Title

A new perspective on the role of physical salience in visual search: Graded effect of salience on infants attention.

Journal

Developmental Psychology, 59(2)

Authors

DeBolt, Michaela
Mitsven, Samantha
Pomaranski, Katherine
et al.

Publication Date

2023-02-01

DOI

10.1037/dev0001460

Published in final edited form as:
Dev Psychol. 2023 February ; 59(2): 326–343. doi:10.1037/dev0001460.

A New Perspective on the Role of Physical Salience in Visual Search: Graded Effect of Salience on Infants' Attention

Michaela C. DeBolt[1,2], **Samantha G. Mitsven**[2], **Katherine I. Pomaranski**[1,2], **Lisa M. Cantrell**[3], **Steven J. Luck**[1,2], **Lisa M. Oakes**[1,2]

[1]Department of Psychology, University of California, Davis

[2]Center for Mind and Brain, University of California, Davis

[3]Department of Child Development, California State University, Sacramento

Abstract

We tested 6- and 8-month-old White and Non-White infants (N = 53 total, 28 girls) from Northern California, USA, in a visual search task to determine whether a unique item in an otherwise homogeneous display (a *singleton*) attracts attention because it is a unique singleton and "pops out" in a categorical manner, or whether attention instead varies in a graded manner on the basis of quantitative differences in physical salience. Infants viewed arrays of 4 or 6 items; one item was a singleton and the other items were identical distractors (e.g., a single cookie and 3 identical toy cars). At both ages, infants looked to the singletons first more often, were faster to look at singletons, and looked longer at singletons. However, when a computational model was used to quantify the relative salience of the singleton in each display—which varied widely among the different singleton-distractor combinations—we found a strong, graded effect of physical salience on attention and no evidence that singleton status per se influenced attention. In addition, consistent with other research on attention in infancy, the effect of salience was stronger for 6-month-old infants than for 8-month-old infants. Taken together, these results show that attention-getting and attention-holding in infancy vary continuously with quantitative variations in physical salience rather than depending in a categorical manner on whether an item is unique.

Keywords

Infancy; Visual attention; Visual search; Eye tracking; Physical salience

Visual attention plays a key role in infants' ability to learn about the world by allowing them to focus on certain stimuli for further processing (Oakes & Amso, 2018). Research has examined which types of stimuli are effective in attracting attention in both infants and adults (Richards, 2010; Theeuwes, 1992), as well as the factors that determine how long

---

Correspondence concerning this article should be addressed to Michaela C. DeBolt or Lisa M. Oakes, Center for Mind and Brain, 267 Cousteau Pl., University of California, Davis, CA 95618, mdebolt@ucdavis.edu; lmoakes@ucdavis.edu.

CRediT author statement. **MCD**: Conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing—original draft. **SGM**: conceptualization, investigation, project administration, writing--review and editing. **KIP**: data curation, writing—review and editing. **LMC**: conceptualization, investigation, writing—review and editing. **SJL**: conceptualization, writing—review and editing. **LMO**: conceptualization, funding acquisition, methodology, supervision, writing—review and editing.

attention is maintained once it has been oriented (Gaspelin et al., 2016; Geng & DiQuattro, 2010). The features that *attract* and *hold* infants' attention are multiply determined and reflect a set of complex and dynamic features (Cohen, 1972; Colombo, 2001). For example, infants' looking is influenced by factors as varied as familiarity (Fantz, 1964), complexity (Brennan et al., 1966), dynamic movement (Colombo et al., 2004), and physical salience (Frank et al., 2009). Moreover, the systems involved in controlling infants' visual attention develop rapidly over the first postnatal year, shifting from what has been traditionally described as primarily obligatory, or "bottom-up", control (Stechler & Latz, 1966) to voluntary, or "top-down", attentional control (Colombo & Cheatham, 2006). Indeed, several studies have shown that the effect of physical salience on infants' visual attention decreases significantly over the first year and the effect of higher-level factors (e.g., social relevance) increases (Frank et al., 2009; Kwon et al., 2016).

However, much of the prior literature treats objects as either being salient or non salient, and this dichotomous view may be overly simplistic. Specifically, researchers often use a "winner-take-all" approach and identify the most physically salient item in an array and examine whether infants attend more to that "most salient" item than to the other items in the array (Amso et al., 2014; Gluckman & Johnson, 2013; Kwon et al., 2016). By contrast, the adult literature on attention has converged on the idea that physical salience varies continuously over a wide range and is one of several factors that contribute to attentional orienting (Fecteau & Munoz, 2006). The goal of the present work was to systematically examine infants' attention in this more nuanced framework, specifically examining the extent to which variations in physical salience are related to variations in infants' attention.

Visual attention in adults

Much research in adults examining the mechanisms that guide attention has used the *visual search* paradigm (Wolfe, 1998a, 2015). In most experiments, a specific target is defined, and observers make a speeded button press response for each stimulus array to indicate whether the target is present or absent (or to report some property of the target). Under some circumstances, the observer's covert and/or overt attention is automatically captured by specific objects, irrespective of the search goal. For example, attention can be involuntarily captured by a *singleton* item, such as a yellow item in a field of blue items. This involuntary capture of attention has been classically described as a "pop out" effect, reflecting the phenomenal experience that the singleton automatically forces itself into awareness. Classic theories proposed that feature singletons are automatically processed and attract attention regardless of the observer's goals and the number of items in the array (Bravo & Blake, 1990; Duncan & Humphreys, 1989; Treisman & Gormican, 1988). This type of search is often contrasted with slow, effortful search that is guided by "top-down" search goals (i.e., locate a green square in an array of blue squares and green rectangles).

Whether and when attention is involuntarily captured by a singleton has been the subject of intense debate (see, e.g., Luck et al. 2021). In particular, research with adults has indicated that a simple bottom-up/top-down dichotomy is problematic for characterizing attentional control (Awh et al., 2012). That is, the allocation of attention is not solely determined by the viewer's top-down goal *or* the physical characteristics of the display. Instead, research with adults has converged on the idea that the visual system computes a “priority map” representing the attentional priority of each location in the input (Awh et al., 2012). The degree of physical salience is one factor that determines the priority value (Fecteau & Munoz, 2006). Implicit learning and explicit goals can also increase or decrease the priority of a given location (Jiang et al., 2018; Theeuwes, 2019). Ultimately, a winner-take-all process selects a single location for the next shift of covert or overt attention, but the probability and speed of this shift depends in part on the degree of physical salience, which is a feature that is not all-or-none but instead varies continuously (Fecteau & Munoz, 2006). In this view, the efficiency of search — or how quickly attention is engaged— varies continuously rather than falling into dichotomous classes of fast/efficient and slow/inefficient (Wolfe, 1998b).

Moreover, there is overwhelming evidence, going back at least to Treisman and Gormican (1988), that the salience of feature singletons varies continuously. For example, response times for feature singletons vary as a function of the singleton’s physical salience (Gaspar et al., 2016; Luck et al., 2006; Töllner et al., 2011). In addition, the latency of attention-related event-related potentials are modulated by the relative physical salience of a singleton (Töllner et al., 2011). The physical salience or discrepancy of the target from other items in the array also influences adults’ search in natural scenes and other complex stimuli (Malcolm & Henderson, 2009; Wolfe et al., 2011) and in visual search arrays with complex stimuli (Alexander & Zelinsky, 2011; Malcolm & Henderson, 2009; Wolfe et al., 2011). Moreover, items that are not singletons can capture attention if they are sufficiently salient (Itti, 2005; Koehler et al., 2014). In adults, therefore, the status of an item as a singleton does not appear to be the key factor driving the capture of attention, but rather the physical salience of a given object relative to the rest of the display. That is, singletons are often highly salient, but it may be the quantitative salience of the singleton that matters rather than its status as the one unique item in an otherwise homogeneous display.

Visual attention in infancy

Researchers have also observed that feature singletons attract or capture attention in infants (Adler & Gallego, 2014; Adler & Orprecio, 2006; Colombo et al., 1995; Gerhardstein et al., 1999) regardless of the *set size* (the number of items in the array). The observation that singletons elicit rapid orienting of attention in both infants and adults raises the questions motivating the present work: Do singletons capture and hold attention in infants because they are unique or because they are physically salient? To the extent that physical salience influences attention in infants, do quantitative variations in salience lead to graded changes in attention? And how does the effect of salience vary over development?

Previous research has demonstrated that infants’ attention is influenced by physical salience in other contexts. For example, infants are more likely to look at salient locations in video stimuli (Frank et al., 2009) and complex natural scenes (Pomaranski et al., 2021). In a variation of visual search using arrays composed of heterogeneous items, Kwon et al. (2016) observed that 4-month-old infants’ attention was captured by the most physically salient item in the array, even though the array also contained a non-salient social stimulus (a human face). The 4-month-old infants in this study looked first and fastest to the physically salient item. Thus, as has been observed with adults, highly salient stimuli capture infants' attention. However, infants' attention was *held* by the social stimulus; they looked longest at the human face. This pattern confirms the classic distinction between *attention-getting* features and *attention-holding* features of stimuli (Cohen, 1972). According to Cohen, physical salience (e.g., size) determines attention-getting and complexity or information content determines attention-holding.

The features that capture and hold infants' attention change over development, however. Physical salience has a larger effect on attention capture in infants 6 months and younger than on older infants (Frank et al., 2009; Kwon et al., 2016; Pomaranski et al., 2021). With increasing age, infants become able to suppress or inhibit responding to the most physically salient item and direct their attention to more meaningful items. For example, Di Gorgio et al. (2012) found that 6-month-old infants, but not 3-month-old infants, selectively attended to faces in arrays of complex objects (e.g., shoes, clocks, cars). Although Kwon et al. (2016) observed that 4-month-old infants' attention was more likely to be captured by the most *physically salient* item in the array, 6- and 8-month-old infants were more likely to orient first to the most *socially meaningful* item in the array, as determined by top-down processes (i.e., faces). These findings are consistent with the general conclusion that infants' attention develops from being primarily driven by bottom-up (physical salience) factors to more determined by top-down factors (Colombo & Cheatham, 2006).

The question remains whether feature singleton items capture and hold infants' attention because they are singletons per se, not because of their relative physical salience. That is, previous studies have asked whether singletons attract attention without attempting to quantify the physical salience of the singleton (Adler & Gallego, 2014; Adler & Orprecio, 2006). In such studies, it is likely that the singletons captured infants' attention because they were highly salient, not because of their status as singletons per se. As we will show, an item may be a singleton without being the most physically salient item in the display. Thus, it is possible to separately quantify the effects of singleton status and salience on the ability of a stimulus to capture and hold infants' attention. This is the primary goal of the present work.

The current study

Here we conducted an experiment designed to examine infants' attention to singletons in arrays of complex objects. Although the study was not preregistered, it was originally designed to assess the factors that influence infants' attention. We used a broad set of different objects, such that the physical salience of the singleton varied widely across displays. This made it possible to separately assess the effects of physical salience and singleton status. To our knowledge, no previous work with infants, children, or adults has quantified the separate contribution of singleton status and physical saliency on attention.

We used a version of visual search that has been adapted for use with infants. Because preverbal infants cannot easily be given instructions to search for specific targets, in these tasks infants' looking behavior is evaluated as they freely view arrays realistic photographs of complex objects constructed by researchers to include one or more experimenter-defined "targets" (i.e., items that are of particular interest to the experimenters, not items that the infants have been instructed to find) (Di Giorgio, Turati, et al., 2012; Gliga et al., 2008; Kwon et al., 2016; Simpson et al., 2014). In previous studies, the target was not necessarily defined as being a singleton, but rather as being significant in some other way, often because it is a social stimulus (e.g., a face) in an array of non-social stimuli (e.g., household items).

