

**EFFECTS OF BOTTOM-UP VERSUS TOP-DOWN CUEING  
ON CONJUNCTION SEARCH IN  
3-MONTH OLD INFANTS**

MARIA CHRISTINA FUDA

A THESIS SUBMITTED TO  
THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY  
YORK UNIVERSITY  
TORONTO, ONTARIO

October 2012

© Maria Christina Fuda, 2012

**Effects of Bottom-Up Versus Top-Down Cues on  
Conjunction Search in 3-Month-Old Infants**

By

**Maria Christina Fuda**

A thesis submitted to the Faculty of Graduate Studies of York  
University in partial fulfillment of the requirements for the degree  
of

**MASTER OF ARTS**

© 2012

Permission has been granted to: a) YORK UNIVERSITY LIBRARIES to lend or sell copies of this thesis in paper, microform or electronic formats, and b) LIBRARY AND ARCHIVES CANADA to reproduce, lend, distribute, or sell copies of this thesis anywhere in the world in microform, paper or electronic formats *and* to authorize or procure the reproduction, loan, distribution or sale of copies of this thesis anywhere in the world in microform, paper or electronic formats.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

## Abstract

Previous research with infants have suggested that they are fully capable of performing a feature search in a manner nearly identical to adults (Adler & Orprecio, 2006), but are developmentally immature in localizing a target in a conjunction search (Fuda & Adler, 2012). An explanation for the difference in infants performance between feature and conjunction searches was attributed to Wolfe's (1989) Guided Search model of visual search, in which feature searches are thought to rely mainly on bottom-up attentional resources to localize a target, whereas conjunction searches are theorized to require both bottom-up and top-down attentional resources. Because infants have been shown to perform a feature search like that of adults but have been shown not to be able to perform a conjunction search in a similar manner, the current study attempted to show that bottom up attentional mechanisms develop before top-down mechanisms. To this end, 3-month-old infants were presented with two types of cues prior to a conjunction search array that will provide them with prior bottom-up or top-down information that might facilitate their performance in a conjunction search task. The bottom-up cue consisted of four rectangular frames indicating *where* the possible location of the target will be, while the top-down cue consisted of the flashing *what* the target will be in the center of the array. Infant saccadic eye movement latencies were recorded for three different set sizes of conjunction search arrays (5, 8, & 10) when the target was either present or absent. When the target was present, the eye movement latency that localized the target was measured, while when the target was absent the first eye movement latency was measured. Results showed that the top-down cue, but not the bottom-up cue, facilitated the exhibition of a more adult-like conjunction search function in which latencies increased with increasing set sizes. More specifically, the bottom-up cue

resulted in relatively flat search functions for both the target-present and target-absent trials. In contrast, the top-down cue results showed that in the second half of all infant trials, target-present latencies increased with increasing set sizes, while target-absent latencies decreased with increasing set sizes. These results show that infants are developmentally mature in their bottom-up processing, but immature in their top-down processing abilities, and as such the top-down cue provided the facilitation that they needed in order to localize a target in a conjunction search. The current study is the first of its kind to show that 3-month-old infants' top-down processing mechanisms are developmentally immature compared to their bottom-up mechanisms in visual search tasks.

## Acknowledgment

I would like to acknowledge and sincerely thank my research supervisor and mentor, Dr. Scott A. Adler, who through his guidance and encouragement has allowed me to develop as an independent scientific research. I greatly appreciate the opportunity to work in his lab while also pursuing my own research interests. I would also like to thank Dr. Stuart Shanker, my committee member, and the other two members of my committee, Dr. Joseph DeSouza and Dr. Lauren Sergio, for their feedback and comments for this project. I would like to thank my family and friends for all of their support, and lastly, I would like to thank the families and the infants who graciously contributed their time to participate in this study.

## Table of Contents

Abstract	iii
Acknowledgement	iv
Table of Contents	v
List of Figures	vii
List of Tables	viii
1. INTRODUCTION	1
Feature versus Conjunction Search in Infants	6
Why the developmental difference?	12
The Present Study	16
2. EXPERIMENT 1	17
2.1 Eye Movements and Selective Attention	18
2.2 Methods	19
2.2.1. Participants	19
2.2.2. Stimuli & Apparatus	20
2.2.3. Procedure	24
2.2.4. Data Reduction & Analysis	27
2.3 Results and Discussion	29
2.3.1. Preliminary Analysis	29
2.3.2. Main Analysis	31
2.3.3. Number of Eye Movements	34
2.3.4. Split-Half Data Analysis	35
3. EXPERIMENT 2	38

3.1 Methods	39
3.1.1. Participants	39
3.1.2. Stimuli & Apparatus	39
3.1.3. Procedure	39
3.1.4. Data Reduction & Analysis	42
3.2. Results & Discussion	42
3.2.1. Preliminary Analysis	42
3.2.2. Main Analysis	44
3.2.3. Number of Eye Movements	49
3.2.4. Split-Half Analysis	50
4. DISCUSSION	54
4.1. Bottom-up Cueing in Infants	56
4.2. Top-Down Cueing in Infants	60
4.3. Relation to Infants' Visual Expectations	64
4.4. Similarity to Autistic's Visual Search	65
4.3. Conclusion	67
5. REFERENCES	70

## List of Figures

Figure 1. Target location (12, 3, 6 and 9 o'clock), set size (5, 8, and 10), and condition (target present and target absent).	23
Figure 2. Testing crib apparatus	24
Figure 3. Sample trial sequence for the visual displays for Experiment 1.	27
Figure 4. Infants' mean saccade latencies in each set size X array type (target-present and target-absent) condition. Latencies to both the array types remained relatively flat as set size increased.	34
Figure 5. Split half regression analysis for first and second half of trials in both the target-present and target-absent conditions.	38
Figure 6. Sample trial sequence for the visual displays in Experiment 2.	42
Figure 7. Mean saccade latencies for the identity cue for both target-present and target-absent arrays across set sizes. The target absent condition resulted in a significant negative slope.	49
Figure 8. Mean saccade latencies plotted as a function of set size for target-present and target-absent conditions for the first and second half of infants' trials.	54



## List of Tables

Table 1. Preliminary analyses ANOVA results for spatial cue	31
Table 2. Preliminary analysis ANOVA results for identity cue	44
Table 3. Pairwise comparisons for interaction effect between set size and array type.	46

## Effects of Bottom-Up Versus Top-Down Cues on Conjunction Search in 3-Month-Old Infants

Everyday our perceptual world is filled with a variety of simultaneously available objects that are in constant competition for input into our visual and cognitive processes. Our resources for processing these competing objects, however, are limited in that only a finite amount of visual information can be processed at any particular moment. Rather than experiencing overwhelming and continuous streams of visual stimuli, specific items in space are selected as targets in order for visual processing and behaviour to proceed efficiently (Broadbent, 1982; Deutsch & Deutsch, 1963; Treisman, 1964; Yantis & Johnson, 1990). This process of extracting certain information from our environment while ignoring other information is known as selective attention – a process that is critical for our cognition and knowledge construction as this information-processing mechanism serves as the gate for the determination of which available information reaches higher levels of processing. For infants, the selective attention process is particularly critical for their cognitive development because this is when they are first constructing a knowledge base of their world of which they have little or no experience (Adler, 2005). So, the nature of the developmental state of infants' selective attention will play a large role in determining what information is processed and, hence, the nature of their cognitive development and construction of their knowledge base.

Typically selective attention, regardless of developmental maturation, is thought to consist of two distinct but complimentary processes. Neisser (1966) was the first to propose this dichotomous relationship for the attentional mechanisms engaged in processing visual information. He termed these two processes as pre-attentive and attentive (Neisser, 1966). The pre-attentive mechanism was proposed to occur before attention and process all visual stimuli in

parallel, whereas the attentive mechanism required active allocation of attention, serially scanning each item in the visual field (Neisser, 1966). Since Neisser's (1966) original dichotomous model of selective attention, others have developed and expanded upon it using various types of visual search paradigms. These paradigms generally require a participant to localize a unique target amidst various types and number (set sizes) of distractors. For instance, Treisman and Gelade's (1980) Feature-Integration theory (FIT) expanded upon Neisser's (1960) attentional dichotomy by suggesting that features such as color, orientation, spatial frequency, brightness, and motion are registered pre-attentively and thus are detected automatically and in parallel. These features exhibit an automatic engagement of attention when a target consists of one unique level on one of these perceptual features that is absent from its surrounding distractors; in other words, these aforementioned features are said to "pop-out" from surrounding distractors. For example, a red target popping out from green distractors. Feature-integration theory, therefore, suggests that in this type of feature or parallel search, identification of the target occurs prior to its localization (Treisman & Gelade, 1980). This is in contrast to an attentive/serial or conjunction search, which occurs at a later stage, where focused attention is required to serially search for objects in a complex array. In this type of search the target contains a conjunction of features that are shared with some, but not all of the surrounding distractors (Treisman & Gelade, 1980). For example, a red X surrounded by red "O"'s and green "X"'s. In this type of serial search, localization of the target is required for identification.

In order to test their FIT model, a visual search paradigm was used to represent and control for targets and competing stimuli in the environment. Typically in visual search paradigms, there are two commonly used measurements: percent correct, and reaction time (RT). In a feature search, because the target is perceptually unique it pops out from its distractors

irrespective of the number of distractors (set size) surrounding it, and thus results in relatively flat reaction time slopes. For example, in two of the experiments in Treisman and Gelade's (1980) original study, they showed that the color pink popped out from the colors brown and purple, and that an "O" popped out from distractors "N" and "T", in set sizes ranging from 1 to 30. Pop-out is exhibited by flat RT regression slopes equaling or resembling zero, indicating that as the set size increases, the time to detect the target does not increase. Wolfe (1994) has stipulated that in order for a search to be parallel and pop out to be exhibited, the search function should be less than 20ms/item, indicating a relatively flat search function.

Conversely, a conjunction visual search occurs when the target is no longer perceptually distinct from its distractors by the presence of a unique feature. In this type of visual search, the target shares one or more features with one set of distractors and one or more features with other sets of distractors. For example, Treisman and Gelade (1980) showed that a pink "O" did not pop out from amongst varying numbers of green "O" and pink "N" distractors. Instead, participants needed to serially allocate attention to each item in the array until they came across the target. Results of this experiment indicated that as the set size increased, so did the reaction time to detect the target. This finding revealed that unlike feature searches, focused attention is required for conjunction search tasks and that identification requires localization of the target (Treisman & Gelade, 1980). Wolfe (1994) suggests that in order for a search to be considered a serial search, then regression slopes must be greater than 20ms/item, indicating an increasing search function.

An important component of visual search studies in determining the functioning of the pre-attentive and attentive mechanisms is the comparison of search when the target is present versus when the target is absent. In a feature search when the target is present it is detected

automatically, thus using few attentional resources (Treisman & Gelade, 1980; Wolfe, Cave, & Franzel, 1989). However, if the target is absent then participants need to scan the display using more attentional resources and thus increasing RT to identify the absence of a target. This is because of what Sternberg (1969) refers to as a self-terminating search, in that attentional resources are used only until the target has been found or all of the items have been scanned. The magnitude of the RT slopes when the target is absent depends on the difficulty of the visual search task. When the task is relatively simple (i.e. small set size with large perceptual discrepancy between target and distractors) then both target present and target absent RT's produce a relatively flat function. As the task becomes more difficult, the RT slopes increase as a function of set size. Generally, target absent trials produce RT slopes that are about twice as steep as target present trials (Wolfe, 1994).

There have been some issues, however, with the FIT model in that some studies have determined that conjunction searches can occur in parallel, whereas certain feature searches can occur serially (Nakayama and Silverman, 1986). For example, Nakayama and Silverman (1986) used stereoscopic disparity or two separate spatial dimensional planes to identify a conjunction search for a target based on both color and stereo disparity (CS), or motion and stereo disparity (MS). For the CS condition, the distractors in the front plane were always blue arrows and distractors in the back plane were always red arrows. The target would thus be either a red arrow in the front plane or a blue arrow in the back plane. For the MS condition the distractors in the front plane were always moving up and the back plane distractors were always moving down. The target in this condition was thus either an arrow moving down in the front plane or up in the back plane. The results of both of these experiments indicated that these target searches were both conducted in a parallel manner, in that RTs were constant across 15-40 set sizes of

distractors. Such findings suggest that the visual system can restrict its attention to a single spatial dimension and not be distracted by distractors in an opposing depth (Nakayama & Silverman, 1986).

In an attempt to address some of the issues with FIT, Wolfe and colleagues (1989, 1994) proposed the Guided Search model in which attention is guided for parallel and serial searches by simultaneously using both bottom-up and top-down processes and then summing together the amount of activation. Only the item receiving the highest activation will be processed further. These dual processes are represented by activation maps, which are critical for the functioning of visual search and attention. More specifically, the goal of Guided Search is that early parallel processes extract limited information from the visual input, but that limited information then guides subsequent processes to find more complex targets (Wolfe, 1994). A parallel search, where all the items in the visual field are processed simultaneously, is considered efficient because less attentional resources are required to attend to a target amongst surrounding items. On the other hand, a serial search, where each item is processed individually, is considered inefficient because more attentional resources are required in order to attend to a target amongst surrounding items. Consequently, a parallel or efficient search occurs when the target exhibits high bottom-up, stimulus-driven activation relative to the distractors in a single feature map (e.g. feature search) or combined across multiple feature maps (e.g. efficient conjunction search), whereas a serial or inefficient search occurs when the level of activation is uniform across the visual display and requires additional top-down activation. This occurs because each feature creates a spatial activation map of all the features and distractors in the visual search array, then each activation map is scanned serially. However, some features have a strong combined activation of features across activation maps or even in a single map, which consequently trigger

the perceptual system's localization of a target (Wolfe, 1994). High activation of a single feature across an activation map, therefore, is located efficiently, whereas lower activation of a target sharing multiple features across multiple activation maps is located inefficiently (Wolfe et al., 1989). From this model, it is suggestive that early pre-attentive search allows for detection of a target, but that later attention is required to identify the identity of the target (Wolfe et al., 1989). Thus in the case of Nakayma and Silverman's (1986) study, each depth plane would have resulted in its own activation map, with high activation of the target in that depth plane's map, resulting in an efficient search process.

