 1 and block 2 to the valid and invalid stimulus streams. Ten-month-old infants, in contrast, looked similarly to the two types of streams in block 1, but looked longer at the valid streams than to the invalid streams in block 2. Given that other studies have revealed effects in later but not in earlier trials (e.g. Bahrick et al., 1996; Moore et al., 1987), we conducted separate ANOVAs on each block. Not surprisingly, the ANOVA conducted on the looking times in Block 1 revealed no significant effects or interactions, indicating that during these first trials infants at both ages looked about the same amount of time, and they looked equivalently at valid and invalid trials.

The ANOVA conducted on the looking times in Block 2, in contrast, revealed a significant interaction between age and trial type, $F(1, 50) = 4.80$, $p = .03$, $\eta^2_p = .09$. As suggested by the figure, in this block 10-month-old infants, but not 5-month-old infants, looked longer to the valid than to the invalid trials.

Next, we conducted a series of $t$-tests comparing looking times for the valid trials to looking times for the invalid trials for each block at each age. These comparisons revealed that although 10-month-old infants looked approximately equally to the valid and invalid trials in the first block, $t(25) = -.48$, $p = .64$, they looked significantly longer to the valid trials than to the invalid trials during the second block, $t(25) = 2.45$, $p = .02$, $d = .48$ (all reported comparisons are two-tailed). In other words, 10-month-old infants differentiated the valid and invalid trials, but only in the second block.

In contrast to the 10-month-old infants, the 5-month-old infants failed to differentiate the valid and invalid trials in either block. They looked approximately equivalently during the valid and invalid trials in both block 1, $t(25) = .58$, $p = .57$, and block 2, $t(25) = -.45$, $p = .15$. These comparisons confirm the impression from the ANOVA on block 2; 10-month-old infants, but not 5-month-old infants, differentiated the valid from the invalid trials in the second block.

**Discussion**

Clearly, these results show that infants’ VSTM is influenced by attention, at least by 10 months of age. In the second block of trials – presumably after infants have become accustomed to the blinking stimuli and the testing situation in general – 10-month-old infants looked longer if the color-changing square was cued (valid trials) than if one of the non-color-changing squares was cued (invalid trials).

Therefore, not only can exogenous cues lead to an orienting of gaze (e.g. M.H. Johnson, 1995b), but they also increase selectivity in 10-month-old infants, causing them to prefer cued changes to uncued changes. This effect is particularly remarkable because 10-month-old infants can detect changes in stimulus streams involving arrays of three objects in the absence of any cuing (Ross-Sheehy et al., 2003). It is possible that these infants detected the changes in both the valid and invalid streams, but the valid cue enhanced their recognition, perception, or interest in that change. Alternatively, the cue may have caused infants to focus on just the cued item to the exclusion of the uncued items; in this case, infants may not have detected the change in the invalid trials. Regardless of which of these two possibilities is correct, the fact that these infants preferred the valid stream to the invalid stream suggests that the exogenous cue was extremely compelling, automatically capturing infants’ attention, and constraining the input to their capacity limited VSTM resources.

These results also revealed developmental change in the interaction between VSTM and attention. Although 10-month-old infants preferred valid trials to invalid trials, 5-month-old infants looked equal durations to these two types of events, suggesting that the attention cue used here did not enhance younger infants’ selection of a single item, or facilitate their preference for changes in these multiple-object arrays. This raises the possibility that although some types of exogenous spatial cues can effectively capture infants’ attention, resulting in a gaze shift (M.H. Johnson & Tucker, 1996), this attentional capture does not translate into young infants’ selective encoding of the cued item into VSTM. As was observed previously, the present results suggest that infants younger than 6 months do not encode the color of any items in multiple-item arrays, even when their attention is pointed to one of the items by a compelling exogenous cue.
However, there are several alternative possibilities for the ineffectiveness of the cue in Experiment 1, and it is possible that cues can be effective at facilitating younger infants’ selective attention under other circumstances. For example, the type of cue we used in Experiment 1 may have been ineffective for younger infants because they may have had difficulty disengaging from the cue stimulus in order to encode the attended item into VSTM (Hood & Atkinson, 1993). That is, attention may have been ‘stuck’ on the cue, preventing them from encoding the colored square presented at this location.