As illustrated in Figure 1, we constructed stimulus displays that included one instance of one object (the *singleton*) and three or five instances of another object (the *distractors*). Different singleton and distractor objects were used in different displays, and the salience of the target (measured using a computational model) varied continuously depending on which object was used as the singleton and which object was used for the distractors. We derived *relative salience scores* (defined in more detail below; see Figure 2) that quantify the salience of the singleton relative to the distractors in a given array. In addition, we varied the *set size* because as the set size increases, the salience of the singleton increases (Duncan & Humphreys, 1992, 1989). Thus, our analyses examined three factors that might all contribute to attention capture: singleton status (i.e., a unique object in an array of identical distractors), relative singleton salience, and set size. Because we assume that multiple factors interact to determine both attention-getting and attention-holding, in this work we also examined how these three factors may contribute to infants' attention holding.

We sought to distinguish between three competing hypotheses. One possibility is that attentional allocation would be determined solely by singleton status, with no additional impact of physical salience. Note that, because infants cannot be given an explicit task, they might find singletons "interesting" because they are the one unique item in an otherwise-homogeneous display, even if they are not physically salient. An alternative hypothesis is that infants' attentional allocation would be determined solely by physical salience (along with set size, which indirectly impacts salience), with no effect of singleton status once salience is accounted for. A third hypothesis is that both salience and singleton status would influence the allocation of attention. In addition, because the extant literature suggests important transitions in attentional control around 6 to 8 months of age, we tested whether the effects of these factors would differ across this age range. Specifically, a number of studies suggest a transition at about 6 months in the role physical salience plays on infants' attention allocation (Di Giorgio, Turati, et al., 2012; Frank et al., 2009; Kwon et al., 2016). Thus, it is possible that physical salience will have less influence on 8-month-old infants' attention allocation compared to 6-month-old infants.

Finally, in the supplemental materials, we provide a modified replication study with a sample of 66 6-month-old infants viewing set size 4 arrays only. The purpose of this modified replication study is to examine the robustness and generalizability of the effect of physical salience on attention in a separate sample of infants. The specific details for this replication sample are available at https://osf.io/k6qev/.

Method

Participants

Infants were recruited from the Sacramento Valley (including both urban and rural communities) located in Northern California, USA. At the time these data were collected

(March 2015 - July 2015), sample sizes were determined based on rule of thumb and historical convention. The target sample size for this study was 24 infants per age group, based on the existing work using this procedure (e.g., Gliga et al., 2008; Kwon et al., 2016; Mitsven et al., 2018). Using the *pwr* package in R (Champely, 2020), we conducted an a posteriori power analysis and established that a sample size of 24 infants provides 80% power to detect a difference from chance with an effect of $d = .60$ using one-sample t-tests, and a significant difference between two groups with an effect of $d = .83$ with an independent samples t-test. Consequently, our sample size was sufficient to detect effects of the size reported by Kwon et al. (2016) with 80% power.

Names of infants for the experiments reported here were originally obtained from the State Office of Vital Records. All parents who lived within a 30-mile radius of our laboratory were sent informational mailings. Parents who wished to volunteer for studies contacted us and were included in our database of potential research participants. When infants reached the appropriate age for this study, we contacted parents about participating and set up an appointment if the parents were interested and available. All infants who participated received a small toy or T-shirt and a certificate.

The final sample included 53 infants: 26 6-month-old infants (16 girls; mean age = 184.96 days, $SD = 6.91$) and 27 8-month-old infants (12 girls; mean age = 243.15 days, $SD = 7.42$). Thirty-four infants (14 6-month-old infants and 20 8-month-old infants) were White, 3 were Asian or Asian American (2 6-month-old infants and 1 8-month-old infant), 13 were mixed race (8 6-month-old infants and 5 8-month-old infants), and the race of 3 infants was not reported (2 6-month-old infants and 1 8-month-old infant). Across these racial groups, 8 6-month-old and 4 8-month-old infants were reported to be Hispanic. All of the 53 mothers who reported their educational attainment had graduated from high school and 82% had earned at least a Bachelor's degree.

To achieve this sample, we tested an additional 14 6-month-old and 6 8-month-old infants who were excluded from the analysis due to being ineligible (i.e., family history of colorblindness, $N = 3$), yielding poor tracking or a failure to calibrate ($N = 14$), falling asleep during the session ($N = 1$), or not being interested in looking at the screen ($N = 2$).

Apparatus

We measured eye movements using an SMI-Red-M eye tracker with a sampling rate of 120 Hz. The eye tracker was attached to the bottom of a 22 in (1680 px by 1050 px) LCD monitor. The monitor was affixed to an ergo arm that allowed the monitor to be positioned to optimally locate each infant's eyes in the center of the detection radius of the eye tracking system. A Logitech Carl Zeiss Tessar 2.0/3.7 2MP Autofocus web camera mounted on the top of the stimulus display screen allowed the experimenter to monitor the infant's behavior. A Dell laptop supplied by SMI was used to monitor the infant and run the experiment. A large white cloth screen was placed behind the monitor to obstruct the infant's view of the additional equipment and the experimenter.

Stimuli

The stimuli were created from a set of 32 photographs of common items expected to be familiar to infants (e.g., toys, clothes, fruit, animals; see Figure 1). Items were edited to be equivalent in size—approximately 4.8° wide by 4.8° high at a viewing distance of 60 cm. We created 16 pairs of images by pseudo-randomly selecting items that did not overlap in category. These image pairs were arranged into visual arrays containing either four or six items on a light gray background. Note that we did not select the pairs on the basis of specific differences relative salience. Instead, we capitalized on the intrinsic differences between the objects in our stimulus set. As will be clear later, our randomly-generated stimulus pairs did in fact lead to a wide range of variation in physical salience between the singleton and the distractor items in the arrays.

All objects were centered approximately 5° from the center and equidistant from each other. One item in each pair was designated to be the *singleton*; this item appeared only once in each array and the other item in the pair was used to create 3 or 5 identical distractors in the array (see Figure 1). All stimuli used in this experiment are available at https://osf.io/k6qev/.

Note that our designation of the non-singleton items as “distractors” is by analogy to the adult attention literature in which participants are instructed to find the singleton and ignore the non-singleton items. Because we are unable to provide instructions to infants about what to search for, we ask here whether a singleton is effective at acting as a target—i.e., preferentially capturing and holding infants’ attention in these arrays.

We also used attention-getting stimuli between trials to focus infants’ attention to the center of the monitor prior to each stimulus array. At the start of each trial, infants were presented with a black fixation cross (3.34° h × 2.86° w) that blinked on and off, accompanied by a ringing sound. In addition, we periodically presented brief animated clips (e.g., short clips from Sesame Street, Teletubbies, and Blue’s Clues) to re-engage infants’ general attention to the task.

Procedure

All procedures were approved by the IRB at the University of California, Davis (protocol number: 220219; project name: Understanding cognitive development in infancy: Attention and visual short-term memory). We obtained informed consent and explained the procedures to the parents in our waiting room. Parents and infants were then escorted to a sound attenuated testing room, where infants were seated on their parent’s lap or infant high-chair, approximately 60 cm from the monitor. Parents wore felt-covered sunglasses to occlude their view of the stimuli during the experimental trials, thus minimizing parental biasing of infant behavior.

The experimenter located the infants’ eyes using SMI software iView, adjusting the position of the monitor (using the ergo arm) as necessary, a child-appropriate video was presented to help the infant maintain looking at the monitor during these adjustments. Next, the experimenter initiated the SMIs automatic calibration sequence, which involved presenting stimuli expected to be visually interesting to infants (a looming circle) at each of 5 calibration points. Immediately after calibration, a validation procedure was initiated in which an image of a yellow rubber duck accompanied by a chirping sound was presented in the four corners of the screen. Using the feedback from the validation, the experimenter repeated the calibration until the infant's gaze was within approximately 2° of the validation stimulus locations (the mean validation accuracy was 1.45 [*SD* = 1.28]). We visually inspected the data for the 6 infants with validation scores greater than 3 and determined that the high discrepancies reflected inattention during the validation procedure and that their data would contribute only a minimal amount of noise ($n = 6$).

Immediately following calibration, infants were presented with the experimental trials using Experiment Center software. Before the presentation of each experimental trial, a fixation cross was presented in the center of the screen. When the SMI system detected 200 ms of consecutive looking within a 4° by 4° AOI surrounding this fixation stimulus, the attention-getting stimulus was removed and replaced with an experimental stimulus. If infants did not look at the cross within 4 s, it was replaced with a clip from a child-appropriate movie (e.g., Blue's Clues, Teletubbies). This stimulus was removed when the system detected 200 ms of looking to the AOI as just described. This procedure ensured that infants (1) were looking at the screen at the start of each experimental trial, and (2) were not looking at the location of any of the images in the array when the trial began. When the fixation cross was removed, one of the stimulus arrays was presented for 5 s.

The stimulus arrays were presented in blocks of four trials; each block had two arrays of set size 4 and two arrays of set size 6. Order of arrays within each block was randomly selected for each infant. We created four orders of blocks; in each of these orders a different set of pairs was presented in the first block, second block, third block, and so on. Each block was followed by a short video clip (e.g., Sesame Street) to maintain interest. Once all 16 arrays had been presented, a second sequence of blocks was presented in which the same pairs were presented, but in different arrays (e.g., the singleton would be in a different location). This procedure continued until the infant became fussy and lost interest, or until all 32 arrays had been presented.

Data processing

Relative salience scores.—First, we calculated the relative salience of the singleton in each array (see Figure 2 for a schematic illustration). We generated a salience map for each stimulus array using the Graph-Based Visual Saliency model (GBVS; Harel et al. 2006) in MATLAB with the default settings. There are many models of salience (Bylinskii et al., 2019; Judd et al., 2012; Kümmerer et al., 2018), and many factors must be considered when choosing a particular model to answer a given research question. Because our goal was to quantify the physical salience of the singleton item, we selected a model that was a relatively pure model of physical salience. Specifically, many modern salience models include not only physical features (e.g., color, edge, orientation), but also more "top down" abstract features such as *objectness* (e.g., Kümmerer et al., 2016). In addition, we sought a model that is based on the known properties of low-level areas of visual cortex, which would be more appropriate for studying infant attention than models that are optimized to explain adult gaze patterns. In fact, Kiat et al. (2022) found that eye gaze patterns in younger infants were best predicted by the properties of lower-level areas of visual cortex. The GBVS model was therefore ideal because it is a biologically plausible model that is based on the known properties of the low-level visual cortex. Moreover, this model has been commonly used in infant and adult visual attention research (Kadooka & Franchak, 2020; Kwon et al., 2016; Mahdi et al., 2018). Indeed, Pomaranski et al. (2021) found it predicted infant eye gaze better than other salience models. Importantly, the GBVS model determines salience on the basis of both the individual features (color, edge orientation, and intensity) at each location and differences between these features across locations, which increases the salience for items that are different from the rest of the array (Harel et al. 2006). Thus, the salience values generated from the model are relative and are meaningful only within the context in which they were derived.

The resulting salience maps were the same dimensions as the stimulus arrays (1050 px by 1680 px) and contained values ranging from 0–1 for each pixel. We created AOIs that were approximately 6.3° by 6.3° (6.7 cm h by 6.7 cm w) for each of the items in each array (see Figure 1B), and calculated the salience value in each AOI by averaging the salience values at each pixel value within each AOI. This yielded 4 *raw salience scores* for the set size 4 arrays and 6 *raw salience scores* for the set size 6 arrays, corresponding to the object in each location in the arrays. Next, we calculated the *relative salience score* for the singleton by subtracting the average salience of the distractors from the salience score for the singleton, and dividing that by the sum of the singleton salience and the average distractor salience. This standardized value represents how much more (or less) salient the singleton is compared to the average of the distractors. The formula was: (singleton salience - average distractor salience) / (singleton salience + average distractor salience). This is analogous to the *Michelson contrast* formula that is widely used to quantify stimulus contrast in vision science (Pelli & Bex, 2013).