#### *Feature Versus Conjunction Search in Infants*

Other than relatively few disparities, research has collectively shown that the pop-out effect in adults appears to be relatively robust across studies. In conjunction with the fact that pop out is thought to be due to a primitive, pre-attentive system that analyzes visual information in a bottom-up fashion, a reasonable supposition would be that pop-out effects should also be exhibited in the early stages of life (Columbo, Ryther, Frick & Gifford, 1995). The first identifiable study conducted on whether or not infants' exhibit pop-out was reported by Salapatek (1975). Salapatek (1975) found that 3-month-old infants always oriented to a unique patch of squares embedded in an array of horizontal lines (or vice versa), whereas 2-month-olds did not; suggesting that the mechanism responsible for pop-out might not develop until around 3 months of age. Since this original study there have been a number of related studies attempting to assess pop-out in infancy, most of them using the preferential-looking (Sireteanu & Rieth, 1992) or the habituation/dishabituation paradigms (Atkinson & Braddick, 1992). For instance, Atkinson and Braddick, similar to Salapatek (1975), found that after 4-month-olds are habituated to one pattern of lines oriented at a 45° angle that were surrounded by lines oriented 135° they

are able to notice differences between the alternatively oriented lines; a task that 2-month-olds were unable to do. Consistent with this finding, Sireteanu and Rieth (1992) showed that infants as young as 2 months preferentially oriented towards a discrepant patch when it was defined by size, suggesting that it popped out. However, when the discrepant patch was defined by orientation, preference was not detected until approximately 12 months.

Alternative methods, in addition to the methods stated above, have been used to assess the capacity to detect non-textural stimuli that pop out in infancy. Rovee-Collier, Hankins, and Bhatt (1992) used the mobile-conjugate reinforcement paradigm to show that “+”s pop out when surrounded by distractor “L”s presumably because the “+” contains a unique perceptual feature of the line crossing (Julesz, 1981). In this study, Rovee-Collier et al. (1992) trained 3-month-olds to kick to move an overhead, seven-block crib mobile that displayed either all “L”s or “+”s on every block side and then tested them with a mobile that consisted of either a single unique “L” block among six “+”s or vice versa. They found that when infants were trained with all “+”s they exhibited recognition when tested with one “+” among six “L”s, and vice versa when trained with “L”s. Conversely, when infant’s were trained with all “+”s and were tested with one “L” block among six “+” blocks, they exhibited discrimination, and vice versa when trained with “L”s. These results show that the familiarity or novelty of the unique characters on the single block irrespective of the familiarity or novelty on the surrounding blocks determined infants’ recognition or discrimination performance. That is, overall, this finding that the unique block controlled infants’ recognition performance suggests that the unique character block popped out from amidst the surrounding dissimilar character blocks, which closely resembles adult studies using similar stimuli (Julesz, 1981).



A couple of issues arise, however, with these aforementioned paradigms that limit their capacity to be informative about the development of the attention mechanisms that form the foundation for the visual search behavior. For instance, the discussed infant paradigms generally use measurements on the order of seconds or even minutes, which are on scales much larger than that measured in the visual search paradigm used in adult research. Recall that visual search paradigms used with adults typically measure RT, which is based on requiring the participant to have enough motor control to push a button and make a response. Infants, however, do not possess the motor control to press a button indicating target detection, creating an issue when comparing infants' visual attention directly with that of adults'. Therefore none of the previously mentioned infant paradigms provide measures (e.g. RT) comparable to those obtained with adults because they measure visual search over larger time scales thereby making comparison of the underlying attentional and perceptual mechanisms difficult (Adler, 2005).

To overcome the limitations of the previous infant paradigms in assessing the development of visual search mechanisms, Adler and Orprecio (2006) designed a means for studying infant pop-out by measuring their saccadic eye movement latencies – a measure that can be directly compared with adults. Using the same “+”s and “L”s used in Julesz’s (1981) and Rovee-Collier et al.’s (1992) studies, Adler and Orprecio (2006) showed infants and adults either a single “+”, or a “+” surrounded by 2, 4, or 7 “L”s for 1000 msec. Every infant and adult was presented with interleaved trials of target-present or target-absent arrays. They found that for infants when the target was present, saccadic latencies did not increase as a function of set size with a slope of 5.2 ms/item (Adler & Orprecio, 2006). In contrast, when the target was absent, latencies increased as a function of set size with a slope of 23.9 ms/item (Adler & Orprecio, 2006). The adults responded similarly to the infants in that latencies to target-present arrays did

not increase as a function of set size with a slope of 7.2 ms/item, compared to target-absent arrays with a slope of 27.5 ms/item. This study represents the first definitive study to show that 3-month-olds exhibit parallel attentive search mechanisms like that of adults.

In contrast to feature search and pop-out, there have been relatively few studies investigating conjunction search in infancy. For adults, color-shape conjunctions are easier to discriminate than shape-shape conjunctions (Wolfe et al., 1989). Rovee-Collier et al. (1992) suggested that although color-shape conjunctions are easier for infants to distinguish than shape-shape conjunctions, they are both still not easily distinguishable for infants. Bhatt, Bertin and Gilbert (1999) examined 3-month-olds' and 5.5-month-olds' ability to detect color-shape conjunctions using a novelty-preference paradigm. Infants were first exposed to arrays of red Xs and blue Os and later shown arrays with a discrepant patch of red O's embedded in the red Xs and blue Os (Bhatt et al., 1999). They found that 3- and 5.5-month-old infants were unable to detect the color-shape conjunctive discrepancies but that both age groups were able to detect featural discrepancies (Bhatt et al., 1999). Three additional experiments by Bhatt and colleagues also revealed that 5-month-old infants are able to detect a feature target but are unable to detect conjunction targets. These findings suggest that the different mechanisms responsible for feature and conjunction searches proposed by Treisman and Gelade (1980) proposal may exhibit different developmental paths.

As with earlier feature studies with infants, the conjunction studies by Bhatt and colleagues (1999) suffer from the same problems for evaluating the development of the underlying visual search mechanisms in that they do not take set size into account and typically provide measurements in seconds rather than milliseconds. In order to measure RT and account for set size, Gerhardstein and Rovee-Collier (2002) used older infants and toddlers. Here they

showed 3 different age groups of 12, 24, and 36-month-olds different set sizes of red and green dinosaurs. The toddlers were trained to search for the dinosaur that was unique in both color and shape from distracting dinosaurs (i.e., a red “Barnette” dinosaur surrounded by red “Dino” dinosaurs and green “Barnette dinosaurs) and then respond by touching them on an LED touch-sensitive monitor (Gerhardstein & Rovee-Collier, 2002). Upon completion of training, they were shown both feature and conjunction arrays with set sizes of 2, 4, 8, and 12. Results showed that for the feature search task, RT was significantly affected by age, in that the younger children were subsequently slower than the older children, but RT was unaffected by set size. For the conjunction search task, there was insufficient data for the 12-month-olds’ conjunction search, but for 24 and 36-month-olds, RT was also significantly affected by age, where again the younger children were subsequently slower. In contrast to feature search, however, RT was affected by set size, in that the time to touch the target increased as a function of set size (Gerhardstein & Rovee-Collier, 2002). This study suggests that feature and conjunction search in older infants and young children shows the same functions as adults (Gerhardstein & Rovee-Collier, 2002). The question then remains as to why infants younger than 24 months, such as the 5.5-month-olds in the Bhatt et al. (1999) study, do not exhibit evidence of adult-pattern conjunction search. Is it due to the methodological considerations stated above or is there a more fundamental developmental issue?

Though the Gerhardstein and Rovee-Collier (2002) study with older infants correctly accounted for set size in their conjunction search task and used RT measurements that are more comparable to adult research, they were unable to address infants younger than 24 month’s ability to perform or not perform the conjunction search task. To this end, using the same methods as Adler and Orprecio (2006), Fuda and Adler (2012) compared feature and conjunction

searches in both 3-month-old infants and adults by assessing their eye movement latencies to localize a target amidst distractors. For the feature search, both infants and adults were required to locate one of four targets (a green or red “X” or “O”) amongst distractors that differed in either color or shape, for example a red “X” amongst green “X”s, or a red “X” amongst red “O”s. Both adults and infants were required to localize the target amongst varying set sizes of 5, 8 and 10, where the target was either present or absent. Results for the feature search were similar to Adler and Orprecio’s (2006) feature search study, in that infants’ and adults’ saccadic latencies did not increase as a function of set size, and that their functions were not significantly different between the two age groups (Fuda & Adler, 2012). The only difference that did occur between the two age groups was that infants were approximately 100 msec slower than the adults to both the target-present and target-absent arrays. This recurring trend is likely related to infant ocularmotor functions being less developed rather than due differences in their underlying attentional mechanisms (Adler & Orprecio, 2006).

For the conjunction search, both adults and infants were required to locate one of the same four targets used in the feature search task; however, half of the distractors shared one feature with the target (color or shape) while the other half of the distractors shared the other corresponding feature with the target (e.g. a red “X” surrounded by red “O”s and green “X”s) (Fuda & Adler, 2012). Both adults and infants were required to localize the target amongst the set sizes of 5, 8 and 10 with interleaved target-absent and target-present arrays. Results of the conjunction search revealed that adults’ saccadic latencies did significantly increase as a function of set size when the target was present with a slope of 64.26 msec per item, whereas the infants’ saccadic latencies did not significantly increase. Also, recall that for the feature search, infants were consistently 100 msec slower than adults, but in the conjunction search infants were closer

to 200 msec slower (Fuda & Adler, 2012). These data suggest that there is a developmental difference in being able to detect a conjunction target versus a feature target, which is consistent with previous conjunction search findings with infants (Bhatt et al., 1999; Gerhardstein & Rovee-Collier, 2002). Infants are likely hitting an ability ceiling with their available attentional capacities, in that the task is too difficult for them to perform any faster on the smaller set sizes than the larger ones. Regardless of set size, infants are performing at the maximum capacity that is allowed by the current developmental level of their attentional mechanisms. In particular, these findings suggest that the attentional mechanism responsible for the developmental difference between infants and adults in conjunction search performance is the serial processing mechanism, with serial processing developing at a much later stage than parallel processing.

*Why the developmental difference?*

In both the Adler and Orprecio (2006) and the Fuda and Adler (2012) studies, both infants and adults performed identically on a feature search. On a conjunction search, however, they performed very differently. One possible explanation for the developmental difference might be related to Wolfe et al.'s (1989) Guided Search model of visual search. Recall that in Guided Search parallel feature activation maps feed into the serial attentive map, so that each feature map is scanned serially and the location that has strongest combined activation across feature maps is selected. When a location has a very large activation relative to other locations either in a single feature map (as occurs in a feature search) or combined across feature maps, then search is efficient and the item at that location proceeds to pop-out. When there is no large activation difference at any one location in a single feature map or combined across feature maps (as typically occurs in a conjunction search), then inefficient, serial mechanisms are engaged.

While Triesman and Gelade (1980) were correct in showing that there are two separate processes, Guided Search expands upon this by showing that early parallel processes guide the serial processes through an efficient/inefficient search component (Wolfe et al., 1989). More importantly for the purpose of this study is that Wolfe et al.'s (1989) model is based on the concepts of bottom up stimulus-driven or exogenous activation combined with top-down cognitively driven or endogenous activation. Basically, searches rely on basic bottom-up information that feed into more complex searches, which also require a top-down component. That is, most searches, according to Wolfe et al. (1989), are accomplished through a combination of bottom-up stimulus feature activation and top-down down attentional allocation. When the bottom-up feature activation is sufficiently large, then little to no top-down attention is required to distinguish the target from surrounding distractors. As a consequence, the item pops out and is found irrespective of the number of distractors. In contrast, when bottom-up feature activation is not sufficiently large perhaps due to sharing of features across a number of visually available items as in conjunction search, more top-down attention is required, an inefficient search ensues, and an increasing function with the number of distractors is produced. Infants are capable of performing a feature search, likely because the bottom-up feature activation maps are fairly mature early in development due to their lower level and more primitive nature. Consequently, infants' apparent inability to similarly perform a conjunction search like adults is likely due to their top-down attentional mechanisms not being fully mature.

In order to assess whether infants' top-down attentional mechanisms are not fully matured, we will turn to some adult literature which has focused on making an inefficient conjunction search an efficient one by presenting cues before the onset of a visual search array that attempt to aid the observer in finding the target more efficiently. A brief bottom-up

peripheral spatial cue provides participants with information as to where the location of the target will be (Chastain, Cheal, & Lyon, 1996; Posner, 1980; Sobel et al., 2009). This type of cue is considered a stimulus-driven or bottom-up cue because the observer is not provided with any information that enables them to form an internal knowledge structure about what or where the target is that they can use to voluntarily allocate attention. Instead, the spatial cue automatically primes attentional allocation to that location merely based on the external information that the cue provides (Posner, 1980). Conversely, a top-down cue provides observers with specific information about what (or where) the target will be (Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004). This type of cue is considered a user-driven or top-down cue because the observer is required to internalize the cue's information that can then be used to voluntarily allocate attention in a search plan that utilizes specific characteristics about the target, such as color or shape or location, which have been provided by the cue (Richards, 2000).

Posner (1980) has shown that when a target is preceded by a top-down cue such as an arrow pointing to the possible target location, if the target is in the indicated location (a valid trial), then observers will localize the target much faster than if the same target has not been cued or the target has been preceded by a cue indicating a different location (an invalid trial). In terms of visual search, Wolfe et al. (2004) examined the effect of two types of user-driven, top-down cues, an actual picture of the target, or words explaining what the target will be, on search functions. For instance, the observers were shown in the picture cue condition a picture of a red vertical line prior to a search array that consisted of a red vertical line target surrounded by green vertical and red horizontal line distractors (Wolfe et al., 2004). The word cue conditions contained the same search array, but prior to the array observers were shown the words "red vertical" (Wolfe et al., 2004). Results of this study showed that in a standard conjunction search,

observers were able to localize a target in a search array much faster when it was preceded by a picture cue of the target, rather than a word cue, or no cue. This is because picture cues provide the observer with a picture of the actual target, requiring less top-down information processing than the word cues (Wolfe et al., 2004). Similarly, a study by Sobel, Pickard and Acklin (2009) took a very interesting approach and combined peripheral spatial cues with identity cues indicating the target's features such as color or orientation prior to showing observers a conjunction search array. Here they used two types of cues, with the first being a circular array indicating only the orientation (not color) of all of the arrows that would be shown in the following conjunction search array (i.e. vertical or horizontal). The second cue was a circular array indicating only the color (not orientation) of all of the arrows that would be shown in the following conjunction search array (i.e. red or green). Finally, a control group of observers were not given any cue prior to the search array. Results indicated that the color cue speeded search performance compared to receiving no cue, and that the orientation cue did not affect performance compared to the control condition (Sobel, Pickard & Acklin, 2009). These results suggest that when observers are given a preview of the color of the arrows at each location in the search array, observers are able to get a "head start on pre-attentive processing", which does not seem to occur for orientation cues (Sobel, Pickard & Acklin, 2009, p. 25). Sobel et al. (2009) have suggested that the difference occurs because both top-down and bottom-up processing work together for color spatial cues, but not for orientation spatial cues (Sobel et al., 2009). This study has also shown that bottom-up cueing can affect visual search performance like that of top-down cues.