It is also possible that younger infants directed their attention and then actually encoded the attention cue itself in VSTM, thus overloading their limited capacity. As a result, they may not have been able to encode the color of the object contained in the cue. That is, the cue may have been effective at capturing infants’ attention, and infants may have encoded precisely one object from the multiple-object array: the cue. The spatial cue used in this experiment might constitute an object, taxing the extremely limited capacity of 5-month-old infants’ VSTM. Furthermore, even if infants were able to selectively encode the target into VSTM, it is possible that they lost the fragile representation when the cue subsequently reappeared after the delay interval. This limited capacity account could explain why 10-month-old infants succeeded in this task whereas the 5-month-old infants failed, as the older infants have sufficient capacity to encode the cue and the cued square (Oakes et al., 2009; Oakes et al., 2006; Ross-Sheehy et al., 2003).

Finally, the small differences in stimulus timing parameters in this procedure as compared to that used in previous studies (Oakes et al., 2009; Oakes et al., 2006; Ross-Sheehy et al., 2003) may have increased forgetting, as the delay period for each cycle was 300 ms longer than in previous VSTM tasks due to the insertion of the attentional pre-cue. Thus, young infants may have failed to prefer the valid streams because they were unable to maintain that memory during the 300-ms retention period and the 300-ms cue period.

In Experiment 2 we tested these possibilities by observing 5-month-old infants in this task in which one square was cued by making it rotate during the period in which it was visible. This manipulation addressed the competition and timing concerns raised in Experiment 1; because the cue was a property of one of the objects, this version eliminated the possibility that the cue itself would either inhibit attention to the colored square at the cued location or compete with this item for representation in VSTM. In addition, it enabled us to eliminate the pre-cue period, so that the stimulus streams in Experiment 2 had the same timing parameters as in previous studies of VSTM in this age range (Oakes et al., 2009; Oakes et al., 2006; Ross-Sheehy et al., 2003). If 5-month-old infants in Experiment 1 failed to discriminate between the valid and invalid streams due to competition, speed of processing demands, VSTM load, or timing differences, then cuing the square using rotation of the square itself should be sufficient to allow infants to encode the cued item into VSTM. This would demonstrate for the first time that young infants can use VSTM to encode the properties of one item when faced with multi-item arrays, and that exogenous attention can influence both perception and VSTM encoding.

**Experiment 2**

**Method**

**Participants**

Twenty-four 5-month-old infants (M = 25.55 weeks, SD = .92, 12 boys and 11 girls) served as participants. All of the infants’ mothers had graduated from high school, and 16 had completed at least a bachelor’s degree. Twenty-two infants were White, and two infants were of mixed race. Six additional infants were tested but excluded from the analysis due to fussiness (n = 5) or parental interference (n = 1).

**Stimuli**

The stimuli were identical to those used in Experiment 1 except that instead of a pre-cue, one of the three squares rotated slowly (approximately .06 rotations per second) when the array was visible, and the 300-ms pre-cue period was removed from the sequence of events (see Figure 3). For half of the streams the targets rotated clockwise, and for half the targets rotated counterclockwise. Like a spatial cue, motion in an otherwise static display is a strong elicitor of an orienting response (Dannemiller & Nagata, 1995; Eizenman & Bertenthal, 1998; S.P. Johnson & Mason, 2002).

**Design and procedure**

The apparatus, procedure, and design were identical to Experiment 1.

**Results**

Mean looking times were calculated for each trial type and block (see right panel of Figure 2). Unlike Experiment 1, there was little overall decrease in looking from Block 1 (M = 5.64 s, SD = 2.53) to Block 2 (M = 5.13 s, SD = 2.53). Indeed, the ANOVA conducted on infants’ looking time did not reveal a significant main effect of block, F(1, 23) = .55, p = .47.