The relative salience scores for each singleton at each set size are presented in Figure 3, and it is clear that the relative salience of the singleton varied considerably from display to display. For most arrays, the singleton was more salient on average than the distractors (i.e., most of the scores were positive), but the salience varied widely across displays, and the singleton and distractor were approximately equal in salience in some displays (i.e., relative salience scores near zero). The singleton was actually *less* salient than the distractors in one case (the cup in an array of bowls). Because this was the only such case, we removed these trials from all analyses[1]. This examination of the relative salience scores demonstrates that the status of an item as a singleton is dissociable from its physical salience in this stimulus set, which will allow us to test our questions about how the relative salience contributes to the effectiveness of a singleton to capture and hold infants' attention. In addition, as seen in Figure 3, the relative salience of a singleton was typically greater when presented in arrays of 6 items, $MM_{6-items} = .25$ ($SD = .17$), than when the singleton was presented in arrays of 4 items, $MM_{4-items} = .21$ ($SD = .17$), $t(14) = 5.74$, $p < .001$, $d = .1.48$. Thus, our set size manipulation had the intended effect on salience.

---

[1] We repeated all the main analyses presented including the cup and bowl pair and the results are generally consistent, with the exception of some higher-order interactions. See supplemental results at https://osf.io/k6qev/

Infant gaze data.—The same AOIs were used for both computing relative salience (as described previously) and for quantifying eye position. The AOIs were slightly larger than the actual stimulus images to compensate for error in calibration. On each trial, we exported the number of samples recorded to have eye gaze in each of these AOIs, and converted the number samples to ms for analysis.

The initial data file contained 2747 trials (aggregated across all infants in the final sample). We removed the 129 trials involving the cup-bowl stimulus pair (rationale described above). We then inspected each remaining trial to ensure that infants started the trial looking at the screen and that they looked at the screen for a sufficient amount of time. We excluded 608 of these trials because they did not meet one or more of the following inclusion criteria: (1) a minimum of 60 ms of looking at the center of the array at the onset of the trial (to ensure that infants were looking at the center when the stimuli were presented), (2) a minimum of 200 ms of looking to the 4 or 6 stimulus items (combined) during the entire trial duration, and (3) a maximum latency of 3000 ms for the first recorded gaze to any of the items (because very long latencies indicate that that the infant was not attending to the display or that the eye tracker temporarily lost the infant’s eye gaze at the beginning of the trial). Thus, our final analyses were based on 2010 trials across the 53 infants. Overall, both groups of infants contributed similar numbers of trials, 6-month-old infants, $M = 40.19$ ($SD = 14.70$, range: 7 – 60), 8-month-old infants, $M = 35.70$ ($SD = 14.20$, range: 9 – 59), $t(50.74) = 1.12$, $p = .27$.

The duration of infants’ looking to each of the AOIs during these trials was used to derive several different measures. First, our measure of general interest was the *total looking* infants devoted to the array as a whole, calculated by summing the amount of looking to all of the AOIs on each trial. Second, to measure attention capture, we identified the *first gaze sample* recorded to one of the AOIs on each trial. From this we determined two scores: (1) the *location* of the first look and (2) the *latency* to look at the singleton. The location of infants’ first look was coded as a binary score, with a value of 1 for trials in which the first look was recorded to the AOI for the singleton and a score of 0 for trials in which the first look was recorded to an AOI for a distractor. Latency was defined as the amount of time from the start of the trial to the first moment the eye gaze was recorded in the AOI for the singleton. Note that this was not the latency of the first *eye movement* in the trial, but the time from the start of the trial until the gaze reached the *singleton*. Longer latencies, therefore, likely include one or more looks to a distractor item before looking at the singleton. Fewer trials were included in the latency analysis (1379 trials) because gaze did not enter the singleton AOI on every trial.

Our measure of attention holding was the *proportion of looking time* to the singleton; this was calculated separately for each trial by dividing the amount of looking to the singleton by the total amount of looking devoted to all the items on that trial. At set size 4, infants would be expected to look at the singleton with a proportion of 0.25 by chance (i.e., given equal looking to the items); the corresponding chance proportion is 0.167 at set size 6. We recognize that the proportion of looking time to the singleton is not independent from the probability that infants fixate the singleton first. That is, on trials in which infants fixate the singleton first necessarily allows more time within a given trial for infants to look at the singleton. However, it is possible that infants may look to the singleton first, but do not *continue* to look at the singleton. Thus, this measure reflects how effectively the singleton held infants' attention and allows for comparisons to the established literature that shows infants preferentially looked to arrays containing a singleton (Colombo et al., 1995; Quinn & Bhatt, 1998).

Data analysis

All statistical analyses were conducted in *R* (R Core Development Team, 2020); all data used in the analyses here and the corresponding analysis code can be found at https://osf.io/k6qev/. We conducted linear mixed effects analyses using the packages *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017). For each linear mixed-effects model (LMM), we conducted a follow-up robust linear mixed-effects analysis using the *robustlmm* package to investigate the robustness of the LMM results (Koller, 2016) and to ensure that any significant effects from the LMMs were not driven in part by linear model assumption violations (e.g., non-normally distributed residuals and/or heteroscedasticity). Robust mixed effects models allow for more reliable parameter estimation by downweighting extreme data points (Field & Wilcox, 2017). We used a generalized linear (logistic) mixed-effects model (GLMM) on the location of the first look, which is a binary outcome measure. Fixed effects and interactions from the models were probed for pairwise comparisons using the *emmeans* and *multcomp* packages(Hothorn et al., 2014; Lenth, 2016). For the GLMM analysis, probabilities were back-transformed from the logit scale to aid in interpretation and visualizations, but contrasts were performed on the original logit scale.

For all analyses, we included a random intercept for the grouping variable Participant to account for the repeated measures nature of the data. We also included a random intercept for each stimulus to account for the fact that the stimuli we selected to use in this study represent a *random* sample from a population of all possible singleton-distractor pair possibilities within our set of objects (Baayen et al., 2008; Westfall et al., 2017) and to increase the generalizability of our findings (Yarkoni, 2019). We also included the *set size* of the stimulus arrays (categorical; four = 4-item arrays, six = 6-item arrays), *age* (categorical; 6 = six-months, 8 = eight-months), and the relative *salience* score of the singleton for each stimulus array (continuous; range: −.12 to .64) as fixed effects in all models. In addition, we included the 2-way interactions between each of these variables and the 3-way interaction term between age, set size, and singleton salience to test whether singleton salience has a different effect on infants' looking behavior to the singleton as a function of age and/or set size. Thus, all models were specified as follows:

$$\text{DV} \sim \text{Age} * \text{ Set Size } * \text{ Relative Salience } + (\ldots|\text{Participant}) + (\ldots|\text{Stimulus})$$

Results

General characterization of infants' looking

We began by examining task engagement. All supplemental tables and results can be found here: https://osf.io/k6qev/. Overall, 6-month-old and 8-month-old infants were approximately equally engaged in this task. They completed approximately the same number of trials and exhibited similar amounts of looking to the different array types (summed across all AOIs for each array, see Table S1 in supplemental materials). This impression was confirmed by an LMM (specified above) with infants' total amount of looking on each trial as the dependent variable (DV); this model revealed no significant main (fixed) effects of set size, age, or singleton salience and no significant interactions (see Table S2 in supplemental materials for full results). Thus, 6- and 8-month-old infants had the same general interest in the arrays, infants did not look longer overall to the arrays containing a relatively *more* salient singleton, and arrays with larger numbers of items did not elicit different overall amounts of looking.

Attention-capture: First looks to the singleton

Our next analyses examined attention capture by the singleton. We first examined where infants looked *first* on each trial, collapsed across all stimulus arrays. Because there were more distractors in each array than singletons, by chance more trials should have first looks directed to the distractor — that is, 25% at set size 4 and 16.67% at set size 6. Collapsed across age and relative salience, at set size 4 there were 423 first looks to the singleton and 597 first looks to a distractor, and at set size 6 there were 318 first looks to the singleton and 672 first looks to a distractor. For each age we conducted *t*-tests to compare to chance the proportion of first looks directed to the singleton, separately for each set size. Overall, infants looked first to the singleton more often than expected by chance for both set sizes in both age groups (see Table 1). Thus, our stimulus displays led to the same pattern observed in prior studies, in which singletons attracted infants' attention (Adler & Gallego, 2014; Adler & Orprecio, 2006; Colombo et al., 1995; Gerhardstein et al., 1999).

Next we asked how attention capture, as measured by the proportion of first looks to the singleton, was related to the relative salience of the singleton. The proportion of trials with first looks to the singleton, broken down by the relative salience of that singleton, are presented in Figure 4A. A visual inspection of Figure 4A shows that the location of infants' first eye gaze was influenced by the level of relative salience of the singleton (for ease of evaluating differences due to the relative salience of the singleton, within each set size the data are arranged in ascending order of relative salience along the *x*-axis). In general, the probability that infants directed their first look at the singleton increased with the relative salience of that singleton.

To confirm this impression, we fitted a GLMM (specified in the Data Analysis section earlier) with the direction of the first look (toward the singleton versus toward any distractor) on each trial as the dependent variable. The results from this model are presented in Table 2. Confirming the pattern shown in Figure 4A, we observed a positive effect of Salience, indicating the probability that the first look was directed toward the singleton increased significantly as the difference in salience between the singleton and the distractors increased. The estimated marginal probabilities in Figure 4B show that across both ages, the probability that the first look was directed to the singleton increased systematically as the relative salience of the singleton increased.

In addition, there was a significant main (fixed) effect of set size, revealing that in general infants were more likely to direct their gaze to the singleton first at set size four compared to set size six (the Set Size effect in Table 2), as would be expected given the different chance rates at the two set sizes. Figure 4B shows that both groups of infants directed *fewer* first looks to the singleton when the set size was larger (although this might simply reflect differences in chance levels between the two set sizes). We did not find a significant effect of age in this measure.

We also conducted *t*-tests on the marginal means to determine whether infants' looking to the singleton differed from chance at the specified levels of relative salience in Figure 4B. Because the previous model included both set sizes (and thus there were complications due to different chance levels at each set size), to generate the marginal means for this analysis we refit the model separately for each set size, omitting the fixed effect of set size. When the singleton and distractors were equally salient (zero relative salience), the estimated probability that the first look was directed to the singleton was numerically near chance (see leftmost bars in each panel of Figure 4B) and did not differ significantly from chance for either age group, all Bonferroni-corrected *p*s >.99. When the salience difference was .14 at set size 4, estimated probabilities were significantly different from chance for 8-month-old infants, Bonferroni-corrected $p = .05$, and marginally significantly different from chance for 6-month-old infants, Bonferroni-corrected $p = .06$, but not significantly different than chance at set size 6 (*p*s >.99). For all the other levels of salience depicted in Figure 4B, the estimated probabilities were significantly greater than chance for both set sizes and both age groups, all Bonferroni-corrected *p*s < .001 (see detailed statistics in Table S3 in supplementary materials at https://osf.io/k6qev/). These results show that variations in the physical salience of the singleton impacted the likelihood that gaze was directed first to the singleton. Moreover, when the singleton was no more salient than the distractors (relative salience of zero), there was no evidence that gaze was more likely to be directed to the singleton than to the distractors. That is, the estimated probability was near chance when the relative salience was zero (leftmost bars in each panel of Figure 4B).