## The Present Study

There has been evidence demonstrating that attentional mechanisms underlying visual search are present in early infancy (Adler et al., 1998; Gerhardstein & Rovee-Collier, 2002; Rovee-Collier et al., 1992), but there are limitations to this evidence because of methods and paradigms that do not allow for a direct comparison between infants and adults of the attention mechanisms previously reported in the adult literature on conjunction searches. Adler and Orprecio's (2006) study demonstrated that infants exhibit pop-out and feature search similarly to adults, suggesting that the responsible pre-attentive mechanisms are already functioning in early infancy. Fuda and Adler's (2012) study has indicated a different story, namely, that infants are capable of localizing a target in a conjunction search, but not in the same manner as adults. In our study, the 3-month-old infants seemed to have exhausted their own abilities, localizing the target as efficiently as they could, when conducting a conjunction search as indicated by high overall saccadic latencies across all set sizes, yet no increasing function as the set size increased. A likely cause for this impaired performance by infants might be differential development of bottom-up versus top-down mechanisms, with top-down attentional mechanisms showing a prolonged developmental timeline.

The present study, therefore, aims to provide 3-month-old infants with two types of cues that will provide them with prior bottom-up or top-down information that might facilitate their performance in a conjunction search task. Using the paradigm designed by Adler and Orprecio (2006) which allows direct comparison with adults and their visual attention mechanisms, infants were shown the same stimuli used in Fuda and Adler's (2012) study, but were additionally given either a stimulus-driven bottom-up cue where infants were flashed frames indicating *where* the possible location of the target will be, or a user-driven top-down cue where infants were shown

*what* the target will be prior to the search arrays. If, as hypothesized, infants can perform a feature search but not a conjunction search, relative to adults, because they are developmentally immature in their top-down attentional processes, then the top-down target cue condition will provide them with more support than the bottom-up spatial cue condition needed to perform a conjunction search like that of adults. As a result, the top-down cue, but not the bottom-up cue, should facilitate the exhibition of a more adult-like conjunction search function in which latencies increase with increasing set sizes.

### **Experiment 1 – Stimulus-Driven Spatial Cues**

In the first experiment, we assessed the effect of stimulus-driven spatial frame cues in a conjunction search by measuring 3-month-old infants' saccadic eye movement latencies. As previously discussed, visual search research with adults has demonstrated that the efficiency of a search is determined by the relative use of attentional resources (Wolfe, 1994; Treisman & Gelade, 1980). An efficient feature search occurs when minimal attentional resources are required to perform a search for a target among distractors. During this search, the time and accuracy required to detect the target is relatively unaffected by the number of distractors on the visual display (Wolfe et al., 1989). In contrast, an inefficient (conjunction) search requires more attentional resources, thus the accuracy and speed of detecting the target among distractors increases as a function of set size. However, when provided with information about the location of the target with spatial cues, target efficiency does increase in adults, though not to the level of an efficient search (Posner, 1980).

Also, in regards to Fuda and Adler (2012), infants' latencies for a conjunction search were relatively flat, but overall significantly slower than adults. These results could be indicating a floor effect in that infants' could not perform any faster irrespective of set size.

Consequently, there are two main hypotheses in Experiment 1. First, we hypothesize that when provided with possible spatial locations for the target, infants will not exhibit an inefficient search like that of adults, however, the bottom-up cue should facilitate performance by lowering the latencies at each set size. These attentional resources would be confirmed if saccade latencies to find the conjunction target remain flat with increasing set sizes but are lower than Fuda and Adler's (2012) results that provided no cue. Secondly, we also hypothesize an inefficient search where the target is absent. This would be confirmed if the saccadic latencies to the stimuli increase linearly with increasing set sizes.

#### *Eye Movements and Selective Attention*

Why use eye movements and the latency to initiate an eye movement as a measure of underlying attentional mechanisms? Previous research has indicated that there is a tight linkage between eye movements and attentional allocation (Adler, Bala & Krauzlis, 2002; Hoffman & Subramanian, 1995; Posner, 1980; Zelinsky, 1997). For instance, Hoffman and Subramanian (1995) found that target detection prior to eye movement initiation is superior when target and saccade location coincide than when they do not, suggesting that the allocation of spatial attention is an important element of generating a saccadic eye movement. Similarly, Kowler et al. (1995) found that attention could not be allocated to one location while at the same time preparing to make a saccadic eye movement to a different location. In particular, there have been a number of studies that have measure the relation between adult's eye movements and attentional processing in visual search tasks. Zelinsky (1997) for example compared adults oculomotor RT's to manual RT's in both parallel (Q among O's) and serial (O among Q's) visual search tasks. The main purpose of this experiment was to determine if both dependent variables (manual and oculomotor) were highly correlated during a visual search task. Results showed

that oculomotor RT was highly correlated with the traditional manual RT motor behavior in both parallel and serial search tasks (Zelinsky, 1997). Together, these findings (Adler, Bala & Krauslis, 2002; Hoffman & Subramaniam, 1995; Kowler et al., 1995; Zelinsky & Sheinberg, 1997) provide significant evidence that eye movements probably assess attentional mechanisms used in visual search. As a result, the dependent variable used to assess attentional processing in 3-month-old infants and adults were the saccadic latencies exhibited in localizing a target among distractors (see Adler & Orprecio, 2006).

## Methods

### *Participants*

Infants were recruited every month from a purchased mailing list (Z Retail Marketing Inc., Toronto, Canada) that included names and addresses of 500 families with 2-month-old infants. Once names were obtained, parents were sent a letter inviting them to have their infants participate in the study. Interested participant then either returned postage paid post-card to the lab, telephoned, or e-mailed suggesting their willingness to participate. These parents were then contacted by phone or e-mail to set up an appointment. The data from 8 infants (4 males, 4 females) who ranged in age from 96 to 112 days ( $M = 104.1$ ,  $SD = 5.9$ ) were included in the study. Infants were Caucasian ( $n = 5$ ), Asian ( $n = 1$ ), and otherwise non-specified ( $n = 2$ ) and were primarily from middle SES backgrounds. Infants were all born at full term, in good health with no apparent visual, neurological or other abnormalities. The data from an additional 4 infants were excluded from the study because of equipment or software failure during testing ( $n = 1$ ) and infant fussiness or inattentiveness ( $n = 3$ ). All parents were required to fill out a consent form and a brief demographic questionnaire before testing begins. Parents were also given information regarding the experiment before commencement of the experiment and debriefed

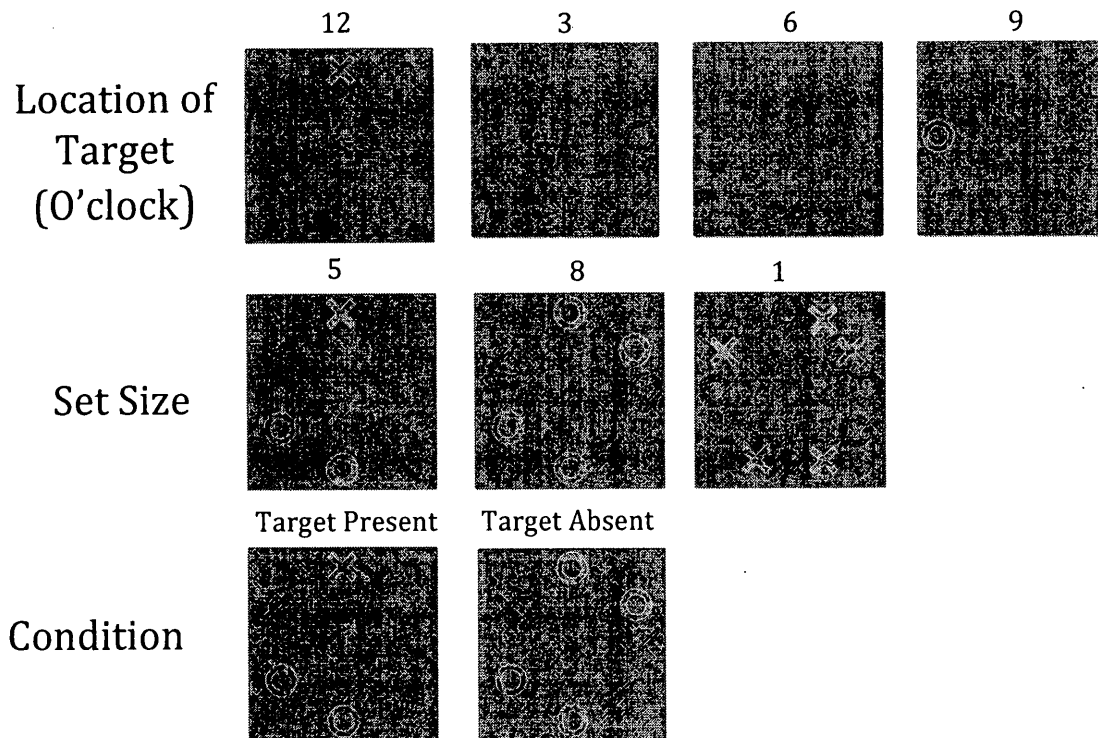
following its completion. In exchange for participation, participants were given a diploma and parent a photo shoot offer from Sears worth \$19.99.

### *Stimuli and apparatus*

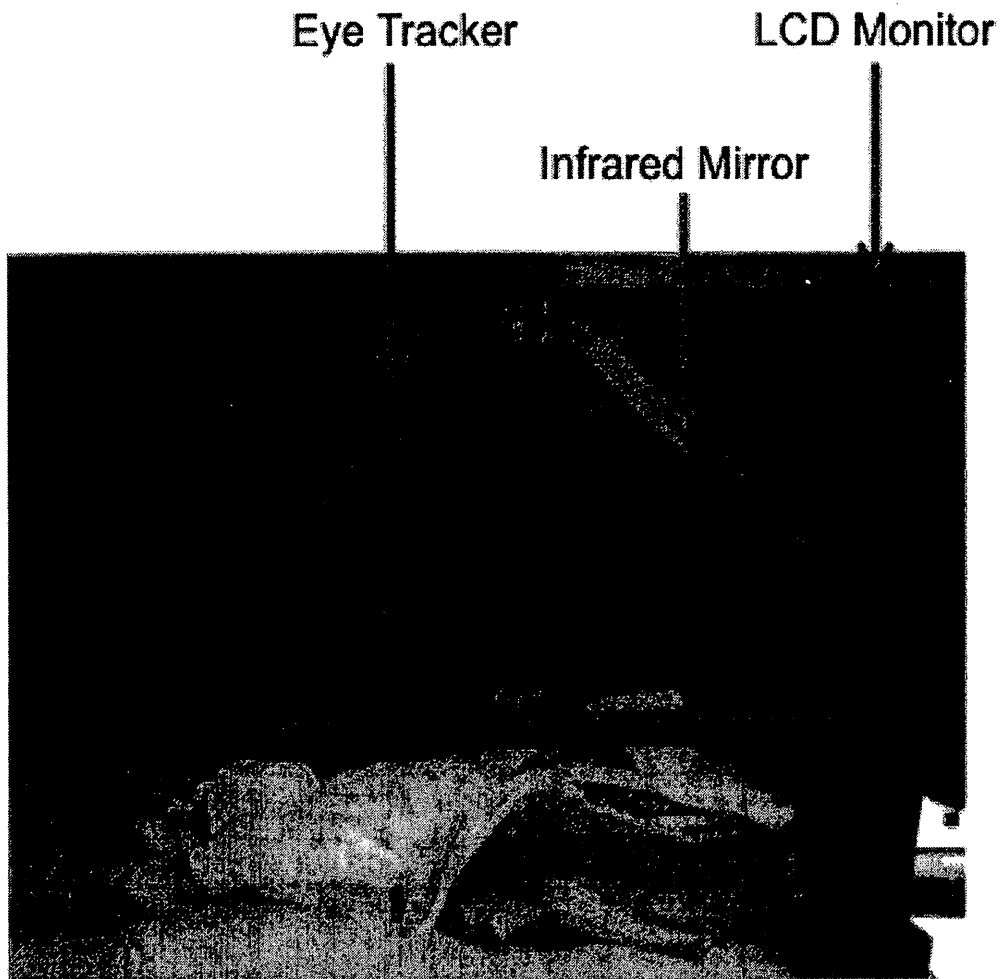
The stimuli were computer-generated graphic images of a pink and yellow fixation triangle, black rectangular frames, and red or green X's or O's. The stimuli were presented on a 19-inch IBM color monitor with 1024 X 768 pixel resolution, a refresh rate of 75 Hz and an 8 bit/pixel gray scale. The infants viewed the images from a distance of 48 cm. The stimuli were arranged on a circular grid with a 7° radius with the rectangle frames, and X's and O's having a subtended visual angle of 4° and the triangle having a subtended visual angle of 1°. There were 4 conjunction search array conditions: a green X amongst green O's and red X's; a red X amongst red O's and green X's; a green O amongst green X's and red O's; and a red X amongst red O's and green X's (see Figure 1). Among the four types of target conditions there were two types of visual search arrays, a target-present array and a target-absent array (see Figure 1). Furthermore, the number of distractors in the two types of arrays varied in that they consisted of set sizes containing 5, 8, or 10 stimuli (see Figure 1). Target-present arrays consisted of a single green or red X or O randomly presented at one of four possible locations (12, 3, 6, or 9 o'clock positions) on a circular pattern amidst 4, 7, or 9 distractors (see Figure 1). Target-absent arrays consisted of 5, 8, or 10 distractors that correspond with the target-present distractors randomly presented at any of the 12 positions on the circular pattern.

Infants were laid in a supine position in a specialized crib in the lab where they viewed images 48 cm below a monitor placed directly overhead. A visible-transmitting infrared mirror was located between the infant and the monitor in order to reflect eye images to the eye-tracking camera (Model 504, Applied Science Laboratories [[www.a-s-l.com](http://www.a-s-l.com)], Bedford, Mass.) also

located directly above the infant (See Figure 2). The eye tracker recorded the participant's eye movements at a temporal resolution of 60 Hz. Infrared light emitted from diodes on the camera were reflected back from the participant's retina through the pupil producing a backlit white pupil while also producing a point reflection from the corneal surface of the eye. Overhead lights in the testing room were dimmed and black felt curtains were hung around the immediate testing area to reduce distraction. Calibration was conducted in order to equate recorded eye tracker values of eye location to known locations on the stimulus screen. The calibration stimulus was a stationary blue square with yellow, black and green squares layered within it. The calibration stimulus was first presented at the top left corner (point 1) of the screen followed by the bottom right corner (point 9) of the screen. All subsequent eye data was filtered through these calibration values.



*Figure 1.* Target location (12, 3, 6 and 9 o'clock), set size (5, 8, and 10), and condition (target present and target absent). The darker shade of gray represents the color red and the lighter shade of gray represents the color green.



*Figure 2.* Testing crib apparatus.