The means in Figure 2 do suggest, however, that, like the 10-month-old infants in Experiment 1, 5-month-old infants responded differently to the valid and invalid trials in the first block and second block. The ANOVA on the looking times confirmed this impression, and revealed a Block by Trial type interaction F(1, 23) = 4.65, p = .04, ηp² = .17.
As in Experiment 1, we also conducted separate analyses on infants’ looking at valid and invalid trials for each block. The ANOVA on the looking times during block 1 revealed no significant main effects or interactions. The ANOVA on the looking times for Block 2 revealed a significant main effect of trial type, \( F(1, 23) = 5.10, p = .03, \eta^2_p = .18 \), indicating that 5-month-old infants in Experiment 2 did look longer to the valid than to the invalid trials during the second block.

We also conducted a series of paired \( t \)-tests comparing the valid trials and invalid trials. Just as was observed for the 10-month-old infants in Experiment 1, 5-month-old infants looked significantly longer to the valid trials than to the invalid trials during the second block, \( t(23) = 2.26, p = .04, d = .46 \), but not during the first block, \( t(23) = .67, p = .50 \).

Supplemental analyses

The separate analyses for Experiments 1 and 2 suggest that the different types of attention cues used in those two experiments had a substantial effect on 5-month-old infants’ performance. We found no evidence of infants’ differentiation of valid and invalid trials in Experiment 1, in which the attention cue could be encoded as a separate object, whereas infants preferred the valid trials in Experiment 2, in which the cue was an integral part of the object itself. To examine these apparent differences, we conducted a supplemental ANOVA on looking times for the 5-month-old infants from the two experiments with Trial type (valid or invalid) and Block (1 or 2) as within-subject variables, and Experiment (1 or 2) as the between-subjects variable. Importantly, there was no main effect of Experiment, \( F(1, 48) = .04, p = .84 \), indicating that overall the two groups of infants watched the events for similar amounts of time. This analysis did, however, reveal a significant main effect of Block, \( F(1, 48) = 8.70, p = .005, \eta^2_p = .15 \), and a Block by Experiment by Trial type interaction, \( F(1, 48) = 5.36, p = .03, \eta^2_p = .10 \), confirming the qualitative conclusion that 5-month-old infants in Experiment 2 were increasingly selective during the experiment, whereas 5-month-old infants in Experiment 1 were not.

Discussion

This experiment showed for the first time that infants under 7 months of age can detect changes within multiple-object arrays in a VSTM task, at least under certain conditions. The success of the 5-month-old infants in this experiment suggests that previous failures were due, at least in part, to the attentional demands imposed by multiple-item arrays. In this experimental context, infants selectively attended to and encoded the rotating object, looking longer when the rotating object changed color than when the rotating object did not change color. Thus, the present results indicate that infants younger than 6 months can use VSTM to detect a change in multiple-object arrays, but that they require attentional scaffolding to do so.

This experiment also revealed that an exogenous cue can influence the encoding of information in VSTM in infants as young as 5 months. That is, just as the results of Experiment 1 showed for 10-month-old infants, the results of Experiment 2 showed that attention cues not only direct attention to objects, but can determine what information is encoded in infants as young as 5 months of age.

Moreover, the comparison of Experiments 1 and 2 demonstrates that the characteristics of the exogenous cue contribute to whether it will influence infants’ VSTM performance, at least at 5 months of age. We cannot determine from these experiments whether the key factor was the direct cuing of the to-be-remembered item or the duration of the retention interval. However, VSTM representations in adults do not decline until at least 4–5 seconds have elapsed (Vogel, Woodman & Luck, 2001; Zhang & Luck, 2009), and it therefore seems likely that the key factor in the present study was the direct cuing of the object. Thus, these results suggest that younger
infants may require cues that are not spatially distinct from the target in order to rapidly encode the target into VSTM, and that the previous findings of 4- and 6-month-old infants’ failures in VSTM tasks may have been due in part to the attentional demands of the task.