Attention-capture: Latency to look at the singleton

Our next analysis examined attention capture by evaluating the latency for gaze to reach the singleton. The latencies from each infant for each individual array are presented in Figure 5A as a function of infant age, relative salience, and set size. Each point in the figure represents the latency for one infant on one trial, and is the amount of time from the start of the trial to the first time the infant's eye gaze was recorded in the singleton AOI. Latencies are shown as a function of the difference in salience between the singleton and distractors. Relative salience is shown on the *x*-axis. Latency to look at the singleton is on the *y*-axis, with shorter latencies at the bottom and longer latencies at the top. Note that chance is undefined for the latency variable, so the results were not compared to chance.

Several things are clear from a visual inspection of Figure 5A. First, infants were generally slower to look at the singleton when the singleton was similar in salience to the distractors. In addition, latencies were longer for the larger set size compared to the smaller set size. Figure 5A also suggests an age effect. Specifically, the slopes for the 6-month-old infants appear to be steeper than those for the 8-month-olds, particularly for the set size 4 arrays, pointing to the possibility that salience had a larger effect on the younger infants.

To confirm these observations, we fitted both an LMM and a robust LMM (specified in the Data Analysis section earlier) with infants' latencies to look to the singleton on each trial as the dependent variable. Because the DV and main predictor variable of interest (relative singleton salience) are on very different measurement scales, we adopted best practices (Robinson & Schumacker, 2009) and scaled the DV using the *scale* R function directly in the model statement. Thus, the results from this model are presented in standardized units in Table 3. The effect of relative salience on latency was significant for both the LMM and robust LMM models. As can be seen in Figure 5B, latencies declined as the relative salience of the singleton increased (the negative Salience effects in Table 3). This effect was qualified by a significant interaction between salience and age in both models (see Table 3), confirming that the effect of salience was greater for the younger infants than for the older infants.

The estimated marginal means from the main LMM analysis are presented in Figure 5B. When there was a large relative difference in salience, both older and younger infants had fast latencies to look to the singleton. As the difference in salience decreased, infants' latencies decreased, and this effect was larger for the younger infants than for the older infants. Although inspection of Figure 4 suggests this age difference is true primarily in the set size 4 condition, the 3-way interaction between age, salience, and set size was not significant in either the LMM or the robust LMM. Thus, although the interaction appears to be driven by the set size 4 condition, we have insufficient evidence to conclude that the effect varies as a function of set size. It is also worth noting that the estimates for the interaction differ substantially for the LMM (1.36) and robust LMM (.68), suggesting that although the interaction was significant for both analyses, the LMM may overestimate the size of the effect. Nonetheless, these results provide additional evidence that the degree to which a given singleton captures attention varies according to the physical salience of that singleton.

The main and robust LMMs yielded somewhat different outcomes for the main (fixed) effects of age and set size. This difference is not surprising given the tendency for latency variables such as response time latencies to contain extreme data points (Csibra et al., 2016). Specifically, the LMM revealed a significant negative fixed effect of age, but the robust LMM did not. Thus, any conclusion about overall differences between younger and older infants (i.e., that older infants were faster than younger infants in general) should be drawn with caution. In addition, the robust LMM revealed a significant positive effect of set size, but the robust LMM did not. These findings provide inconsistent evidence that infants had slower latencies in the larger arrays compared to the smaller arrays across both age groups and at the average level of salience (see Set Size effects in Table 3).

Attention-holding: Maintaining gaze at the singleton

In our final analyses, we examined attention-holding, quantified as the proportion of infants' total time looking at all items in the array devoted to the singleton. As in the attention-capture analysis, we first compared the proportion of looking time to chance, averaged across arrays for each infant. If the singleton was no more interesting than any of the distractors, then infants should devote 25% of their looking to the singleton at set size 4 and 16.67% at set size 6. Table 4 shows that, averaged across all stimulus arrays, the mean proportion of looking toward the singleton was significantly greater than expected by chance at both set sizes for both age groups (see Table 4).

The single-trial, single-participant values are presented in Figure 6A. A visual inspection of Figure 6A clearly shows that attention-holding increased as the relative salience of the singleton increased. Comparison of the illustrative regression lines suggest that there may have been a greater effect of salience in the younger infants (e.g., steeper slope at set size 4) than in the older infants. Moreover, when the singleton and distractors were approximately equally salient (relative salience at or near zero), the illustrative regression lines were close to chance, providing no evidence that singleton status per se impacts attention-holding.

To confirm these observations, we fitted an LMM and a robust LMM (specified earlier) with infants' proportion of looking to the singleton on each trial as the dependent variable (the model results are in Table 5). Both the LMM and the robust LMM revealed a significant main (fixed) effect of singleton salience, driven by a positive relation between infants' looking to the singleton and the relative singleton salience (see the Salience effect in Table 5). As observed for the latency analyses, the salience by age interaction was significant in both the LMM and robust LMM analyses; in this case the estimates from both models are nearly identical. There was a larger effect of salience in the younger infants than in the older infants. As observed in the latency analyses, the 3-way interaction between age, salience, and set size was not significant, but inspection of Figure 6 suggests that the salience by age interaction was most pronounced at set size 4. As we observed for the latency analyses, the LMM and robust LMM yielded some different outcomes. Specifically, the robust LMM, but not the LMM, showed a significant main effect of age.

Lastly, comparisons of the marginal means to chance revealed that when the relative difference in salience between the singleton and distractors was equal (zero relative salience for both set sizes), the estimated proportion of looking to the singleton was not significantly different from chance in either group of infants at either set sizes (all Bonferroni-corrected *p*s ≥.19, see Figure 6B). At all other levels of salience depicted along the *x*-axis in Figure 6B, the proportion of looking differed significantly from chance for both groups of infants at both set sizes (all Bonferroni-corrected *p*s < .0001; see detailed statistics in Table S4 supplementary materials at https://osf.io/k6qev/). Thus, as was observed for attention-getting, we found a strong effect of physical salience on attention-holding but no evidence for an effect of singleton status per se.

Replication study

We report in the supplemental materials a replication of the physical salience effects with a separate group of 6-month-old infants viewing set size 4 arrays (see full results here: https://osf.io/k6qev/). The results from this partial replication study are consistent with the results observed in the primary sample showing that quantitative variations in physical salience were related to quantitative variations in infants' attention. Specifically, the probability and latency of attention capture and the proportion of time infants spent looking to the singleton was related to increases in relative salience of the singleton. These results demonstrate the robustness of the effects of relative salience on infants' attention capture and maintenance.

Discussion

Salience versus Singleton Status

This study examined the influence of physical salience and singleton status on infants' attention, as well as changes in attention between 6 and 8 months of age. The main question was whether attention to singletons would be determined solely by physical salience, solely by whether or not an item was a singleton, or by both salience and singleton status. We found clear evidence for a role of physical salience: quantitative variations in singleton salience had a very large impact on the probability that attention was initially attracted to the singleton, the latency of the initial allocation of attention to the singleton, and the amount of time infants spent looking at the singleton rather than to the distractors.

This effect at first may seem somewhat surprising, and contrary to traditional views that singleton stimuli always "pop out" and automatically capture attention. Further support for our conclusions is found in a replication study we ran with a separate group of 66 6-month-old infants using the set-size 4 arrays used in this study. As we reported in the main analyses here, in this replication sample relative salience predicted where infants looked first, how quickly they looked at the singleton, and how long they fixated the singleton once they looked at it (the full details can be found here https://osf.io/k6qev/). Thus, these two sets of results taken together provide strong support for the conclusion that both attention-getting and attention-holding for a given singleton is influenced by the relative salience of that singleton.

These results are consistent with research in adults showing that quantitative variations in salience lead to quantitative variations in the probability and latency of attention capture (Gaspar et al., 2016; Luck et al., 2006; Töllner et al., 2011). They are also consistent with research in both infants and adults showing that physical salience impacts attention in natural scenes that do not contain singletons (Itti, 2005; Koehler et al., 2014; Pomaranski et al., 2021).

By contrast, we found no evidence that singleton status per se plays a role in infant attentional allocation. Although singletons are often physically salient, this is not universally true, and we found a broad range of relative singleton salience when we constructed stimulus arrays by randomly choosing one object to be the singleton and another for the distractors (see Figure 2). Indeed, the singleton object was sometimes equal or lower in physical salience than the distractors, as quantified by a model of salience based on the physiological properties of low-level visual areas. This made it possible for us to test the effect of singleton status independently of the relative salience of the singleton. When the singleton was approximately equal in salience to the distractors (relative salience near zero), we found no evidence that attention was captured or held more by the singleton than by the distractors. That is, the singleton in these displays was no more likely than the distractors to attract the first shift of gaze or to hold attention throughout the display period. This is a somewhat surprising result, because it is entirely plausible that the mere fact that one item in a display differs from the rest would be intrinsically interesting to infants. Thus, the present results indicate that singletons attract and hold attention in infants only to the extent that they are more salient than the rest of the display and not because they are singletons per se.

These results suggest a more nuanced interpretation of previous studies of singletons in infants, in which the salience of the singletons was not quantified (Adler & Gallego, 2014; Adler & Orprecio, 2006; Colombo et al., 1995; Gerhardstein & Rovee-Collier, 2002). In these studies, the singleton was presumed to be the most salient item in the display (Adler & Gallego, 2014), and pop-out was implicitly assumed to be an all-or-none phenomenon. The present results suggest that effects of the singletons in these studies were a result of salience rather than the target's status as a singleton. Moreover, the present results show that both attention-getting and attention-holding mechanisms are sensitive to the specific level of physical salience, which is a continuous variable rather than an all-or-none property.

Effects of Set Size

We also found that infants' attention is influenced by the number of items in the array (the set size). Specifically, we observed that infants directed more first looks to singletons and were faster to look at singletons at set size 4 than at set size 6. This is not what might be expected for items that pop out in an all-or-none manner. However, because the effect of set size on response latencies was observed only in the robust model, conclusions about set size on latency should be drawn with caution. Moreover, chance varies with set size, which produces set size effects for visually unique targets even in adults (Palmer, 1994).

Similar results were reported by Goldknopf et al. (2019), who found that infants had longer latencies and fewer first looks directed to a singleton in a visual search array as set size increased. Our results and those of Goldknopf et al. contrast with the results reported by Adler and Orprecio (2006), who observed that young infants' latencies to look at a singleton target did *not* vary with set size. However, they used a different measure of latency, which may have contributed to the differences in results. Indeed, the fastest latencies we observed (at the highest values of salience) were overall *slower* than the latencies observed by Adler et al. across all the set sizes presented in their study. More research is needed to determine whether and when set size impacts attention capture in infants.

Contributions to our Understanding of Attention Development

Consistent with the literature, we observed that the effects of salience decreased over age. Although the 3-way interactions including age, set size, and salience were not significant, this finding appeared to be driven by the effect of salience on set size 4 arrays, highlighting the fact that in addition to physical salience, many other factors, including stimulus complexity, impact how infants direct and sustain their attention. It is worth noting that we observed only small age effects, likely because we focused here on the difference between 6- and 8-month-old infants, and there are developmental changes across the first year in the factors that influence attention. However, our results corroborate other findings in the literature that there are larger effects of salience on attention in 4- to 6-month-old infants than in infants who are older than 6 months (Frank et al., 2009; Kwon et al., 2016; Pomaranski et al., 2021). Specifically, in the work presented here, although both groups of infants had shorter latencies to look at the singleton and increased their maintenance of attention on the singleton as relative salience increased, 6-month-old infants showed larger effects than did 8-month-old infants. Note that this pattern cannot be explained by noisier data or more off-task performance in the younger infants, which would have led to weaker rather than stronger effects.