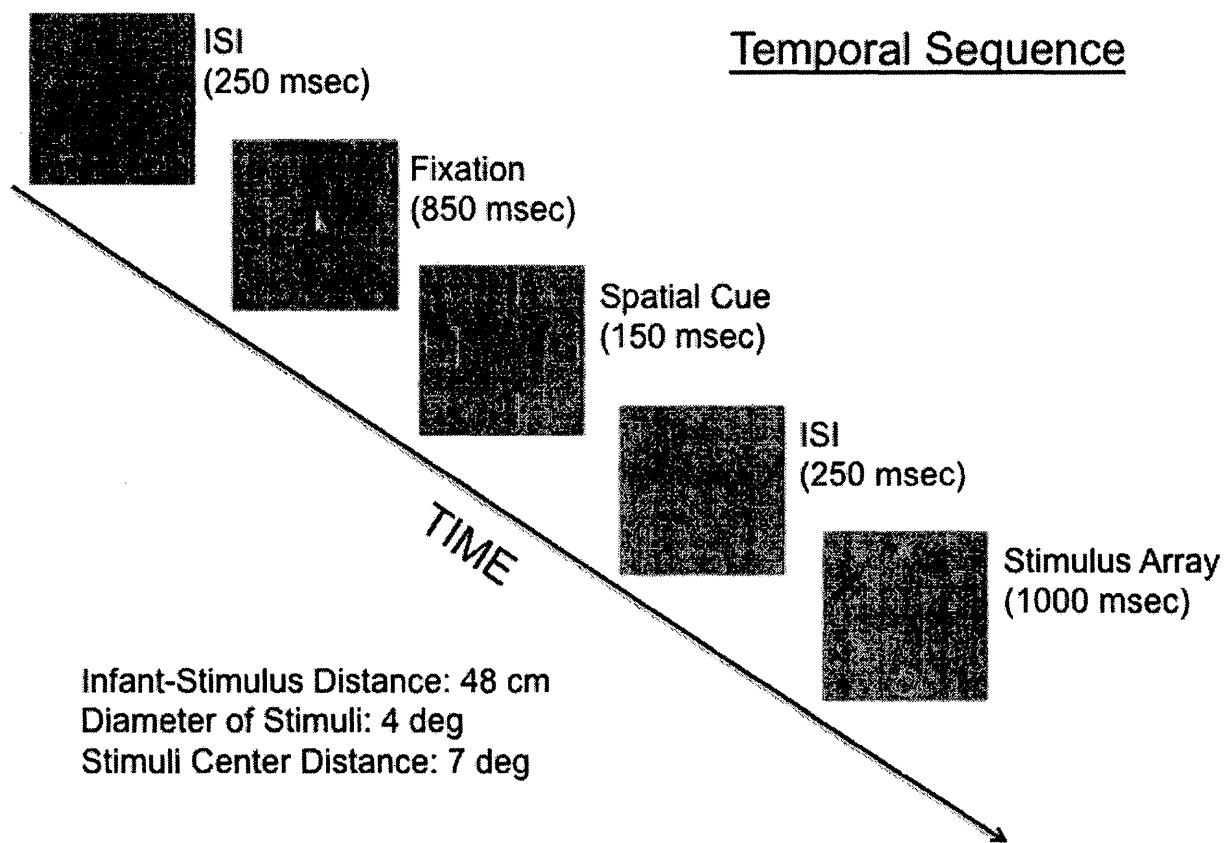


*Procedure*

In the testing room behind the black felt are two Dell computers. One generated the stimuli while the other controlled the eye tracker and collected eye movement data. One experimenter viewed the stimuli and the infant's eye movements on the data-collecting computer's monitor using a picture-in-picture video while controlling the camera using a remote control and starting the stimuli using a keyboard on the stimulus-generating computer. The other experimenter was lightly holding the infant's chin in order to reduce head movements and thereby reduce noise in the eye movement raw data collection. The digital data was written to the data-collecting computer and synchronized in real-time with a unique, stimulus-dependent, numerical code received from the stimulus-generating computer. That is, eye movements were coordinated with the specific stimulus array through a time-stamped numerical code (port code).

The experimental session was programmed using DirectRT research software version 2010 (Empirisoft Corporation, New York, NY). The trial began by presenting the fixation triangle for 850 msec alone and then for an additional 150 msec with the four rectangular frames located at each of the 12, 3, 6, or 9 o'clock positions on the circular pattern, cueing these locations as possible target locations. This was followed by an inter-stimulus interval (ISI) of 250 msec during which the monitor was blank. Immediately after this interval, the visual search array was displayed for 1000 msec during which either a target-present or target-absent array of one of the four conditions and one of the three set sizes were presented on the screen (see Figure 3). Which visual array presented was determined randomly across trials, with the constraints of target location, set size, and search array type being presented in a counter-balanced fashion on an equal number of trials each. The trial ended after an ISI of 250 msec was presented during

which again the monitor was blank. The next trial began when the fixation triangle and subsequent black frames were presented again.



*Figure 3.* Sample trial sequence for the visual displays for Experiment 1.

Infants were randomly assigned to one of the four conjunction-search conditions. When target was present, the target randomly occurred twice at each of the 4 possible locations (12, 3, 6, and 9 o'clock) for each set size (5, 8, or 10) resulting in 24 target-present trials. To be consistent with target-present arrays, target-absent arrays consisted of 4 trials of each set size being presented twice with the positions of the distractors randomized across trials for a total of 24 target-absent trials. Infants therefore received a total of 48 trials of randomly presented interleaved target-present and target-absent arrays for one of the four conjunction search conditions.

#### *Data reduction and analysis*

The raw digital data was recorded by the eye tracker and then imported into a MATLAB derivative called ILAB (Gitelman, 2002) for subsequent analysis. The ILAB toolbox software allows for the analysis of eye movements, separating out and displaying individually the horizontal and vertical components of the eye movement, on a trial-by-trial basis. Moreover, ILAB provides a mean by which to display the scan path of the eye at each trial and thereby determining whether or not the eye first fixates on the target and nature of the eye movement (latency, direction and distance) relative to the search array. Search behavior was assessed in the feature search conditions by measuring the latency value of the eye movement that localized the target in a target-present trial and the first eye movement to one of the distractors in a target-absent trial. Saccade latency was defined as the time taken from the appearance of the visual array to the time of the saccade initiated to localize a target or one of the distractors depending on array type (see Figure 4).

There were a number of criteria that an eye movement must meet in order to be included in the final data sample. First, anticipatory eye movements occurring before array onset and

within the first 167 msec after the onset of the visual search array for all conditions were excluded from the final data sample (even if eye movement correctly fixates on the target in a target-present trial). This onset latency was chosen because it has been determined that 3-month-old infants typically cannot make eye movements in reaction to the onset of stimulus faster than 167 msec (Canfield, Smith, Brezsnyak, & Snow, 1997). Secondly, infants had to remain fixated on a stimulus for at least 100 msec in order for it to be considered an eye movement (Adler & Haith, 2003; Adler & Orprecio, 2006). Thirdly, infants were required to fixate on the central fixation triangle prior to presentation of search array in order to determine the scan path on each trial and eliminate bias of producing eye movements that originate from random locations on the screen. Finally, the eye movement to a stimulus in the target-present and target-absent arrays had to trace a path that was more than 50% of the distance between the fixation triangle and the stimulus. This is because it has been determined that infants typically localize a target by way of a series of saccades of equal magnitude and that saccades are initiated towards the location of the target, but are terminated prior to reaching the target (hypometric saccades) (Shea, 1992). Also, this 50% criterion has been used in previous infant eye movement studies (Adler & Haith, 2003; Adler & Oprecio, 2006) and is typically taken as an indication that the eye movement is intentional and not random.

To summarize, in the target-present trials, the latency of the first eye movements that trace a path of at least 50% of the distance between fixation and target, and remains fixated for at least 100 msec, were measured. If the first eye movement traces a path to a distractor or to a location where no item is present, then this latency was not included and the trial was considered incorrect. Or if an eye movement traces a path to the correct location but does not remain fixated for at least 100 msec, it was not included. In target-absent trials, because there is no target, the

mean latencies of the first eye movements that trace a path at least 50% of the distance between fixation and any one of the stimuli and remains fixated for at least 100 msec in the target-absent array were analyzed.

In order to increase the power of our statistical tests, the data will be pooled across all infants who viewed a given array with a given set size (e.g. 8), with a specific target feature (shape or color), and with a specific target location (e.g. 12 o'clock). This is a common practice when eye movement latencies are the dependent measure (e.g. Adler, Bala, & Krauzlis, 2002; Adler & Oprecio, 2006).

## **Experiment 1 Results and Discussion**

### *Preliminary Analysis*

As there were four different targets (green or red “X” or “O”) in four different locations (3, 6, 9, and 12 o'clock), it was necessary to test that there were no latency differences due to target location, target shape, or target color in the target-present conditions as a function of set size (5, 8, and 10). Thus, a 3 (set size: 5, 8, 10) X 4 (location: 3, 6, 9, 12 o'clock) X 2 (Shape: X or O) X 2 (Color: green or red) ANOVA was conducted. Analyses showed that there were no significant main or interaction effects (see Table 1 for  $F$  and  $p$  values). These results suggest that there were no effects of target location, shape, color, or set size of the visual array on infants' search performance. Consequently, infant's data was pooled across location, shape and color for the main analysis.

Effect	dfs	F Value	p Value
Location	3, 39	1.44	0.24
Shape	1, 39	1.27	0.26
Color	1, 39	0.11	0.73
Set Size	2, 39	1.14	0.32
Shape X Color	1, 39	0.03	0.86
Shape X Location	3, 39	0.70	0.55
Shape X Set Size	2, 39	1.46	0.24
Color X Location	3, 39	0.73	0.54
Color X Set Size	2, 39	0.44	0.65
Location X Set Size	6, 39	0.97	0.46
Shape X Color X Location	2, 39	0.66	0.52
Shape X Color X Set Size	2, 39	0.33	0.72
Color X Location X Set Size	4, 39	0.45	0.77
Shape X Color X Location X Set Size	6, 39	0.62	0.72

*Table 1:* Preliminary analyses ANOVA results showing no significant differences across set sizes for color, shape, and location of the target.

### *Main Analysis*

Based on the preliminary analyses that revealed no significant differences in infants' search performance due to location, shape, or color, the data was pooled across these variables and a 2 X 3 ANOVA for array type (target-present vs. target absent) and set size (5, 8, and 10) was conducted. This analysis revealed a main effect of array type,  $F(1, 194) = 16.98, p < .001$ , indicating that across all set sizes infants' saccade latencies to a stimulus were faster to target-absent arrays ( $M = 447.07$  msec,  $SE = 18.07$ ) than to the target-present arrays ( $M = 579.24$  msec,  $SE = 48.46$ ). The main effect of set size was not significant  $F(2, 194) = 0.92, ns$ , nor was there an interaction effect of array as a function of set size,  $F(2, 194) = 0.18, ns$ , indicating that mean saccade latencies did not differ regardless of set size or across set size.

While these latter analyses revealed no significant latency differences between the set sizes, visual search tasks are typically measured by the RT slopes in relation to set size (Adler & Orprecio, 2006; Treisman & Gelade, 1980). Recall that slopes less than 20 msec per item are indicative of a parallel search in which target detection is unaffected by the number of distractors, whereas slopes greater than 20 msec per item are reflective of a serial search where as the number of distractors increases so does search time (Wolfe, 1989). In order to accurately assess the search phenomenon, therefore, a regression analysis was conducted on infants' saccade latencies as a function of set size. Also, because there was a significant difference between array types, regression analyses were performed separately for the target-present and target-absent arrays.

The dependent variable in the regression analyses was the mean saccade latencies, and the independent variable was set size. For target-present arrays, this analysis revealed a non-significant regression,  $r^2 = 0.43$ , indicating that saccade latencies did not increase with increasing



set size. Further support is provided by the fact that the slope of the regression line was 9.02 msec per item, indicating a relatively flat slope (see Figure 4). These results show that the latency of infants' eye movements were unaffected by set size. For the target-absent arrays this analysis also revealed a non-significant regression,  $r^2 = .40$ , indicating that saccade latencies did not increase with increasing set size. The slope of the regression line was 10.83 msec per item further supporting that infants' eye movements were unaffected by set size (see Figure 4).

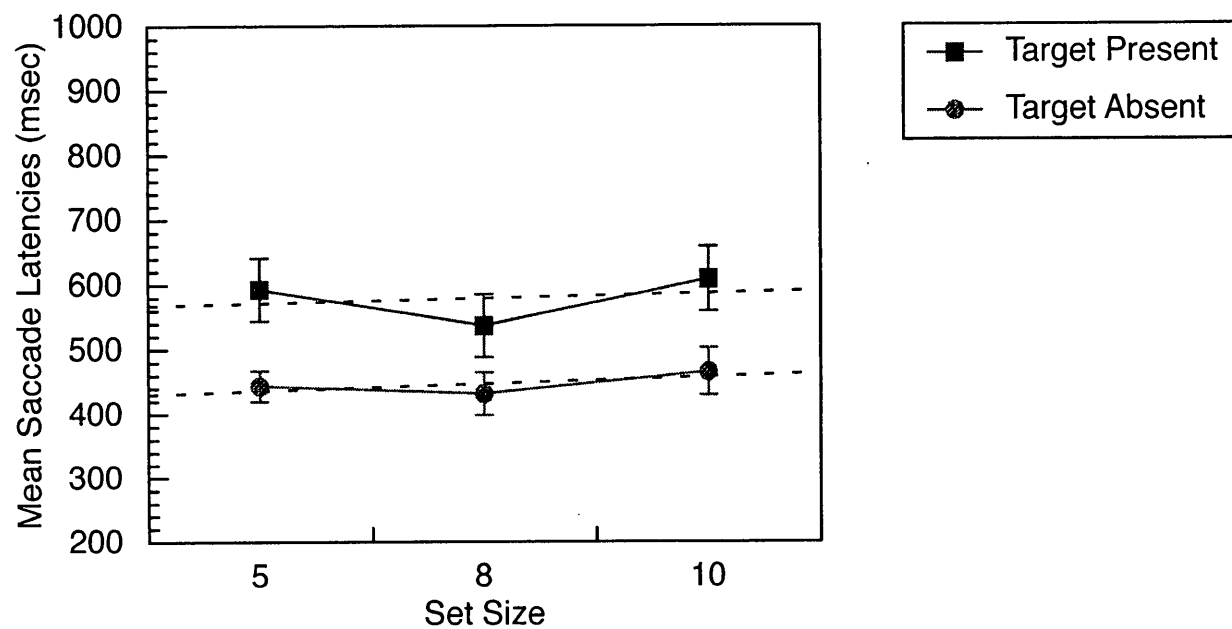


Figure 4: Infants' mean saccade latencies in each Set Size X array type (target-present and target-absent) condition. Latencies to both the array types remained relatively flat as set size increased.