Finally, the comparison of Experiments 1 and 2 shows that the failure of 5-month-old infants to prefer the valid stream in Experiment 1 did not result from a general inability to detect changes in any multiple-item array. Rather, when the conditions provide the right type of support, even young infants can selectively attend to one item in multiple-item arrays. This pattern is consistent with a general model of development in which the emergence of a new skill is jointly determined by the development of multiple independent systems (e.g. Adolph, Vereijken & Shrout, 2003; Spencer, Vereijken, Frederick & Thelen, 2000; Thelen & Ulrich, 1991).

**General discussion**

These data make several contributions to our understanding both of VSTM in infancy and of the interactions between attention and other cognitive processes. First, we showed that infants’ VSTM is influenced by selective attention – in both experiments we influenced infants’ encoding of items in VSTM by providing a compelling attention cue. Ten-month-old infants in Experiment 1 preferred streams in which the changing square was cued to streams in which a non-changing square was cued, despite the fact that both streams contained a changing object. Five-month-old infants showed the same effect, preferring streams in which a changing square rotated to streams in which a non-changing square rotated. These results go beyond demonstrating cue-driven orienting responses by showing that attention actually changes the way information is processed in VSTM at the cued location. Such attentional influences have been demonstrated in adults for perceptual encoding (Luck, Hillyard, Mouloua, Woldorff, Clark & Hawkins, 1994; Yeshurun & Carrasco, 1998) and VSTM encoding (Rensink, O’Regan & Clark, 1997; Schmidt et al., 2002). The results reported here show that attention can play a similar role in infants’ visual cognitive processes.

Experiment 2 also showed for the first time that infants younger that 6 months can, under certain conditions, detect changes in rapidly presented multi-item arrays. These results contrast sharply with previous findings using very similar stimulus streams showing that infants of 6 months failed to encode even a single item’s worth of information when the displays contained multiple distinct items (Oakes et al., 2009; Oakes et al., 2006; Ross-Sheehy et al., 2003). Thus, infants younger than 6 months can encode into VSTM information about one item, even when that item is presented in an array of multiple different items, as long as their attention is scaffolded with exogenous cues. These findings therefore solve a puzzle about VSTM in young infants. Specifically, because almost all natural scenes are multiple-element arrays, prior findings seem to indicate that young infants would be abysmal at storing information in VSTM in daily life. The present finding indicates that things are not quite so dire, and that these infants will at least be able to store salient information in VSTM.

Second, the present experiments provide insight into how previous observations of changes in VSTM capacity over the first year of life might be related to developmental changes in selective attention. We found here that although the use of an attention cue could enhance VSTM for multiple-object arrays in 5-month-old infants, these younger infants appear to have a more limited attentional repertoire, using only non-spatially distinct cues to encode items into VSTM. Thus, although attentional cuing can be effective even at 5 months in this context, this effect is fragile and subject to the specific characteristics of the cue and/or task. Attention cuing therefore is not an all-or-none process, with all attention cues being equally effective. Rather, the particular cue that is effective at 5 months of age depends on factors such as timing and competition between elements of the array.

What does this mean for our understanding of VSTM capacity in infancy? Clearly, previous reported failures do not reflect a complete inability of infants 6 months and younger to encode information in VSTM from multiple-item arrays. Infants younger than 6 months can be induced to encode items from and detect changes in multiple-item arrays if they are exogenously directed to encode one item in the array, although this selection process is limited. Other studies also have reported that young infants’ attentional processes vary depending on the features of the cue and stimuli (Dannemiller, 2000; Ross & Dannemiller, 1999). For example, Kaufman, Gilmore and Johnson (2006) observed that when planning a sequence of eye-movements, 4-month-old infants used a retinocentric frame of reference when the processing of the stimuli was more attention-demanding, and a more sophisticated body-centered frame of reference when the stimuli were less attention-demanding. We propose that the present differences in the effectiveness of the two cue types reflect the development of the selective attention system.