This pattern of results is consistent with the understanding that, over the course of infant development, higher level cortical regions play an increasing role in attentional control, leading to an increasing influence of features other than low-level physical features of the stimulus (Colombo, 2001; Johnson, 1990). In addition, although attention develops quantitatively during infancy, results like those presented here suggest that the general mechanisms of attention in infants are similar to those in adults. Specifically, like the adult visual system, the infant visual system may compute a “priority map” representing the attentional priority of each location in the input (Awh et al., 2012) based on the degree of physical salience, learning, and goals. Developmental changes in the visual system across infancy presumably causes shifts in the relative influence of these factors on the priority map. In addition, because attention is multiply determined, these factors will interact in different ways depending not only on infant age but on other aspects of the context. In fact, the interaction between age and salience observed here contrasts with other results with complex scenes in which salience may become a *stronger* predictor of infants’ attention with age (Amso et al., 2014; Franchak et al., 2016), or in which the predictive effect of salience does not change across age (Renswoude et al., 2019). Future research may uncover how age-related changes in the effect of physical salience depend on the task, stimuli, and models of physical salience.

Acknowledgments

This research and preparation of this manuscript were made possible by NIH grants R01EY022525 and R01EY030127 awarded to LMO. LMC was supported by training grant EY015387. We thank the students and staff in the Infant Cognition Laboratory at the University of California, Davis, for their help with data collection. Samantha Mitsven is now at the Department of Psychology, University of Miami. This study was not preregistered and all data, analysis code, and supplemental materials can be found in the Open Science Framework (OSF; repository: https://osf.io/k6qev/ (DOI 10.17605/OSF.IO/K6QEV, DeBolt, Oakes, & Luck, 2022).

References

Adler SA, & Gallego P (2014). Search asymmetry and eye movements in infants and adults. Attention, Perception & Psychophysics, 76(6), 1590–1608. 10.3758/s13414-014-0667-6

Adler SA, & Orprecio J (2006). The eyes have it: Visual pop-out in infants and adults. Developmental Science, 9(2), 189–206. 10.1111/j.1467-7687.2006.00479.x [PubMed: 16472320]

Alexander RG, & Zelinsky GJ (2011). Visual similarity effects in categorical search. Journal of Vision, 11(8). 10.1167/11.8.9

Amso D, Haas S, & Markant J (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. PloS One, 9(1). 10.1371/journal.pone.0085701

Awh E, Belopolsky AV, & Theeuwes J (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. Trends in Cognitive Sciences, 16(8), 437–443. 10.1016/j.tics.2012.06.010 [PubMed: 22795563]

Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59(4), 390–412. 10.1016/j.jml.2007.12.005

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Dai B, Grothendieck G, Green P, & Bolker MB (2015). Package “lme4.” Convergence, 12(1), 2. http://dk.archive.ubuntu.com/pub/pub/cran/web/packages/lme4/lme4.pdf

Bravo M, & Blake R (1990). Preattentive vision and perceptual groups. Perception, 19(4), 515–522. 10.1068/p190515 [PubMed: 2096369]

Brennan WM, Ames EW, & Moore RW (1966). Age Differences in Infants' Attention to Patterns of Different Complexities. Science, 151(3708), 354 LP – 356. http://science.sciencemag.org/content/151/3708/354.abstract [PubMed: 5903347]

Bylinskii Z, Judd T, Oliva A, Torralba A, & Durand F (2019). What Do Different Evaluation Metrics Tell Us About Saliency Models? IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(3), 740–757. 10.1109/TPAMI.2018.2815601 [PubMed: 29993800]

Champely S (2020). pwr: Basic Functions for Power Analysis. R package version 1.3–0. 2020. URL https://CRAN.R-Project.Org/package=Pwr.

Cohen LB (1972). Attention-Getting and Attention-Holding Processes of Infant Visual Preferences. In Child Development (Vol. 43, Issue 3, p. 869). 10.2307/1127638 [PubMed: 5056611]

Colombo J (2001). The development of visual attention in infancy. Annual Review of Psychology, 52, 337–367.

Colombo J, & Cheatham CL (2006). The emergence and basis of endogenous attention in infancy and early childhood. Advances in Child Development and Behavior, 34, 283–322. 10.1016/S0065-2407(06)80010-8 [PubMed: 17120808]

Colombo J, Kannass KN, Shaddy DJ, Kundurthi S, Maikranz JM, Anderson CJ, Blaga OM, & Carlson SE (2004). Maternal DHA and the development of attention in infancy and toddlerhood. Child Development, 75(4), 1254–1267. 10.1111/j.1467-8624.2004.00737.x [PubMed: 15260876]

Colombo J, Ryther JS, Frick JE, & Gifford JJ (1995). Visual pop-out in infants: Evidence for preattentive search in 3- and 4-month-olds. Psychonomic Bulletin & Review, 2(2), 266–268. 10.3758/BF03210968 [PubMed: 24203663]

Csibra G, Hernik M, Mascaro O, Tatone D, & Lengyel M (2016). Statistical treatment of looking-time data. Developmental Psychology, 52(4). 10.1037/dev0000083

DeBolt MC, Oakes LM, Luck SJ. (2022). A new perspective on the role of physical salience in visual search: Graded effect of salience on infants' attention. 10.17605/OSF.IO/K6QEV

Di Giorgio E, Leo I, Pascalis O, & Simion F (2012). Is the face-perception system human-specific at birth? Developmental Psychology, 48(4), 1083–1090. 10.1037/a0026521 [PubMed: 22142186]

Di Giorgio E, Turati C, Altoè G, & Simion F (2012). Face detection in complex visual displays: an eye-tracking study with 3- and 6-month-old infants and adults. Journal of Experimental Child Psychology, 113(1), 66–77. 10.1016/j.jecp.2012.04.012 [PubMed: 22695758]

Duncan J, & Humphreys G (1992). Beyond the search surface: visual search and attentional engagement [Review of *Beyond the search surface: visual search and attentional engagement*]. Journal of Experimental Psychology. Human Perception and Performance, 18(2), 578–588; discussion 589–593. 10.1037//0096-1523.18.2.578 [PubMed: 1593236]

Duncan J, & Humphreys GW (1989). Visual serach and stimulus similarity. Psychological Review, 96(3), 433–458. [PubMed: 2756067]

Fantz RL (1964). Visual Experience in Infants : Decreased Attention to Familiar Patterns Relative to Novel Ones. Science, 146(3644), 668–670. [PubMed: 14191712]

Fecteau JH, & Munoz DP (2006). Salience, relevance, and firing: a priority map for target selection. Trends in Cognitive Sciences, 10(8), 382–390. 10.1016/j.tics.2006.06.011 [PubMed: 16843702]

Field AP, & Wilcox RR (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. Behaviour Research and Therapy, 98, 19–38. 10.1016/j.brat.2017.05.013 [PubMed: 28577757]

Franchak JM, Heeger DJ, Hasson U, & Adolph KE (2016). Free Viewing Gaze Behavior in Infants and Adults. Infancy: The Official Journal of the International Society on Infant Studies, 21(3), 262–287. 10.1111/infa.12119 [PubMed: 27134573]

Frank MC, Vul E, & Johnson SP (2009). Development of infants' attention to faces during the first year. Cognition, 110(2), 160–170. 10.1016/j.cognition.2008.11.010 [PubMed: 19114280]

Gaspar JM, Christie GJ, Prime DJ, Jolicœur P, & McDonald JJ (2016). Inability to suppress salient distractors predicts low visual working memory capacity. Proceedings of the National Academy of Sciences of the United States of America, 113(13), 3693–3698. 10.1073/pnas.1523471113 [PubMed: 26903654]

Gaspelin N, Ruthruff E, & Lien M-C (2016). The problem of latent attentional capture: Easy visual search conceals capture by task-irrelevant abrupt onsets. Journal of Experimental Psychology. Human Perception and Performance, 42(8), 1104–1120. 10.1037/xhp0000214 [PubMed: 26854530]

Geng JJ, & DiQuattro NE (2010). Attentional capture by a perceptually salient non-target facilitates target processing through inhibition and rapid rejection. Journal of Vision, 10(6), 5–5. 10.1167/10.6.5

Gerhardstein P, Renner P, & Rovee-Collier C (1999). The roles of perceptual and categorical similarity in colour pop-out in infants. The British Journal of Developmental Psychology, 17(3), 403–420. 10.1348/026151099165366

Gerhardstein P, & Rovee-Collier C (2002). The Development of Visual Search in Infants and Very Young Children. Journal of Experimental Child Psychology, 81(2), 194–215. 10.1006/jecp.2001.2649 [PubMed: 11786009]

Gliga T, Elsabbagh M, Andravizou A, & Johnson M (2008). Faces attract infants' attention in complex displays. In Infancy (Vol. 7, Issue S1). 10.1080/15250000903144199

Gluckman M, & Johnson SP (2013). Attentional capture by social stimuli in young infants. Frontiers in Psychology, 4(AUG), 1–7. 10.3389/fpsyg.2013.00527 [PubMed: 23382719]

Hothorn T, Bretz F, Westfall P, Heiberger RM, Schuetzenmeister A, & Scheibe S (2014). Multcomp: simultaneous inference in general parametric models. R Package Version, 1–3. http://www2.uaem.mx/r-mirror/web/packages/multcomp/

Itti L (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. Visual Cognition, 12(6), 1093–1123. 10.1080/13506280444000661

Jiang YV, Sha LZ, & Sisk CA (2018). Experience-guided attention: Uniform and implicit. Attention, Perception & Psychophysics, 80(7), 1647–1653. 10.3758/s13414-018-1585-9

Johnson MH (1990). Cortical maturation and the development of visual attention in early infancy. Journal of Cognitive Neuroscience, 2(2), 81–95. 10.1162/jocn.1990.2.2.81 [PubMed: 23972019]

Judd T, Durand F, & Torralba A (2012). A Benchmark of Computational Models of Saliency to Predict Human Fixations. Mit-Csail-Tr-2012, 1, 1–7. https://doi.org/1721.1/68590

Kadooka K, & Franchak JM (2020). Developmental changes in infants' and children's attention to faces and salient regions vary across and within video stimuli. Developmental Psychology, 56(11), 2065–2079. 10.1037/dev0001073 [PubMed: 32915049]

Kiat JE, Luck SJ, Beckner AG, Hayes TR, Pomaranski KI, Henderson JM, & Oakes LM (2022). Linking patterns of infant eye movements to a neural network model of the ventral stream using representational similarity analysis. In Developmental Science (Vol. 25, Issue 1). 10.1111/desc.13155

Koehler K, Guo F, Zhang S, & Eckstein MP (2014). What do saliency models predict? Journal of Vision, 14(3), 14. 10.1167/14.3.14

Koller M (2016). robustlmm : An R Package for Robust Estimation of Linear Mixed-Effects Models. Journal of Statistical Software, 75(6), 24. https://doi.org/Koller Manuel (2016). robustlmm : An R Package for Robust Estimation of Linear Mixed-Effects Models. Journal of statistical software, 75(6), pp. 1–24. UCLA Statistics 10.18637/jss.v075.i06 <10.18637/jss.v075.i06> [PubMed: 32655332]

Kümmerer M, Bylinskii Z, Judd T, Borji A, Itti L, Durand FD, Oliva A, & Torralba A (2018). MIT/Tübingen Saliency Benchmark. MIT/Tübingen Saliency Benchmark. https:/saliencey.tuebingen.ai