### *Number of Eye Movements*

Results from the regression analysis revealed a flat search function when the target was both present and absent. This flat search function suggests an efficient search. However, the saccadic latency that localized the target in the feature-present condition was not always infants' first eye movement, suggesting that the target did not pop out and an inefficient search was produced. If the target popped out amidst the distractors, generally only one eye movement is needed for infants (Adler & Orprecio, 2006). Thus, we need to compare the mean number of eye movements made across set size to determine if it was an efficient or inefficient search. Because there were only distractors in the target absent condition the first eye movement always localized one of the distractors. For the target-present arrays, infants' generally required between one to two eye movements to localize the target ( $M = 1.44$ ,  $SE = .07$ ). This shows that although the infants' search slope remained relative flat across set sizes, suggesting a relatively efficient search, the target item did not pop out since they needed to make more than one eye movement to localize the target.

To analyze the relation between number of eye movements and set size, a one-way ANOVA was conducted with the number of eye movements as the dependent variable and set size as the independent variable. Results revealed no significant differences in the number of eye movements required to find the target across the set-sizes 5, 8, and 10,  $F(2, 77) = 0.04$ , *ns*. This finding suggests that infants' executed approximately the same number of fixations 95% CI [1.29, 1.60] across each set-size prior to finding the target. These results may provide a suggestion as to why infants' saccade latencies do not increase with the increasing set size, which is typical for adult conjunction search. This will be further discussed in the discussion section.

### *Split-Half Data Analysis*

Typically in adult cueing visual search studies, participants are told that there will be a cue prior to the onset of the search array (Sobel et al., 2009; Posner, 1980). Since we are looking at search behavior of 3-month-olds, they lack the developmental capacities to be told what the cue or target is. Thereby they may need some time to implicitly learn the cue and utilize it. In order to assess whether infants' learned and then used the cue, a split-half analysis was conducted comparing the first half of every infants' trials with the second half of all infants' trials. For this split half analysis we conducted a 2 (array type: target-present and target-absent) X 2 (halves: first half and second half) X 3 (set size: 5, 8, and 10) ANOVA. There were no significant main effects for set size,  $F(2, 185) = .94$ , *ns*, or for halves,  $F(1, 185) = 0.01$ , *ns*. These results indicate that there was no significant difference between infants' eye movement latencies in the first half of trials and the second half. Nor did latencies differ as a function of increasing set size. There was again a significant main effect for array type,  $F(1, 185) = 18.62$ ,  $p < .001$ , with target-absent latencies ( $M = 447.07$  msec,  $SE = 18.07$ ) being faster than target-present latencies ( $M = 580.36$  msec,  $SE = 28.22$ ). This result shows that infants are faster at making an eye movement to one of the distractors when there is no target present compared to making an eye movement to the target when it is present. Finally, there were no significant interaction effects between array type and split halves,  $F(1, 185) = 0.00$ , *ns*; array type and set size,  $F(2, 185) = 0.33$ , *ns*; split halves and set size,  $F(2, 185) = 1.51$ , *ns*; and array, split halves, and set size,  $F(2, 185) = 0.75$ , *ns*.

As with the data for all of the trials, regression analyses were conducted for the split half results in order to observe the slopes of the latencies as a function of increasing set sizes. As with the full results, the dependent variable was the mean saccade latencies with set size as the

independent variable. For the first half target-present regression analysis revealed a non-significant  $r^2 = 0.13$ , with the slope being -3.54 msec per item. The second half target-present also resulted in a non-significant  $r^2 = .51$  with a slope of 8.36 msec per item. For the first half of target-absent analysis, there was a non-significant  $r^2 = .66$  with a slope of 3.60 msec per item. And finally for the second half of the target-absent analysis there was a non-significant  $r^2 = .96$  with a slope of -2.09 msec per item (see Figure 5). These analyses all resulted in relatively flat search functions, indicating that latencies did not increase as a function of set size. Also, latencies did not differ from the first or second halves of infant trials.

Thus, the results from this experiment reveal that with exposure to the bottom-up spatial cues, infants' performance in searching for and localizing the target did not change. Such a finding suggest that the nature of infants' immaturity in producing adult-like conjunction search behavior is likely not due to a lack of development in their bottom-up attentional mechanisms for processing the visual input. Instead, infants' immature conjunction search behavior may reflect an early developmental phase in their top-down attentional mechanisms.

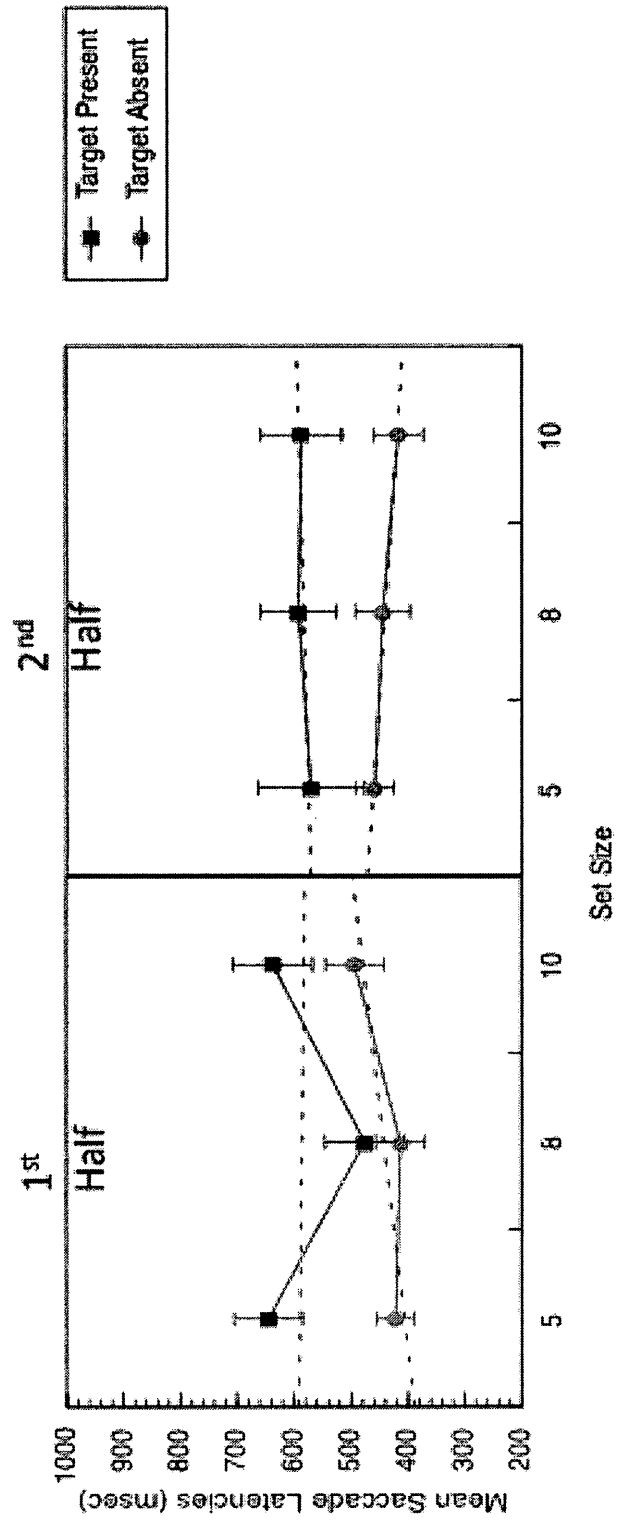


Figure 5: Split half regression analysis for first and second half of trials in both the target-present and target-absent conditions.

## **Experiment 2: User-Driven Top-Down Cues**

In the present experiment, we intend to assess the effect of top-down user-driven identity cues in the development conjunction search by measuring 3-month-old infants' saccadic eye movement latencies. As previously discussed 3-month-old infants' were unable to perform a conjunction search like that of adults without the facilitation of cues (Fuda & Adler, 2012). Our interpretation of those findings was that there is differential development of bottom-up versus top-down attentional mechanisms, with top-down mechanisms showing a prolonged maturational timeline. The previous experiment is meant to assess the relative role of bottom-up mechanisms by testing whether spatial stimulus-driven cues can facilitate infants' performance in a conjunction search. In contrast, the current experiment is intended to assess the relative role of top-down mechanisms by testing whether infants' conjunction search performance is facilitated when provided with top-down information as to the color and shape of the target before the search arrays are presented. Previous work with adults has shown that observers respond quicker to targets when they are provided with prior information such as the color of the target, rather than orientation (Sobel et al., 2009). Or that, when observers are provided with a picture cue rather than a written cue of the target prior to the onset of a conjunction search array, then they localize the target quicker (Wolfe et al., 2004). Consequently, if infants' top-down attentional mechanisms are under-developed, then there are three main hypotheses as to the effect of top-down cues. First, we hypothesize that when infants are provided with top-down cues indicating what the target is (color and shape) prior to the conjunction search arrays, they will need to interpret the target features and thus be able to search for that specific target amongst varying numbers of distractors. We predict therefore that with this extra information, infants will exhibit a conjunction search function more like that of adults, where such attentional resources would be

shown if saccade latencies to find the conjunction target increased linearly with increasing set sizes, producing a positive latency-by-set size slope. Secondly, we also hypothesize an inefficient search when the target is absent. This would be confirmed if the saccadic latencies to the stimuli increase linearly with increasing set sizes. Finally, we hypothesize that because infants are able to perform a feature search similarly to adults, but not a conjunction search, they will need more assistance in top-down attentional guidance, thus these top-down target cues will aide in their search more so than the bottom up spatial cues from Experiment 1. This last hypothesis will be confirmed if the saccadic latencies of Experiment 2 are faster and show a greater positive latency-by-set size slope than Experiment 1's results.

## Methods

### *Participants*

Nine (7 males, 2 females) infant participants who ranged from 91 to 112 days ( $M = 103.3$ ,  $SD = 7.4$ ) were included in this study. Infants were recruited in the same manner as in Experiment 1.

### *Stimuli and apparatus*

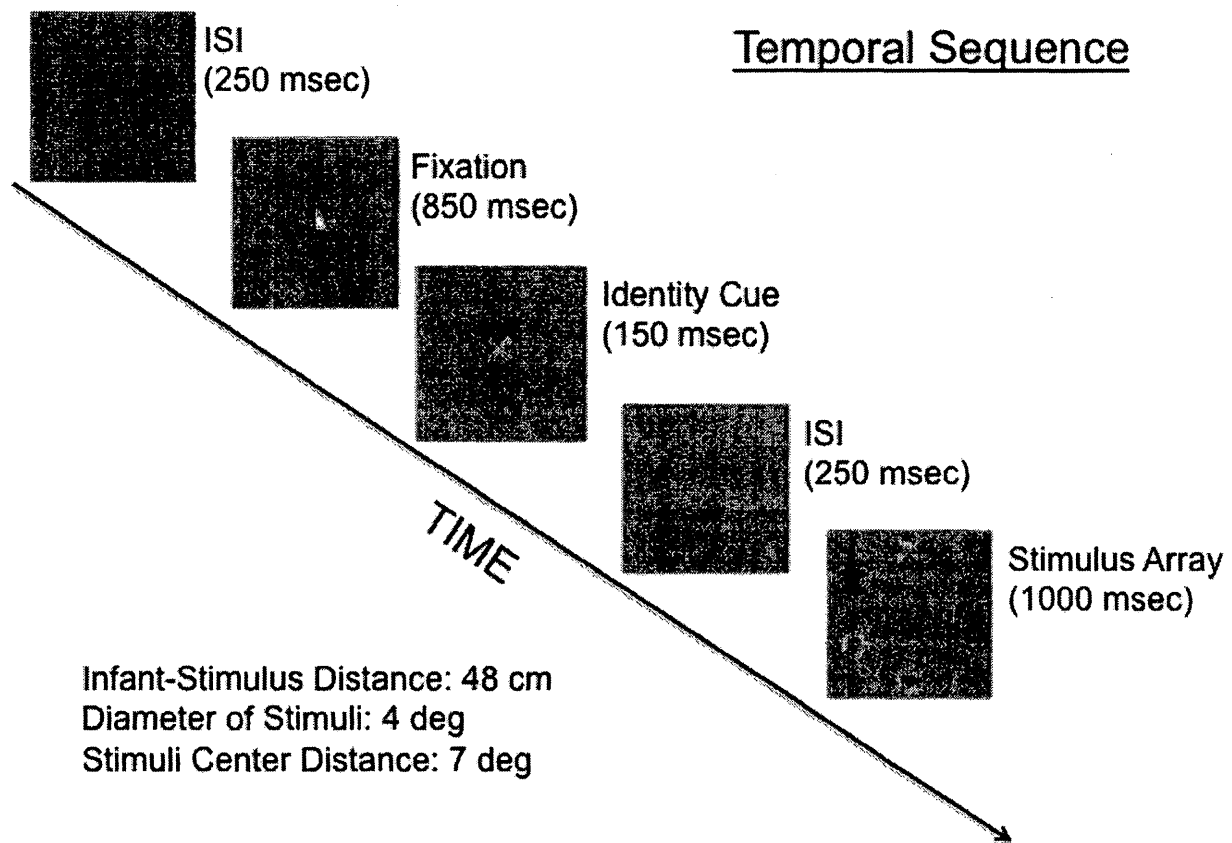
The apparatus was the same as Experiment 1. The stimuli consisted of the same fixation triangle, and targets of green or red X's or O's from Experiment 1. However the four rectangular frames were not included, and instead one of four possible top-down identity cues (green X, green O, red X, red O) were presented in the center of the array.

### *Procedure*

The procedure was the same as Experiment 1 with a couple of exceptions. First, infants were no longer shown the spatial frame cues prior to the search array indicating the possible locations of the target. Instead, for this experiment, infants were shown what the target is that



they are searching for prior to the onset of the search array. For instance, as in the previous experiment, a trial began with the fixation triangle being presented for 850 msec. In contrast to Experiment 1, the triangle disappeared after 850 msec and was replaced immediately by one of the four targets (green or red X or O) for 150 msec. There was an ISI of 250 msec, followed by the conjunction search array (see Figure 4). As with Experiment 1, infants were randomly assigned to one of the four conjunction-search conditions. When the target is present, the target occurred twice at each of the 4 possible locations (12, 3, 6, and 9 o'clock) for each set size (5, 8, or 10) resulting in 24 target present trials. To be consistent with target-present arrays, target-absent arrays consisted of 4 trials of each set size being presented twice with the positions of the distractors randomized across trials. Infants therefore received a total of 48 trials of randomly presented interleaved target-present and target-absent arrays for one of the four conjunction search conditions.



*Figure 6.* Sample trial sequence for the visual displays in Experiment 2.