In the present case, there are at least two stimulus-related factors that may have interfered with the effectiveness of the attention cue in younger infants. First, after the initial orienting response to the spatial cue, the abrupt onset of the target array may have catastrophically interfered with attentional selection for the younger but not the older infants. Indeed, younger infants are disproportionately driven by abrupt exogenous events (Posner et al., 1997). There is both functional and anatomical support for this hypothesis (for a review see Atkinson, 1995). Similarly, the onset of a spatial pre-cue following the presentation of the target array and delay interval may have been sufficient to disrupt fragile
VSTM representations in the 5-month-old infants, interfering with their ability to compare the remembered item to the subsequently reappearing target.

Another possibility is that the spatially distinct cue used in Experiment 1 actually interfered with processing by competing with the cued square for access to VSTM. Infants’ automatic fixations can be suppressed when the appearance of the target item overlaps in time with the currently fixated item (Atkinson et al., 1977; Banks & Salapatek, 1978; Hood & Atkinson, 1993). Thus, 5-month-old infants in Experiment 1 may have failed to disengage attention from the cue, orient to the target square, and encode it within the 500-ms time window. In Experiment 2, the relative motion of the cue might actually help the process of encoding because it would induce persistent attention to the colored square, increasing the likelihood that it was encoded. Identifying which of such mechanisms are responsible for the effects observed here will also provide additional understanding of developmental changes in infants’ VSTM for multiple-item arrays.

There are at least two hypotheses about why the attention cue helps infants detect changes in this VSTM task. First, the presence of the attention cue helps by essentially reducing the visual memory load to only a single item. Young infants’ immature selective attention processes may make it difficult for them to reduce multiple-item arrays to a manageable number of items. Indeed, one apparent source of capacity differences among adults is the ability to focus on a subset of items; low-capacity adults seem to have more difficulty ignoring irrelevant items (Fukuda & Vogel, 2009). In the present context, the rotating object in Experiment 2 may have created a visual pop-out effect, helping young infants to attend and maintain attention to that item, thus reducing the set size to one item.

Second, the attention cue may help infants to bind the cued object’s color to its location. Previous work has demonstrated a striking deficit in the ability of infants younger than 8 months to bind color and location in VSTM (Oakes et al., 2006). Binding may reduce VSTM load by combining multiple feature dimensions into a single object representation (e.g. red square on the right). However, for infants with limited binding facility, the number of to-be-remembered features grows factorially as a function of the number items in the display (i.e. red thing, square thing, and something on the right). Thus, attention cues may scaffold binding in infants whose VSTM capacity would otherwise be overwhelmed. It is important to note, however, that binding may not be necessary in this task. That is, because the attention cue provided unambiguous spatial information at all times, infants did not need to bind color information to a specific location in space. Further studies are necessary to more clearly determine the effect of attention on binding in VSTM per se. Similarly, the perceptual distinctness of the cue may have helped infants by serving as a spatial anchor, allowing infants to align their memory for a specific square with the appropriate square on the subsequent array.

In summary, the present results provide evidence about both the development of VSTM in infancy and development of interactions between attention and VSTM. Results like those reported here confirm that when faced with multiple-object arrays, only the most salient, compelling items likely are encoded, and infants’ failure to detect changes in multiple-object arrays likely reflects their inability to inhibit responding to multiple equally salient objects. Clearly, some type of attentional support can help young infants overcome the limitations in their VSTM abilities. Our results suggest that the dynamic interaction between attention and VSTM may offer a potent solution to capacity limits, enabling young infants to rapidly encode items into VSTM despite the presence of multiple competing items in the visual input.

Acknowledgements

A portion of these data were submitted by SRS to the University of Iowa as partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology. This research and preparation of this manuscript were made possible by NIH grants HD49840 awarded to LMO and MH076226 awarded to SJJ; SRS was funded by a predoctoral NRSA (MH068934) and a postdoctoral NRSA (HD055040). We thank committee members Andrew Hollingworth, Jodie M. Plumert, and Steven Anderson. We would also like to thank Shaena McGivern, Sammy Perone, Kristine Kovack-Lesh and the undergraduate students in the Infant Cognition Laboratory at the University of Iowa, for their help with this project.

References


© 2010 Blackwell Publishing Ltd.


© 2010 Blackwell Publishing Ltd.


Received: 18 February 2010
Accepted: 31 May 2010