Kümmerer M, Wallis TSA, & Bethge M (2016). DeepGaze II: Reading fixations from deep features trained on object recognition. In arXiv [cs.CV]. arXiv. http://arxiv.org/abs/1610.01563

Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest package: tests in linear mixed effects models. Journal of Statistical Software, 82(13). https://orbit.dtu.dk/en/publications/id(2e6e11c6-6a44-43df-b809-d96b897d9bf2).html

Kwon M-K, Setoodehnia M, Baek J, Luck SJ, & Oakes LM (2016). The development of visual search in infancy: Attention to faces versus salience. Developmental Psychology, 52(4), 537–555. 10.1037/dev0000080 [PubMed: 26866728]

Lenth RV (2016). Least-Squares Means: TheRPackagelsmeans. In Journal of Statistical Software (Vol. 69, Issue 1). 10.18637/jss.v069.i01

Luck SJ, Fuller R, Braun E, Robinson B, Summerfelt A, & Gold J (2006). The speed of visual attention in schizophrenia: Electrophysiological and behavioral evidence. In Schizophrenia Research (Vol. 85, Issues 1–3, pp. 174–195). 10.1016/j.schres.2006.03.040 [PubMed: 16713184]

Mahdi A, Su M, Schlesinger M, & Qin J (2018). A Comparison Study of Saliency Models for Fixation Prediction on Infants and Adults. IEEE Transactions on Cognitive and Developmental Systems, 10(3), 485–498. 10.1109/TCDS.2017.2696439

Malcolm GL, & Henderson JM (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. Journal of Vision, 9(11), 1–13. 10.1167/9.11.8 [PubMed: 19761316]

Mitsven SG, Cantrell LM, Luck SJ, & Oakes LM (2018). Visual short-term memory guides infants' visual attention. Cognition, 177(July 2017), 189–197. 10.1016/j.cognition.2018.04.016 [PubMed: 29704857]

Oakes LM, & Amso D (2018). The Development of Visual Attention. In Ghetti S (Ed.), The Steven's Handbook of Experimental Psychology and Cognitive Neuroscience (Fourth Edi). Wiley. 10.1002/9781119170174.epcn401

Palmer J (1994). Set-size effects in visual search: the effect of attention is independent of the stimulus for simple tasks. Vision Research, 34(13), 1703–1721. 10.1016/0042-6989(94)90128-7 [PubMed: 7941377]

Pelli DG, & Bex P (2013). Measuring contrast sensitivity. Vision Research, 90, 10–14. 10.1016/j.visres.2013.04.015 [PubMed: 23643905]

Pomaranski KI, Hayes TR, Kwon M-K, Henderson JM, & Oakes LM (2021). Developmental changes in natural scene viewing in infancy. Developmental Psychology, 57(7), 1025–1041. 10.1037/dev0001020 [PubMed: 34435820]

Quinn PC, & Bhatt RS (1998). Visual pop-out in young infants: Convergent evidence and an extension. Infant Behavior & Development, 21(2), 273–288. 10.1016/S0163-6383(98)90006-6

Renswoude DR, Visser I, Raijmakers MEJ, Tsang T, & Johnson SP (2019). Real-world scene perception in infants: What factors guide attention allocation? In Infancy (Vol. 24, Issue 5, pp. 693–717). 10.1111/infa.12308 [PubMed: 32677279]

Richards JE (2010). The development of attention to simple and complex visual stimuli in infants: Behavioral and psychophysiological measures. Developmental Review: DR, 30(2), 203–219. 10.1016/j.dr.2010.03.005 [PubMed: 20526452]

Robinson C, & Schumacker RE (2009). Interaction effects: centering, variance inflation factor, and interpretation issues. Multiple Linear Regression Viewpoints, 35(1), 6–11. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.620.5853&rep=rep1&type=pdf

Simpson EA, Husband HL, Yee K, Fullerton A, & Jakobsen KV (2014). Visual search efficiency is greater for human faces compared to animal faces. Experimental Psychology, 61(6), 439–456. 10.1027/1618-3169/a000263 [PubMed: 24962122]

Stechler G, & Latz E (1966). Some Observations on Attention and Arousal in the Human Infant. Journal of the American Academy of Child Psychiatry, 5(3), 517–525. 10.1016/S0002-7138(09)62098-7

Theeuwes J (1992). Perceptual selectivity for color and form [Review of *Perceptual selectivity for color and form*]. Perception & Psychophysics, 51(6), 599–606. 10.3758/bf03211656 [PubMed: 1620571]

Theeuwes J (2019). Goal-driven, stimulus-driven, and history-driven selection. Current Opinion in Psychology, 29, 97–101. 10.1016/J.COPSYC.2018.12.024 [PubMed: 30711911]

Töllner T, Zehetleitner M, Gramann K, & Müller HJ (2011). Stimulus Saliency Modulates Pre-Attentive Processing Speed in Human Visual Cortex. In PLoS ONE (Vol. 6, Issue 1, p. e16276). 10.1371/journal.pone.0016276 [PubMed: 21283699]

Treisman AM, & Gormican S (1988). Feature analysis in early vision - evidence from search asymmetries..pdf. Psychological Review, 95(1), 15–48. 10.1037/0033-295X.95.1.15 [PubMed: 3353475]

Westfall J, Nichols TE, & Yarkoni T (2017). Fixing the stimulus-as-fixed-effect fallacy in task fMRI. Wellcome Open Research, 1.

Wolfe JM (1998a). What can 1,000,000 trials tell us about visual search? Psychological Science, 9(1), 33–39.

Wolfe JM (1998b). What Can 1 Million Trials Tell Us About Visual Search? In Psychological Science (Vol. 9, Issue 1, pp. 33–39). 10.1111/1467-9280.00006

Wolfe JM (2015). Jeremy Wolfe. In Current Biology (Vol. 25, Issue 9, pp. R348–R349). 10.1016/j.cub.2015.02.028 [PubMed: 26146677]

Wolfe JM-H, Võ ML, Evans KK, & Greene MR. (2011). Visual search in scenes involves selective and nonselective pathways. In Trends in Cognitive Sciences (Vol. 15, Issue 2, pp. 77–84). 10.1016/j.tics.2010.12.001 [PubMed: 21227734]

Yarkoni T (2019). The generalizability crisis. PsyArXiv, 1–26. 10.31234/osf.io/jqw35

Significance Statement

This study shows that how infants focus their attention is influenced by several different factors, including the infants' age and the characteristics of images (such as color and brightness). These results help us understand how infants process visual information.

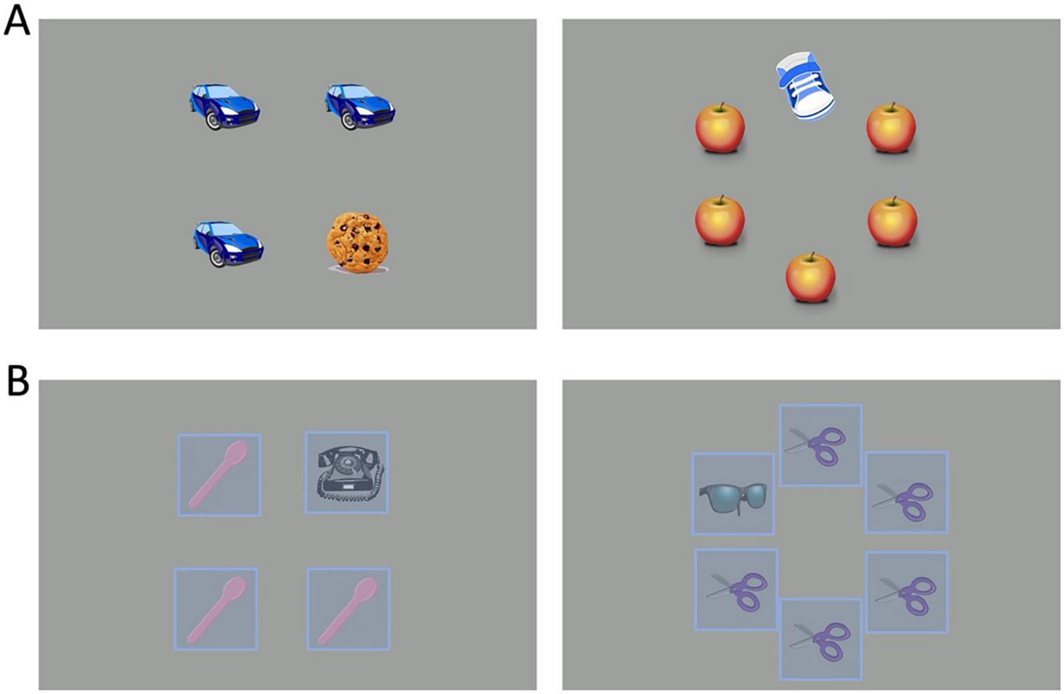

Figure 1.

Schematic displays of stimulus arrays presented to infants (A) and schematic areas of interest (AOIs) used for the eye tracking data analysis (B) for both set size 4 arrays (left) and set size 6 arrays (right). Note that the images composing the arrays were obtained from Pixabay (www.pixabay.com) and are freely available for public use and are included here for illustration purposes only. All the stimuli that were used in the study are available at https://osf.io/k6qev/.

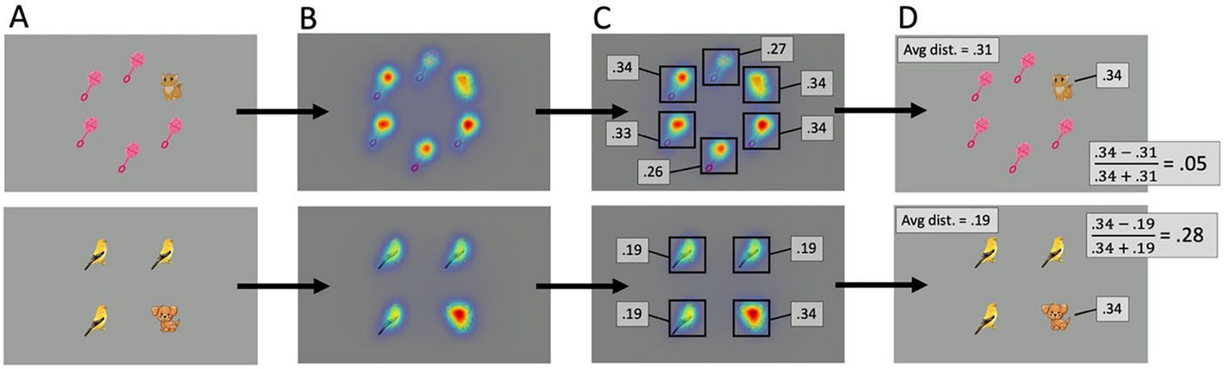

Figure 2.

A schematic depiction of the relative salience score calculation process with the incremental steps illustrated in columns A through D. **A**: Example arrays with a set size 6 array on top and a set size 4 array on bottom. **B**: For each stimulus array, a salience map is generated using the GBVS toolbox in Matlab. **C**: The eye tracking AOIs are applied to the salience maps generated in the previous step and the pixel values *within* the dimensions of each AOI are averaged to create the *raw salience scores.* **D**: Next, the raw salience scores for the distractors are averaged to the *distractor average salience score* (i.e., the .31 value in top array in column D and .19 value in bottom array). The *relative salience score* (what is used in the main analyses) is computed by dividing the *salience difference score* (*singleton salience score - distractor average salience score*) by the sum of the *distractor average salience score* and the *singleton salience score*. Note that the images composing the arrays were obtained from Pixabay (www.pixabay.com) and are freely available for public use and are included here for illustration purposes only. All the stimuli that were used in the study are available at https://osf.io/k6qev/.