### *Data reduction and analysis*

The raw digital data were recorded, reduced, and analyzed in the same manner as in Experiment 1.

## Results

### *Preliminary Analysis*

Similarly to Experiment 1, there were four different targets (green or red “X” or “O”) in four different locations (3, 6, 9, and 12 o’clock). As such, it was again necessary to determine that there was no latency differences due to target location, target shape, or target color in the target present conditions as a function of set size (5, 8, and 10). Thus, a 3 (set size: 5, 8, 10) X 4 (location: 3, 6, 9, 12 o’clock) X 2 (Shape: X or O) X 2 (Color: green or red) ANOVA was conducted. Analyses showed that there were no significant main or interaction effects (see Table 2 for *F* and *p* values). These results show that location, shape, or color of the target did not have any effect on set size, thus this data was pooled across these factors for the main analysis.

Effect	df's	F Value	p Value
Location	3, 46	0.10	0.96
Shape	1, 46	0.10	0.75
Color	1, 46	0.56	0.45
Set Size	2, 46	0.86	0.42
Shape X Color	1, 46	0.20	0.65
Shape X Location	3, 46	0.54	0.65
Shape X Set Size	2, 46	0.29	0.75
Color X Location	3, 46	2.49	0.07
Color X Set Size	2, 46	2.55	0.08
Location X Set Size	6, 46	0.54	0.77
Shape X Color X Location	3, 46	1.04	0.38
Shape X Color X Set Size	2, 46	0.17	0.84
Color X Location X Set Size	6, 46	0.89	0.77
Shape X Color X Location X Set Size	5, 46	0.23	0.94

Table 2: Preliminary analysis ANOVA results for identity cue.

### *Main Analysis*

Since the preliminary analysis revealed no significant differences across set sizes for location, color, or shape of the target, the data was pooled across these factors in order to determine the relationships between latencies, set size (5, 8, and 10), and array type (target-present and target-absent). To this end, a 3 (set size: 5, 8, and 10) X 2 (array type) ANOVA was conducted with saccade latencies as the dependent variable. Results revealed that there was no main effect of set size,  $F = 0.92$ , *ns*, indicating that contrary to our hypothesis, latencies did not change as a function of set size. Consistent with Experiment 1, there was a main effect of array type,  $F(1, 224) = 11.61$ ,  $p < 0.001$ , with target-absent latencies ( $M = 476$ ,  $SE = 17.89$ ) being approximately 100 msec faster than target-present latencies ( $M = 586.74$ ,  $SE = 29.91$ ). There was also an interaction effect between array type and set size,  $F = 4.02$ ,  $p = .019$ . Multiple pairwise comparisons with a Bonferonni correction of  $0.05/9 = .006$  revealed that out of the nine comparisons the only significant difference was between the target-present set size of 10 and target-absent set size of 10,  $t(29) = 3.42$ ,  $p = .002$ , with the target-absent set size of 10 ( $M = 439.93$ ,  $SE = 41.51$ ) being significantly faster than the target-present set size of 10 ( $M = 661.67$ ,  $SE = 52.86$ ) (for all nine comparisons see Table 3).

Conditon	df's	t-test	p-value
TP(5) - TP(8)	27	0.36	0.72
TP(5) - TP(10)	27	-1.6	0.10
TP(8) - TP(10)	27	-1.64	0.12
TA(5) - TA(8)	45	1.64	0.11
TA(5) - TA(10)	45	2.69	0.01
TA(8) - TA(10)	45	0.85	0.40
TP(5) - TA(5)	27	0.49	0.62
TP(8) - TA(8)	28	1.12	0.27
TP(10) - TA(10) *	29	3.42 *	0.002 *

Table 3: Pairwise comparisons for interaction effect between set size and array type where TP(#) is target-present (set-size) and TA(#) is target-absent (set-size)

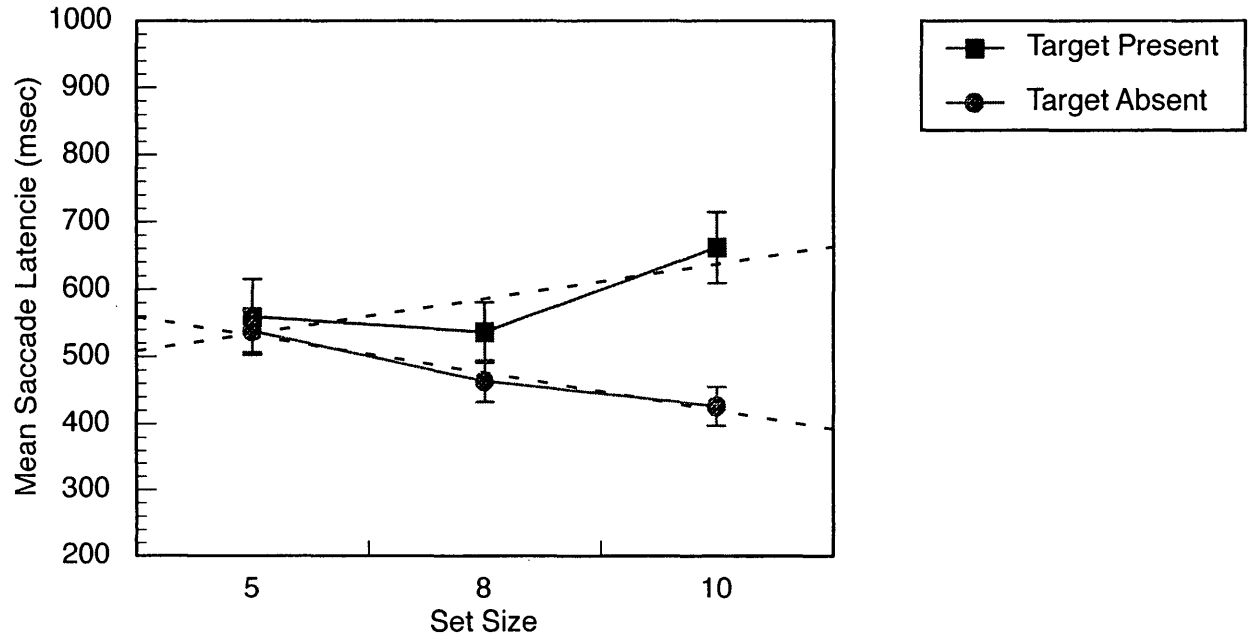
\* represents significant results with a Bonferonni correction of .005.

Although these latter analyses revealed no significant latency differences between the set sizes, as indicated previous, visual search tasks are typically measured by the RT slopes in relation to set size (Treisman & Gelade, 1980). Recall that slopes less than 20 msec per item are indicative of a parallel search in which target detection is unaffected by the number of distractors, whereas slopes greater than 20 msec per item are reflective of a serial search where as the number of distractors increases so does search time (Wolfe, 1989). In order to accurately assess infants' search behavior with top-down identity cues, therefore, a regression analysis was conducted on infants' saccade latencies as a function of set size. Also, because there was a significant difference between array types, regression analyses were performed separately for the target-present and target-absent arrays.

The dependent variable in these analyses was the mean saccade latencies, and the independent variable was the set size. For target-present arrays, this analysis revealed a non-significant regression,  $r^2 = .59$ , indicating that saccade latencies did not increase with increasing set size. The slope of the regression line was 52.24 msec per item, suggesting an increasing search function, even though the relation between latency and set size was not significant. For the target-absent arrays, the regression analysis did reveal a significant relation,  $r^2 = .96$ , indicating that saccade latencies decreased with increasing set size. The slope of the regression line was -55.91 msec per item, indicating a steep negative slope (see Figure 4). This result is contrary to the majority of visual search research (Adler & Orprecio, 2006; Treisman & Gelade, 1980; Wolfe, 1989), and thus contrary to our hypothesis suggesting an increasing search slope when the target is absent. This negative search function, however, is not unseen. Other visual search studies investigating crowding effects as well as priming effects have similarly found search functions with negative slopes (Bacon & Egeth, 1991; Bravo & Nakayama, 1992;

McSorely & Findlay, 2003). Theoretical explanations for such negative slopes will be discussed more in the discussion.





*Figure 7:* Mean saccade latencies for the identity cue for both target-present and target-absent arrays across set sizes. The target absent condition resulted in a significant negative slope.

### *Number of Eye Movements*

As in Experiment 1, the saccadic latency that localized the target in the feature-present condition was not always infants' first eye movement, suggesting that the target did not pop out and an inefficient search was produced. If the target popped out amidst the distractors, generally only one eye movement is needed for infants (Adler & Orprecio, 2006). Thus, we needed to compare the mean number of eye movements made across set size to determine if it was an efficient or inefficient search. Because there were only distractors in the target absent condition the first eye movement always localized one of the distractors. For the target-present arrays, infants' generally required between one to two eye movements to localize the target ( $M = 1.49$ ,  $SE = 0.07$ ). This shows that the target item did not pop out since they needed to make more than one eye movement to localize the target.

To analyze the relation between the number of eye movements as a function of set size to target-present arrays, a one-way ANOVA was conducted with the number of eye movements as the dependent variable and set size as independent variable. Results revealed no significant differences in the number of eye movements required to find the target across the set-sizes 5, 8, and 10,  $F(2, 85) = 1.78$ ,  $p = 0.17$ , *ns*. This finding suggests that infants' executed approximately the same number of fixations 95% CI [1.36 – 1.62] across each set-size prior to finding the target. If infants' saccade latencies had increased as a function of set size as would be expected in a conjunction search, then the lack of a relation between the number of eye movements and set size would have precluded explaining the latency increase as due to an overt shifts (eye movements) of attention. Instead, any latency increase would have been due to top-down serial search mechanisms. That infants do not show a latency increase nor show an increase in the

number of eye movements suggests that the immaturity in their attentional mechanism lies in their application of top-down attentional resources.

### *Split-Half Results*

To repeat, typically in adult cueing visual search studies, participants are told that there will be a cue prior to the onset of the search array (Maljkovic & Nakayama, 1994; Nakayama & Silverman, 1986; Reddy & VanRullen, 2007; Sobel et al. 2009). As before, since we are looking at search behavior of 3-month-olds, they lack the developmental capacities to be told what the cue or target is. In this experiment, infants are given a cue with information about the targets' shape and color. As such they may need time to implicitly learn and remember the cue information in order to use this information when the search array is presented. A split-half analysis was therefore conducted in order to assess whether infants learned and used the cue by comparing their first half of trials with the second half of trials. Infants' data from the first half of trials were compared with infants' data from the second half of trials.

A 2 (array type: target present and target absent) X 3 (set size: 5, 8, and 10) X 2 (split-half: 1<sup>st</sup> versus 2<sup>nd</sup>) ANOVA was conducted, with saccade latencies as the dependent variable. Results revealed a significant main effect for array type,  $F(1, 215) = 11.67, p < 0.001$ , with target-absent latencies ( $M = 476.84, SE = 17.89$ ) being overall faster than target-present latencies ( $M = 586.74, SE = 29.91$ ). These results are consistent with those found in Experiment 1. There were no main effects for split halves,  $F(1, 215) = 0.34, ns$ , or for set size,  $F(2, 215) = 1.17, ns$ , showing that the first half did not differ from the second half and that latencies did not differ across set sizes. There were two significant interaction effects. As in Experiment 1 and with the full results in this experiment, the interaction of array type by set size,  $F(2, 215) = 3.92, p = 0.02$ , was significant. Pairwise comparisons with Bonferroni correction again revealed that this

was due to the difference between target-absent set size of 10 compared with the target-present set size of 10. Additionally, a 3-way interaction was found for array type, split halves, and set size,  $F(2, 215) = 3.37, p = 0.03$ . This result shows that the saccadic latencies for each set size were different for each array type and for each half. As can be seen in Figure 6, this is likely representative of the fact that latencies to target-present arrays increased as a function of set size in the second half when they had not in the first half, whereas latencies to target-absent arrays decreased as a function of set size in the second half when they had not in the first half. To further explore these relations, regression analyses were performed as before.

Four regression analyses were conducted for the 2 array types in both the first and second halves. For the target-present array, the regression over the first half of trials resulted in a non-significant relation,  $r^2 = .91$ , with a slope of -42.96 msec per item. For target-absent array, the regression for first half of trials was also not significant, with an  $r^2 = .90$  and a slope of -54.71 msec per item. For the target-present array, the regression over the second half of trials revealed a significant relation,  $r^2 = 0.73, p = 0.01$ , and a slope of 131.50 msec per item. These results confirm the hypothesis that as set size increases so does the saccade latencies required to localize the target. For the target-absent array, the regression for the second half of trials also revealed a significant relation,  $r^2 = 0.99, p = 0.04$ , but in the opposite direction from the target-present second half regression, as indicated by a negative slope of -57.25 msec per item (see Figure 6).

In contrast the findings of Experiment 1 with bottom-up spatial cues, the results from this experiment with a top-down identity cue reveal that with continued exposure to the top-down cue, infants' performance in searching for and localizing the target exhibited dramatic change. That is, with continued exposure to the top-down cue, infants implicitly learned the nature of the features (color and shape) in the cue and used that information to increase the top-down

activation of the processing of those features in each search array (Wolfe, 1989). As a consequence, when the target was present, their search behavior more similarly resembled adult-like search behavior in the second half of the trial sequence. Such a finding suggests that the nature of infants' immaturity in producing adult-like conjunction search behavior is likely due to their top-down attentional mechanisms being in an early developmental phase.