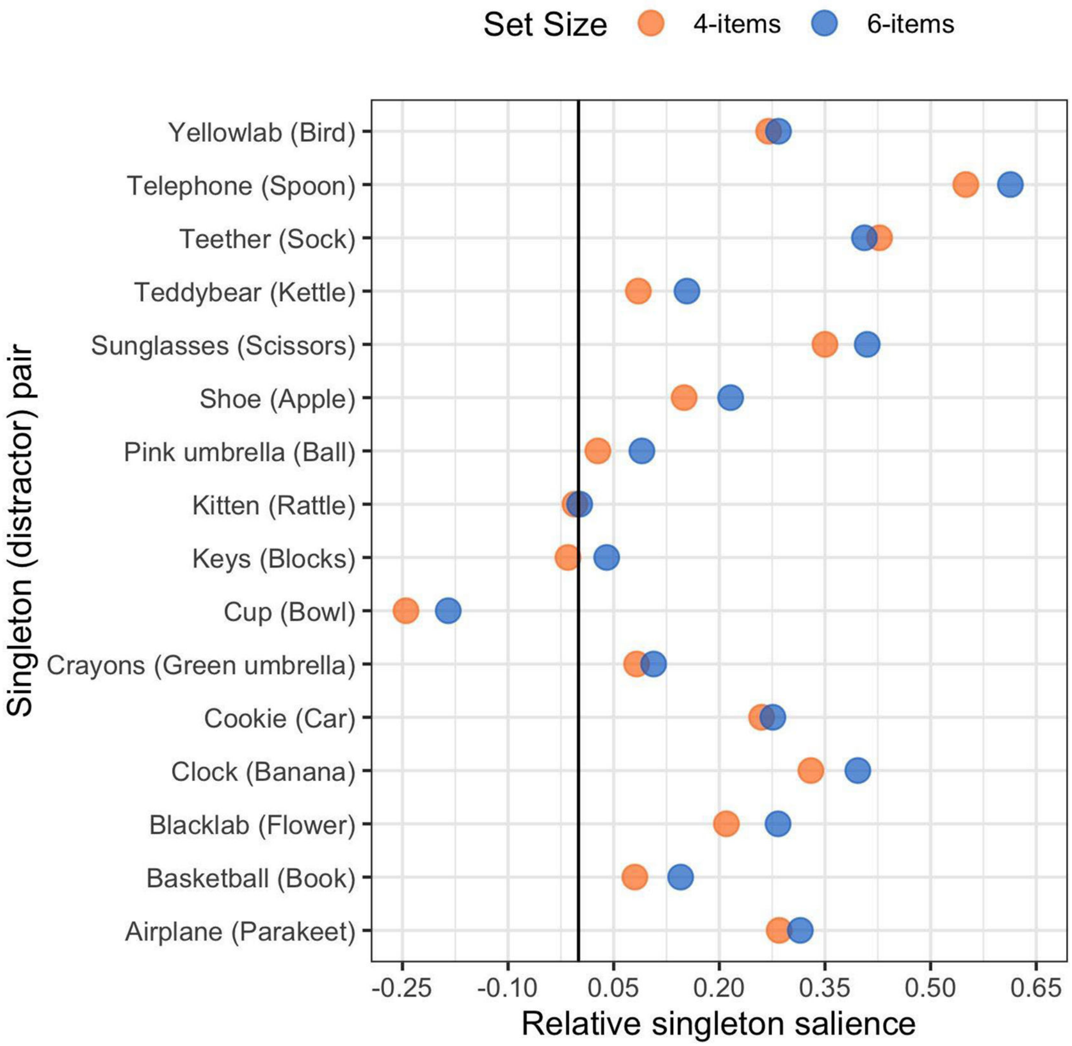

Figure 3.

Relative salience scores (x-axis) for each singleton-distractor pair (y-axis). The set size of each array is indicated by the color of each point. The vertical black line denotes equal salience between the singleton and the average of the distractors. Note that the stimuli differ considerably in the relative salience of the singleton, with the singleton being no more salient than the distractors in some cases. The relative salience of the singleton varied slightly depending on its location within the array; thus the colored points represent the average relative salience across all possible positions in the array.

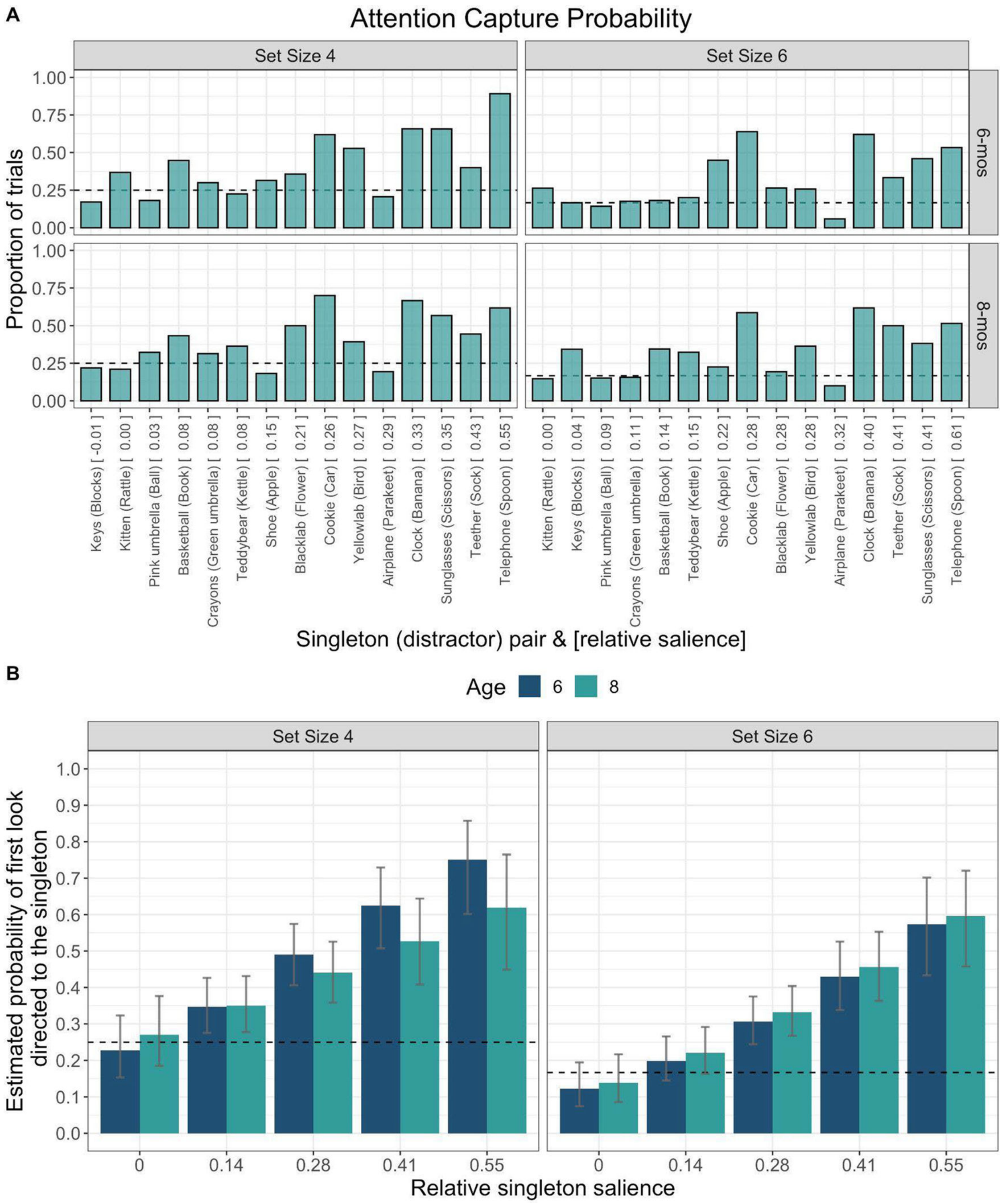

Figure 4.

A: Attention capture probability for each target-distractor pair at each set size (4- and 6-item arrays) and age (6-mos, 8-mos). The height of the bars represent the proportion of total trials in which infants' gaze was first directed to the singleton. The relative singleton salience and the singleton-distractor pair names are denoted along the *x*-axis. The stimuli are arranged in ascending order with respect to relative singleton salience, with salience increasing from left to right. **B:** Statistical modeling results. The height of the bars represent the estimated probability of looking to the singleton first from the GLMM (*y*-axis) for both groups of infants (different colored bars) at specified levels of relative singleton salience (*x*-axis). Probabilities are calculated as odds / (1 + odds). Error bars are back-transformed from the logit scale and represent asymptotic upper and lower 95% confidence limits. The black dashed horizontal lines bisecting both figures denote chance-level responding for set size 4 (.25) and set size 6 (.167).

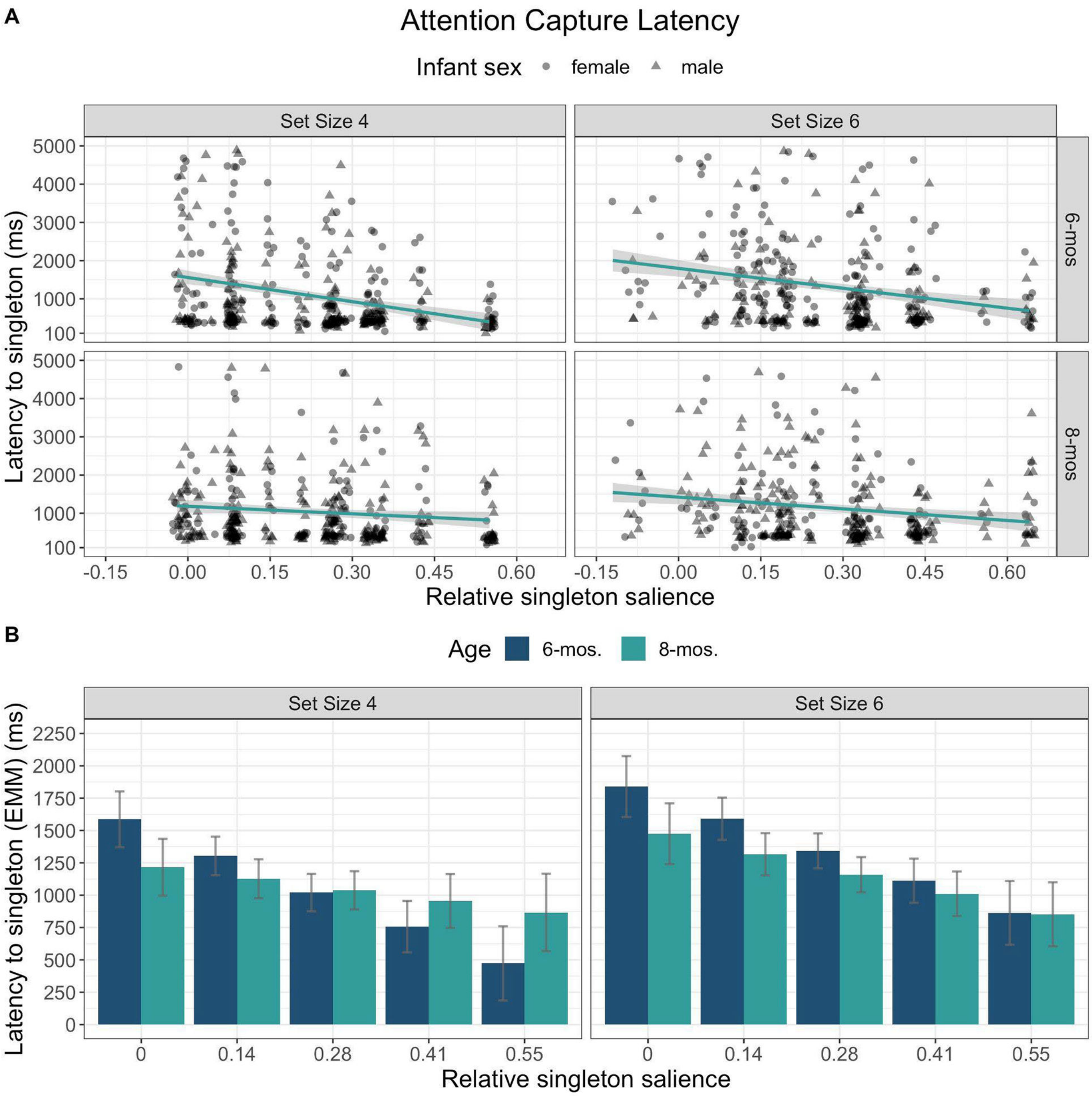

Figure 5.

A: Attention capture latency as a function of the relative salience of the singleton for each set size (4- and 6-item arrays) and age (6-mos, 8-mos). Each point represents a single infant's latency to look to the singleton (*y*-axis) on an individual trial. Although we do not include infant sex in our analyses, as required by our funding source we disaggregate the data by biological sex in the data by indicating female infants with circles and male infants with triangles. The regression lines show a negative relation between infants' latencies and the relative salience of the singleton. Note that the data points (circle and triangle points) were horizontally jittered ±.01 values along the *x*-axis to aid in visualization. **B:** Marginal means from the LMM for the latency to the singleton (*y*-axis) at specified levels of relative singleton salience (*x*-axis). Error bars represent 95% confidence intervals. Intervals are back-transformed from the scale (*M* = 1150, *SD* = 1020).

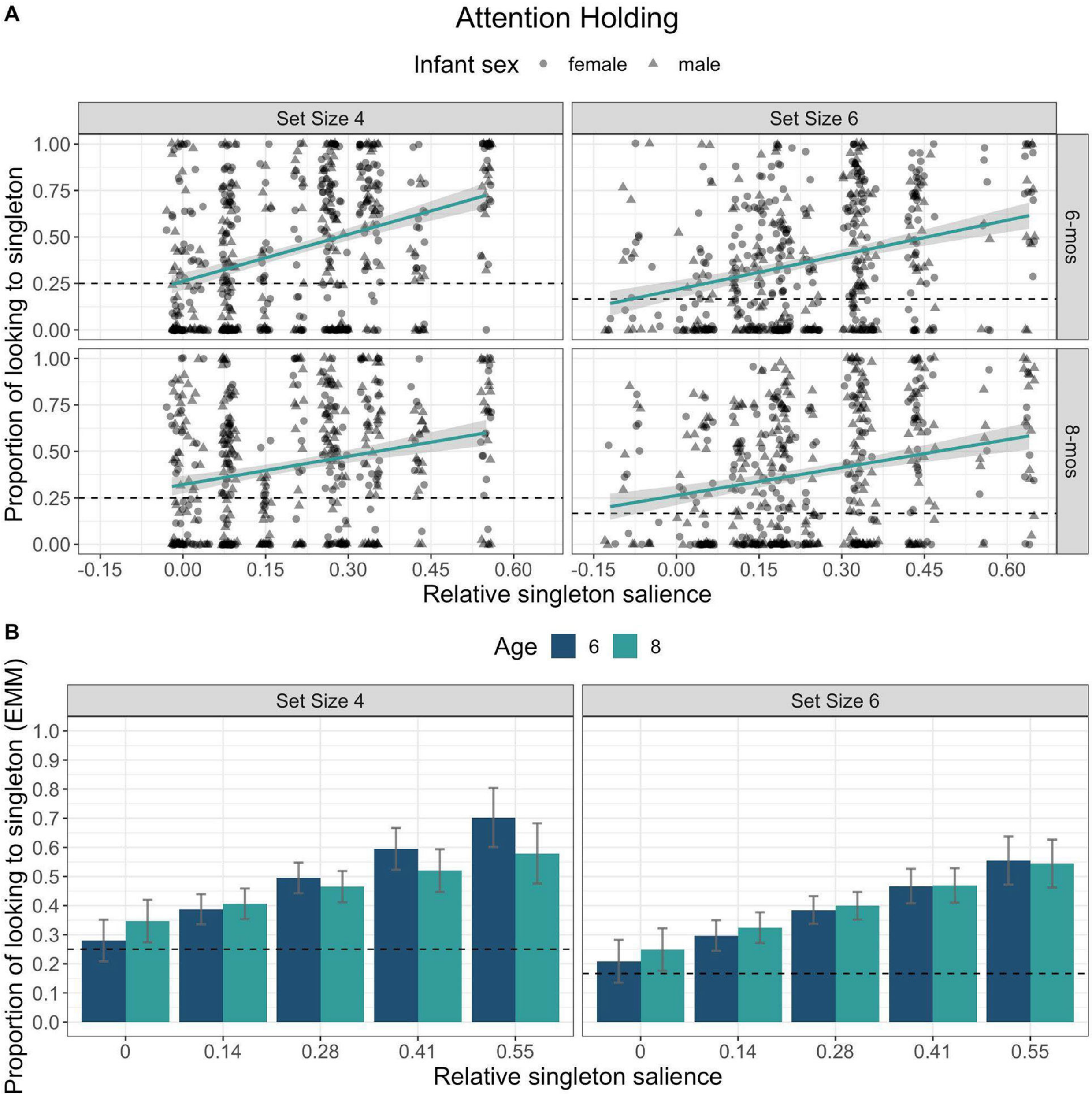

Figure 6.

A: Trial-level data plots of attention-holding for each set size (4- and 6-item arrays) and age (6-mos, 8-mos). Each point represents a single participant's proportion of looking to the singleton (*y*-axis) on an individual trial and, although we did not include sex in our models, we disaggregated our data by biological sex by using different shapes to indicate female and male infants. The relative singleton salience is denoted along the *x*-axis. The regression lines show a positive relation between infants' proportion of looking to the singleton and the relative salience of the singleton. Note that the data points (circle and triangle points) were horizontally jittered ±.01 values along the *x*-axis to aid in visualization. **B:** The height of the bars represent the estimated marginal means from the LMM for the proportion of looking to the singleton (*y*-axis) for both groups of infants (different colored bars) at specified levels of relative singleton salience (*x*-axis). Error bars represent 95% confidence intervals. The black dashed horizontal lines bisecting both figures denote chance-level responding for set size 4 (.25) and set size 6 (.167).

Table 1.

Comparisons of infants' first looks to the singleton to chance averaged across the relative singleton salience values.

Age	**Set size (chance)**	**Mean (95% CI)**	***t*-value**	***p*-value**	**df**	**Cohen's *d***
6-mos	Four (.25)	.43 [.38 – .47]	7.93	<.0001	25	1.56
	Six (.167)	.29 [.24 – .35]	5.14	<.0001	25	1.01
8-mos	Four (.25)	.39 [.33 – .45]	5.04	<.0001	26	.97
	Six (.167)	.32 [.26 – .38]	5.70	<.0001	26	1.10

Note. The trial-level data were averaged to compute a *first-look preference score* by dividing the number of trials in which the infant looked first to the singleton by the total number of trials with either first looks to the singleton or the distractor for each infant (Kwon et al. 2016). These first-look preference scores were then entered in a one-sample *t*-test compared against chance separately for each age and set size combination (all *p*-values are Bonferroni corrected).

Table 2.

Coefficient estimates from the GLMM with first looks to the singleton.

	GLMM			
Predictors	***Odds Ratios***	***SE***	***Statistic***	***p-value***
Intercept	0.29	0.07	−4.94	**<0.001**
Salience	68.90	63.10	4.62	**<0.001**
Age [8-mos]	1.26	0.32	0.92	0.359
Set size [six]	0.47	0.17	−2.07	**0.039**
Salience X Age [8-mos]	0.21	0.19	−1.72	0.086
Salience X Set size [six]	0.89	1.12	−0.09	0.929
Age [8-mos] X Set size [six]	0.92	0.35	−0.22	0.825
Salience X Age [8-mos] X Set size [six]	4.24	5.44	1.13	0.260

Note. The factor levels in the square brackets denote the reference variables in the model. SE = standard error.

Table 3.

Coefficient estimates from the LMM and robust LMM model with latencies to the singleton.

	LMM				**Robust LMM**			
Predictors	***Estimates***	***SE***	***Statistic***	***p-value***	***Estimates***	***SE***	***Statistic***	***p-value***
Intercept	0.43	0.11	4.03	**<0.001**	0.05	0.08	0.64	0.523
Salience	−1.98	0.38	−5.24	**<0.001**	−1.21	0.28	−4.33	**<0.001**
Age [8-mos]	−0.36	0.12	−2.94	**0.003**	−0.16	0.09	−1.75	0.081
Set size [six]	0.25	0.15	1.60	0.109	0.35	0.11	3.04	**0.002**
Salience X Age [8-mos]	1.36	0.42	3.21	**0.001**	0.68	0.32	2.11	**0.035**
Salience X Set size [six]	0.25	0.52	0.47	0.639	−0.24	0.39	−0.62	0.533
Age [8-mos] X Set size [six]	0.01	0.18	0.03	0.972	−0.15	0.14	−1.11	0.267
Salience X Age [8-mos] X Set size [six]	−0.73	0.60	−1.21	0.226	−0.09	0.46	−0.20	0.841

Note. The factor levels in the square brackets denote the reference variables in the model. SE = standard error. Estimates are presented in standardized units.

Table 4.

Comparisons of infants' proportion of looking to the singleton to chance averaged across the relative singleton salience values.

Age	Set size (chance)	Mean [95% CI]	*t*-value	*p*-value	df	Cohen's d
6-mos	Four (.25)	.43 [.40 – .47]	10.19	<.0001	25	2.00
	Six (.167)	.37 [.33 – .42]	9.44	<.0001	25	1.85
8-mos	Four (.25)	.42 [.38 – .46]	9.90	<.0001	26	1.91
	Six (.167)	.39 [.34 – .43]	10.41	<.0001	26	2.00

Note. Infants' trial-level preference scores were averaged to compute a *subject-average* preference score. These subject-average preference scores were then used in a one-sample *t*-test compared against chance separately for each age and set size combination (all *p*-values are Bonferroni corrected).

Table 5.

Coefficient estimates from the linear mixed-effects model (LMM; left) and the robust linear mixed-effects model (right) predicting infants' proportion of looking to the singleton.

	LMM				**Robust LMM**			
Predictors	***Estimates***	***SE***	***Statistic***	***p-value***	***Estimates***	***SE***	***Statistic***	***p-value***
Intercept	0.03	0.04	.82	0.411	0.00	0.04	.04	0.968
Salience	0.77	0.13	5.90	**<0.001**	0.89	0.14	6.21	**<0.001**
Age [8-mos]	0.07	0.04	1.92	0.055	0.08	0.04	2.19	**0.029**
Set size [six]	0.01	0.05	.25	0.801	0.02	0.05	.34	0.734
Salience X Age [8-mos]	−0.35	0.12	−3.00	**0.003**	−0.36	0.13	−2.89	**0.004**
Salience X Set size [six]	−0.14	0.17	−0.80	0.424	−0.20	0.19	−1.05	0.295
Age [8-mos] X Set size [six]	−0.03	0.05	−0.60	0.552	−0.03	0.05	−0.59	0.555
Salience X Age [8-mos] X Set size [six]	0.25	0.16	1.56	0.118	0.24	0.18	1.34	0.182

Note. The factor levels in the square brackets denote the reference variables in the model. SE = standard error.