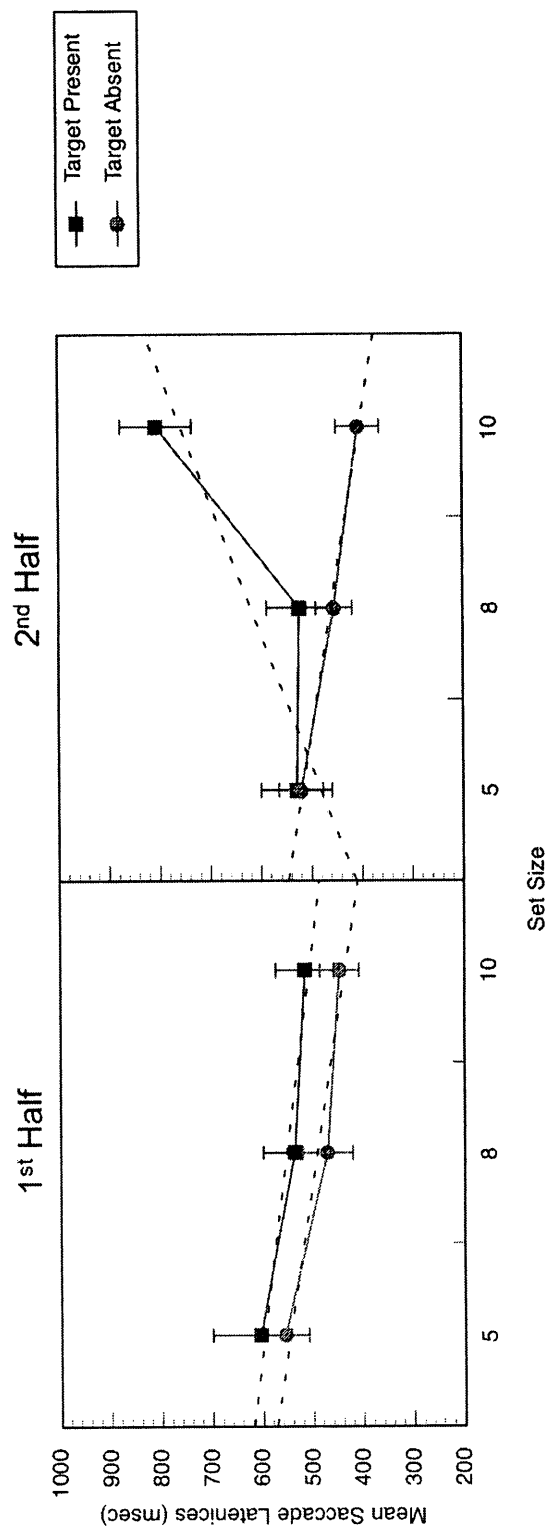


Figure 8: Mean saccade latencies plotted as a function of set size for target-present and target-absent conditions for the first and second half of infants' trials.

## Discussion

The ability to selectively attend is crucial for efficient processing and behavioral responding as it allows for segregation and selective focus on items of interest while at the same time ignoring irrelevant information. In infancy, this ability is especially important, as infants' knowledge base is comprised of information gathered from their environment. The ability to selectively attend, therefore, would allow infants to search and attend to items of interest in a dense and what would otherwise be an overwhelming environment of stimulation (Adler et al., 1998; Columbo et al., 1990). Consequently, investigating the ability to selectively attend to a stimulus in the face of competing distractors not only provides us with the opportunity to better understand infants' visual attention but it may also shed light on the infants' ability to extract specific perceptual information from the environment. The current study sought to investigate whether the mechanisms involved in selective attention develop at different rates or mature at different ages.

Previous research with infants has suggested that they are fully capable of performing a feature search in a manner nearly identical to adults (Adler & Orprecio, 2006), but are developmentally immature in localizing a target in a conjunction search (Atkinson & Braddick, 1992; Bhatt et al., 1999; Gerhardstein & Rovee-Collier, 2002). Many of these conjunction search studies with infants have, however, used methods that were not comparable to those used in adult visual search research. The infant studies used paradigms in which search behavior was measured in time scales that spanned over seconds and minutes, compared to milliseconds in adult studies, allowing the possibility that infants' performance was due to not just attentional resources, but perhaps higher-order processes such as memory. Only recently, Fuda and Adler (2012) directly compared 3-month-old infants' visual search performance with adults' in a

conjunction search by using a paradigm that potentially isolated their selective attentional mechanisms by measuring saccade latencies that occur on the order of milliseconds. Results of their study, consistent with previous infant studies, also concluded that infants lacked the ability to perform a conjunction search in a similar manner to adults. An explanation for the difference in performance between feature and conjunction searches in infancy was attributed to Wolfe et al.'s (1989) Guided Search model of visual search, in which feature searches are thought to rely mainly on bottom-up attentional resources to localize a target, whereas conjunction searches are theorized to require both bottom-up and top-down attentional resources.

More specifically, Guided Search theorizes that in feature searches the target greatly contrasts from its distractors in a particular feature dimension so that a stark difference in bottom-up activation at the location of the target occurs. Consequently, bottom-up mechanisms are sufficient for detecting the target and top-down activation is not necessary. In conjunction searches, however, because the target shares all of its features with various distractors in the array, sharp differences in bottom-up activation are not as strong and thus top-down activation is required to disambiguate the target from the surrounding distractors and thereby localize the target (Wolfe et al., 1989). Because infants have been shown to perform a feature search like that of adults (Adler & Orprecio, 2006), but have been shown not to be able to perform a conjunction search in a similar manner (Fuda & Adler, 2012), perhaps this is due their top down attentional resources being developmentally immature. If this hypothesis were correct then we would expect that bottom-up cues would not facilitate a conjunction search function whereas top-down cues would. The current study, in testing this hypothesis, showed just such a pattern in that bottom-up spatial cues failed to facilitate the exhibition of an adult-like conjunction search function but a top-down identity cue did. In other words, the findings from the present study



would seem to provide evidence that infants' top-down attentional resources are more developmentally immature relative to bottom-up resources (Adler & Orprecio, 2006; Fuda & Adler, 2012).

### *Bottom-Up Mechanisms in Infants*

The findings of the present study proved to be both straightforward and contradictory when compared to past research. Previous research of adult visual search suggests that an efficient search relies primarily on bottom-up attentional mechanisms (Treisman & Gelade, 1980; Wolfe, 1994). Consequently, allocation of attentional resources towards target items occur in a spatially parallel manner resulting in response times (e.g. the latency of saccadic eye movements) that are unaffected by the number of distractors (Treisman & Gelade, 1980; Wolfe, 1989). In contrast, an inefficient search has been suggested to require both bottom-up and top-down attentional mechanisms, in which allocation to a target occurs in a more serial search manner so that response times increase with the number of distractors (Wolfe, 1989). Some adult studies, however, have shown that a search which would typically be categorized as inefficient can result in an efficient search when cues displaying either possible location of the target or certain target features are provided prior to the search array (Posner, 1980; Sobel et al., 2009; Wolfe et al., 2004). The cues are thought to facilitate search behavior by priming the activation of the bottom-up or top-down mechanisms. Such effects have not been demonstrated to date in development. Yet, if infants' conjunction search (typically inefficient) is developmentally immature, then cueing either their bottom-up or top-down mechanism might illuminate which mechanism is responsible for that developmental immaturity. To this end, spatial cues were used to facilitate bottom-up mechanisms and identity cues were used to facilitate top-down mechanisms. Depending on the relative effects of these cues on infants'

search performance, the relative development of infants' bottom-up and top-down mechanisms might be assessed and their contribution to infants' search behavior can be determined.

In the current study, the spatial (bottom-up) cue did not result in infants' saccade latencies increasing with the increasing set size. This would seem to suggest that the spatial cue did not facilitate infants' bottom-up mechanisms performance in such a way as to switch the type of search (efficient versus inefficient) to reflect a more adult-like trend whose latencies tend to increase with increasing set sizes in a conjunction search. Thus, these findings would suggest that infants' bottom-up attentional mechanisms are relatively mature in early development and are not responsible for the lack of an adult-like conjunction search behavior.

Contrary to adult conjunction search and cueing studies, the results of current study would suggest that infants' conjunction search was efficient as their search function was flat across set size, and remained so even after the addition of the bottom-up spatial cue. If infants were exhibiting an efficient conjunction search, in which case the target would have popped out irrespective of the number of distractors, then we would have expected infants' first eye movement to localize the target. Results of the current study, however, revealed that infants typically required approximately two eye movements to localize the conjunction target regardless of set size, suggesting that the target did not pop out -- as would be expected of an efficient search. What might be the reason for a flat function that is not an efficient search, even with a spatial bottom-up cue? The reason that this type of cue did not facilitate infant search may be understood by considering Posner's (1980) general model of attention. Posner (1980) proposed that in order to exhibit overt attention, there needs to be an initial disengagement from one stimulus, movement to another stimulus, and then engagement of the new stimulus. Johnson, Posner, and Rothbart (1991) showed that disengagement from one stimulus to another occurred

significantly more rapidly in 4-month olds compared to 3- and 2-month olds. The issue therein is that perhaps 3-month old infants are taking too long to disengage from their initial eye movement to then find the target on their second eye movement. Thus, the latency of the second eye movement, which actually localizes the target, is the fastest that they can generate within the 1000 msec search window provided, irrespective of set size. Another way of viewing this is that the amount of time available between the fastest that they can make the second eye movement and the amount of time left in the presentation of the search array is not sufficient to allow exhibition of an increasing search function – even when the underlying attentional mechanisms are inefficient. Exposure of the spatial cues was therefore not sufficient to produce enough of a spread in the timing to allow exhibition of an increasing search function, likely due to infants' bottom-up mechanisms being fairly mature (Hood, 1993; Johnson, Posner, & Rothbart, 1991) and the cues not providing enough additional activation to bottom-up mechanisms to facilitate search.

More importantly, however, this explanation could account for why infants in the current study were faster overall than the original Fuda and Adler (2012) conjunction search study. The bottom-up cues in the current study likely did succeed in narrowing the infants' search window from 12 to 4 possible target locations. Such narrowing apparently affected overall search performance, resulting in infants eye movement latencies being faster at each set size than infants' search performance in Fuda and Adler's (2012) study when they were not provided with spatial cues. That is, the spatial cues decreased the possible locations where the conjunction target could appear, thereby narrowing the possible locations infants elicited bottom-up attentional resources before finding the target. As a consequence, since the spatial cues

narrowed the search possibilities equally across all set sizes, there effect was equivalent across set sizes.

The issue, however, still remains in both the original and the current cueing study that infants' conjunction search function did not increase linearly as it does for adults. This could be related to the fact that the four spatial cues did not disambiguate infant search performance as much if there were only two spatial cues. The reason for this being that infants' in general only made 2 eye movements per trial because it was suggested that they cannot disengage from items fast enough in the 1000 msec search window to make further eye movements. If two spatial cues were provided instead of four, then perhaps infants would have been better able to coordinate their two movements surrounding the spatial location cues provided. For instance, if they were utilizing the cue, they would make an eye movement to where one spatial cue was, but if the target was not in this location, they would make an eye movement to where the other spatial cue suggested and hence localize the target. This is an issue that would like to be subsequently addressed in later studies.

Additionally, infants did not show an increasing search function after presentation of the spatial cues because those cues facilitated bottom-up mechanisms. As stated previously, infants likely have sufficiently developed bottom-up attentional mechanisms, which is why previous research has shown they can perform a feature search nearly identical to adults (Adler & Orprecio, 2006; Fuda & Adler, 2012). According to Guided Search theory, when a target is highly contrasted against its distractors, such as in a feature search, bottom-up mechanisms are sufficient to activate the feature map and drive attention to localize the target, thus making it an efficient search. In an inefficient conjunction search, however, the target is not highly contrasting compared to its distractors because the target shares two of its features with alternating

distractors. Thus activation of multiple feature maps (i.e. color and shape maps) would both require activation and top-down guidance to connect the two activations to drive attention. In this study the search array that was presented to them after the cue was a conjunction array, which requires both bottom-up and top-down search mechanisms to detect and localize the target (Wolfe, 1989). Because infants are already equipped with functional bottom-up mechanisms, providing them with bottom-up cues did not provide any additional facilitation of those already functional mechanisms, nor did they facilitate the top-down attentional guidance that is also required to perform the task. Generally these results show that infants' are at minimum capable of using bottom-up cues to speed their overall search performance, but that bottom-up cues do not result in them performing in a similar manner to adults in a conjunction search.

#### *Top-Down Mechanisms in Infants*

The identity cue, which presumably facilitated top-down attentional mechanisms, provided somewhat contrasting results relative to the bottom-up spatial cues. The overall results showed that when a target was present infants' latencies did increase with increasing set sizes; however, these results were not significant. Only after conducting a split-half analysis that it was shown during the second half of infants' trials, latencies significantly increased with the increasing set sizes. This shows that the infants needed time to implicitly learn the cue before they were able to use it. Moreover, this finding suggests that infants' top-down mechanisms are relatively weak and fairly immature and require extensive priming over repeated exposure in order to begin to modulate behavior.

Most adult visual search studies provide participants with either knowledge of the target or practice trials prior to the actual experiment to familiarize them the experiment expectations (Maljkovic & Nakayama, 1994; Nakayama & Silverman, 1986; Reddy & VanRullen, 2007;

Sobel et al. 2009). Even studies with infants have provided their participants with prior training (Johnson et al., 1991) or practice trials (Gerhardstein & Rovee-Collier, 2002) before the actual experiment. In the current study, infants' first half of the trials could be considered their learning or training period in which there was no linear increase of latencies across set size. Once infants learned the characteristics of the target (color and shape) and the temporal sequence of the trial presentation were they able to fully use the cue. As a consequence, the time it took for infants' to localize the target amongst a group of distractors increased as the number of distractors increased. These results confirm our hypothesis suggesting that the identity cue facilitated infants' top-down mechanisms in order to perform a conjunction search similar to adults.

In relation to Wolfe's (1989) Guided Search model, the current study results show that that in order for infants to localize most items in their visual world, they pool activation from both bottom-up activation maps that merely register the saliency of each feature at each location combined with top-down activation maps that bias items in certain feature maps. When providing 3-month-olds with the added knowledge of target shape and color, items in those feature maps are provided with biased activation. Pooling the top-down activation with the bottom-up feature activation enable infants to initiate an adult-like search for the conjunction target. Without such top-down activation, infants are left to function and localize the target on the basis of primarily their bottom-up mechanisms, which are not sufficient to produce adult-like search in the time limitation discussed earlier. These results are substantive in that they show evidence that top-down attentional mechanisms in infancy develop later than bottom-up attentional mechanisms required for visual search.

Where the results with the top-down identity cue differ from past attentional research is when the target was absent. Typically in either a conjunction or feature search, when the target

is absent, an increasing search function is exhibited (Adler & Orprecio, 2006; Wolfe, 1989). In the current study, however, when the target was absent, search was represented by a significant decreasing function. This significant and highly negative slope occurred for both the overall results, and for the second half of the trials. Research with adults have shown decreasing search functions with increasing set sizes (Bravo & Nakayama, 1992; Bacon & Egeth, 1991; Humphreys, Quinlan, & Riddoch, 1989; Humphreys, Riddoch, & Quinlan, 1985) A possible explanation for the decreasing search function may be related to grouping effects, which are more likely to occur as the display size increases (Bacon & Egeth, 1991). Grouping effects occur when items in a search array are visually grouped together either by similarity to one another or proximity to one another (Pomerantz, 1981). Thus, the two factors that generally contribute to grouping of items in the search array are similarity of target and distractors and the proximity distractors to each other and distractors to the target (Pomerantz, 1981). Humphreys et al. (1985) showed that when they presented observers a small circle search array comprised of a target of an upright T with similar distractors (inverted T's and L's), the number of distractors had little effect on RT when a target was present, and when the target was absent arrays were faster than when the target was present across and faster as set size increased. They attributed these quick target-absent responses to preattentive processing being facilitated by grouping combinations of similar homogenous distractor. In other words, as the display size increased, smaller groupings increased within the display size, and when distractors are more homogenous or similar to one another, distractor-distractor grouping is stronger than target-distractor-grouping (Humphreys et al., 1985). In the current study, however, the distractors were not similar to one another. For instance, if the target was a green X, the distractors would be a red X or a green O. The more

likely factor that might affect grouping in the current study, therefore, would be the proximity of similar distractors to one another.

Bacon and Egeth (1991) claim that in a simple feature task, a distractor must be within a close distance to the target in order for the contrast of the target to be seen. Sagi and Julesz (1987) also showed RT's decreasing from a set size of 2 to 30 because in the smaller set sizes, the items were farther than  $2^\circ$  apart. Other studies have shown that when same color distractors are closer to other same color distractors than to the target, the RT's decrease as the number of distractors are added (Bundesen & Pedersen, 1983). This type of grouping has to do with both contrast and proximity, in that distractors and targets in close proximity need have at least one feature (shape, color, orientation) starkly different from other surrounding items. If many of the same distractors are close together and the target is at a distinct location in the array far away from the distractors, then the distance/location contrast of the target relative to the distractors would promote pop-out. This may have been the case with the current study. Here the distractor placement occurred at random, so it's possible that similar distractors were randomly placed at one side of the array while an odd distractor was not within proximity to similar ones allowing that distractor to perhaps pop out from other distractors.

If this explanation of the decreasing search function is accurate then why did decreasing target-absent latencies with set size only occur with the top-down identity cue, whereas latencies in the target-absent trials after bottom-up spatial cues remained relatively flat with set size. Obviously then, the account provided to explain the decreasing search function when targets were absent must somehow be related to the attentional mechanisms involved, namely, top-down mechanisms. An explanation can again be related back to Wolfe's (1989) Guided Search theory in that because infants top-down activation maps were biasing their search, when the target was



absent, either color (red or green) or shape (X or O) activation maps guided their search to a distractor that best fit one of these features. In the bottom-up cueing condition, infants' top-down biasing for activation maps was not enabled, thus allowing them to merely allocate attention to any one of the distractors resulting in the flat search function. Now which distractor feature (color or shape) activated their top-down guidance is unclear at this point. Further research would be required to investigate which top-down feature map (color or shape) biased their attention allocation to. In other words, if infants' first eye movement went to a distractor the same color or shape of the target that they were searching for.

#### *Relation to Infants' Visual Expectations*

The results of this study showed that infants are developmentally mature in their bottom-up attentional mechanisms but relatively immature in their top-down attentional mechanisms. Immaturity of processing was interpreted on the basis of which cue better facilitated infants' search performance to reflect typical adults' performance in a conjunction search. This developmental progression in infants' top-down abilities, however, might be specific to visual search tasks. Other infant research that has investigated their abilities during a visual expectation task would seem to suggest that 3-month-old infants do possess top-down attentional mechanisms (Adler & Haith, 2003; Adler & Haith, 1998; Haith, 1988). For instance, Adler and Haith (2003) presented infants with two types of pictures alternating in a temporally consistent left-right sequence. On one side (either the left or the right) the pictures were always the same combination of colors in one experiment, or combination of shapes in another experiment, (invariant side) whereas on the other side the color or shape varied. Results of these experiments showed that infants made significantly more anticipatory eye movements to the invariant side than to the side that varied in color or shape combinations. In other words, infants were

expecting the constancy of the invariant picture to appear on its corresponding side, so they made an eye movement to that location prior to the picture onset. These results suggest that infants are able to form expectations of where an item will appear in their environment and then use these expectations in a top down manner to modulate their attentional allocation and eye movements. The expectation findings would seem to reveal that infants do not merely react to stimuli in their environment in a bottom-up fashion, but are able to utilize top-down control of their attentional allocation (Haith, 1993).

If this is the case that expectation studies indicate a functioning top-down attentional mechanisms but visual search studies indicate a later developing mechanism, then what might be the difference between top-down attentional control in visual search processing compared to visual expectations? Current research does not have an answer for this problem. Based on the current study's results compared to visual expectation studies with infants, one might hypothesize that two distinct top-down attentional mechanisms exist, one for search and one for expectancy. Or, alternatively, perhaps the difference is not in their attentional mechanisms but instead related to other mechanisms such as working memory or even encoding processes. Further investigation that combines these two tasks would aide in answering this question.

#### *Similarity to Autistic's Visual Search*

Understanding how selective attentional processes work in infancy is critical for increasing our knowledge of how these processes develop. It is critical because infants can only remember and learn items in their environment that they attend to. Thus, selective attention provides a window into how we begin to construct a knowledge base of the world around us. Additionally, such knowledge would be especially critical to understanding the attentional processes do not develop properly – for instance in attentional disorders such as ADHD or ASD

(Autism Spectrum Disorders). Visual search research conducted on toddlers, school-aged children, and adults with autism have shown that they are quicker at finding targets in conjunction searches than typical age-matched controls (Kaldy, Kraper, Carter, & Blaser, 2011; Plaisted, O’Riordan, & Baron-Cohen, 1998; O’Riordan, 2004). Kaldy et al. (2011), for example, presented a conjunction task with three set sizes (5, 9, & 13) to 2.5-year-old toddlers with autism and compared their performance with age matched normal controls. Similar to the current study, they showed toddlers an identity cue of the target (a red apple) prior to the onset of search arrays that varied in set size and distractor location. An eye tracker was used to compare gaze fixations to target and distractors and recorded the percentage of correct trials that toddlers gaze localized the target.

Results of this study showed that ASD toddlers were more successful at finding the target across all set sizes than the controls (Kaldy et al., 2011). Reaction time or saccade latencies were not measured in this study because trials were on for an extensive fixed amount of time, 4 seconds. Interestingly though, in the 4 seconds, ASD children fixated almost twice as many items in the array than the control toddlers, and that as set size increased ASD toddlers fixated on more items while typical controls fixated on the same number of items regardless of set size. This suggests that ASD toddlers are perhaps better at finding the target than controls because they are able to disengage from an item in the display quicker than typical functioning toddlers which allows them to continue searching for the target (Kaldy et al., 2011). These results could also mean that ASD toddlers’ bottom-up attentional mechanisms are not developed and as such they do not process items efficiently in parallel. Instead, because they are lacking the mechanisms to processes items efficiently, their serial search mechanisms have developed to be more efficient in locate items in their environment (Kaldy et al., 2011). In other words, they

have developed the ability to engage and disengage from each item quickly so that they can serially move their search from one item to the next. Or, perhaps the top-down attentional mechanisms that seem to show later development in infants and are likely responsible for infants showing an immature conjunction search are the same attentional mechanisms that are deficient in individuals with ASD and responsible for these individuals not showing typical conjunction search behavior. Regardless, more research comparing individuals with autism to developing infants would be ideal in understanding the mechanisms involved in visual search and how they develop or lack thereof.

### *Conclusion*

In summary, the current study is the first of its kind to show that 3-month-old infants' top-down processing mechanisms are developmentally immature compared to their bottom-up mechanisms. This makes sense in that bottom-up processes are considered more primitive than top-down processes; as a result we would expect that infants would perform better in a feature search task than a conjunction one. This differential functional maturity of the attentional mechanisms was demonstrated by showing that 3-month-old infant's only showed linearly increasing search functions when they were provided with a top-down cue and only after they had time to implicitly learn and remember the cue. An increasing linear search function is typical in adult conjunction searches and is an indication of a fully developed selective attention system. In the current study, however, search functions did not increase when infants were provided with bottom-up spatial cues overall search performance was faster compared to infants in a previous study that were not given any cues (Fuda & Adler, 2012). These combined experimental results show that infants are developmentally mature in their bottom-up processing, but immature in their top-down processing abilities, and as such the top-down cue provided the

facilitation needed to search, detect, and localize a target in a conjunction search in a more adult-like manner.

So what does differential developmental maturity of bottom-up and top-down attentional mechanisms imply for infants' processing stimuli naturally in their environment? Our environment is constantly filled with potential targets that we can attend to, and even a mature attentional system is not equipped to process all possible objects and events in our environment simultaneously. Thus selective attention allows us to choose and process individual items in our environment. The current study's results suggest that highly contrasting unique items in the environment will likely capture infants' immature attentional system through bottom-up mechanisms more so than highly complex stimuli, thus allowing them to build on their growing knowledge base for their visual world. This is not suggesting that complex environmental stimuli will not engage their learning processes of their environment at all, but instead suggests that perhaps at this early age infants are only able to attend to unique high contrasting items that are the source of the functional bottom-up attentional mechanism. Perhaps once these more basic unique contrasting items have been attended to and processed they will feed into and assist more complex and top-down attentional processing mechanisms to develop.

Further research, however, is still necessary to completely understand the nature of the development of top-down attentional mechanisms. For instance, adult research has suggested that not all top-down information is equivalent in their capacity to bias search. For instance Sobel et al. (2009) showed that color cues more so than orientation cues better facilitated conjunction searches. Currently it is unknown which aspect of the top-down identity cue facilitated infants' search performance better (color or shape of the target). Thus, in one possible future direction for research, there would be value in investigating whether infants only require

color or shape or a combination of both identity features to aide in their search performance. The current research opens up the possibility of investigating many different avenues of inquiry that will not only illuminate more brightly the development of selective but will also likely shine the light on where that development goes wrong.

## References

- Adler, S. A. (2005). Visual search and pop-out in infants. In *Neurobiology of Attention*, (Eds.) L. Itti, G. Rees & J. K. Tsotsos. London: Elsevier Academic Press.
- Adler, S. A., & Bala, J., & Krauzlis, R. J. (2002). Primacy of spatial information in guiding target selection for pursuit and saccade. *Journal of Vision*, 2, 627-644.
- Adler, S. A., & Haith, M. M. (2003). The nature of infants' visual expectations for event content. *Infancy*, 4, 389-421.
- Adler, S. A., Inslicht, S., Rovee-Collier, C., & Gerhardstein, P. C. (1998). Perceptual asymmetry and memory retrieval in 3-month-old infants. *Infant Behavior & Development*, 21(2), 253-272.
- Adler, S. A., & Opreco, J. (2006). The eyes have it: Visual pop-out in infants and adults. *Developmental Science*, 9, 189-206.
- Atkinson, J., & Braddick, O. (1992). Visual segmentation of oriented texture by infants. *Behavioural Brain Research*, 49, 123-131.
- Bacon, W. F., & Egeth, H. E. (1991). Local processes in preattentive feature detection. *Journal of Experimental Psychology: Human Perception and Performance*, 17(1), 77-90.
- Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics*, 51(5), 465-472.
- Bhatt, R. S., Bertin, E., & Gilbert, J. (1999). Discrepancy detection and developmental changes in attentional engagement in infancy. *Infant Behavior & Development*, 22, 197-219.
- Bundesen, C., & Pedersen, L. F. (1983). Color segregation in visual search. *Perception & Psychophysics*, 33, 487-493.

- Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development*, 62, 1-145.
- Colombo, J., Ryther, J. S., Frick, J. E., & Gifford, J. J. (1995). Visual pop-out in infants: evidence for preattentive search in 3- and 4-month-olds. *Psychonomic Bulletin and Review*, 2, 266-268.
- Fuda, C., & Adler, S. A. (2012). *Conjunction Search in Infants and Adults: An Eye Movement Study*. Poster session presented at Vision Sciences Society, Naples FL.
- Gerhardstein, P., & Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *Journal of Experimental Child Psychology*, 81, 194-215.
- Gitelman, D. R. (2002). ILAB: a program for post experimental eye movement analysis. *Behavioral Research Methods, Instruments and Computers*, 34(4), 605-612.
- Haith, M. M. (1993). Future-oriented processes in infancy: The case of visual expectations. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development*, 59, 467-479.
- Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics*, 57, 787-795.
- Hood, B. M. (1993). Inhibition of return produced by covert shifts of visual attention in 6-month-old infants. *Infant Behaviour and Development*, 16, 245-254.
- Humphreys, G. W., Quinlan P. T., & Riddoch, M. J. (1989). Grouping processes in visual



- search: effects with single- and combined-feature targets. *Journal of Experimental Psychology: General*, 118(3), 258-279.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1985). Interactive processes in perceptual organization: Evidence from visual agnosia. In M. I. Posner & O. S. M. Marin (Eds.) *Attention and performance XI* (pp.301-318). Hillsdale, NJ: Erlbaum.
- Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency, learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, 3(4), 335-344.
- Julesz, B. (1981). A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 41, 131-138.
- Kaldy, Z., Kraper, C., Carter, A. S., & Blaser, E. (2011). Toddlers with Autism Spectrum Disorder are more successful at visual search than typically developing toddlers. *Developmental Science*, 14(5), 980-988.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35, 1897-1916.
- Maljkovic, V., & Nakayama, K. (1994). Priming of popout: I. Role of features. *Memory and Cognition*, 22, 657-672.
- Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, 320(6059), 264-265.
- Neisser, U. (1966). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry*, 39, 777-783.

- Pomerantz, J. (1981). Perceptual organization in information processing. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 141-180). Hillsdale NJ: Erlbaum.
- Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.
- Reddy, L., & VanRullen, R. (2007). Spacing affects some but not all visual searches: Implications for theories of attention and crowding. *Journal of Vision*, 7(2), 1-17.
- Rovee-Collier, C., Hankins, E., & Bhatt, R. (1992). Textons, visual pop-out effects, and object recognition in infancy. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1178-1187.
- Sagi, D., & Julesz, B. (1987). Short-range limitation on detection of feature differences. *Spatial Vision*, 2, 39-49.
- Salapatek, P. (1975). Pattern perception in early infancy. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (Vol. 1, pp. 133-248). New York: Academic Press.
- Shea, S. L. (1992). Eye movements: Developmental aspects. In E. Chekaluk and K. R. Llewellyn (Eds.). *The role of eye movement in perceptual processes* (pp. 239-306). Elsevier.
- Sireteanu, R., & Reith, C. (1992). Texture segregation in infants and children. *Behavioural Brain Research*, 49, 133-139.
- Sireteanu, R., Rettenbach, R., & Wagner, M. (2009). Transient preferences for repetitive visual stimuli in human infancy. *Vision Research*, 49(19), 2344-2352.
- Sobel, K. V., Pickard, M. D., & Acklin, W. T. (2009). Using feature preview to

- investigate the roles of top-down and bottom-up processing in conjunction search. *Acta Psychologica*, 132, 22-30.
- Theeuwes, J., Van Der Berg, E., & Belopolsky, A. (2008). Detecting the presence of a singleton involves focal attention. *Psychonomic Bulletin & Review*, 15, 555-560.
- Theeuwes, J. (1991). Exogenous and endogenous control of attention – The effect of visual onsets and offsets. *Perception & Psychophysics*, 49, 83-90.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, 135, 77-99.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-106.
- Wolfe, J. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, 1, 202-238.
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology*, 15, 419-433.
- Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, 44, 1411-1426.