



ELSEVIER

Contents lists available at SciVerse ScienceDirect

## Cognition

journal homepage: [www.elsevier.com/locate/COGNIT](http://www.elsevier.com/locate/COGNIT)

## Predictable locations aid early object name learning

Viridiana L. Benitez\*, Linda B. Smith

Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47405, United States

### ARTICLE INFO

#### Article history:

Received 4 August 2011

Revised 26 July 2012

Accepted 17 August 2012

Available online 16 September 2012

#### Keywords:

Attention

Development

Infancy

Word learning

### ABSTRACT

Expectancy-based localized attention has been shown to promote the formation and retrieval of multisensory memories in adults. Three experiments show that these processes also characterize attention and learning in 16- to 18-month old infants and, moreover, that these processes may play a critical role in supporting early object name learning. The three experiments show that infants learn names for objects when those objects have predictable rather than varied locations, that infants who anticipate the location of named objects better learn those object names, and that infants integrate experiences that are separated in time but share a common location. Taken together, these results suggest that localized attention, cued attention, and spatial indexing are an inter-related set of processes in young children that aid in the early building of coherent object representations. The relevance of the experimental results and spatial attention for everyday word learning are discussed.

© 2012 Elsevier B.V. All rights reserved.

### 1. Introduction

The tie between a name and the thing named transcends time and space: A particular object that is a *ball* when on the table will still be a *ball* when seen later on the floor. Given this, one might predict that the best way to teach children names for things would be to expose them to events in which the referent was named in varying locations. However, core properties of memory and attention use the location of objects: (1) to bind elements into unified representations (e.g., Treisman, 1986; Treisman & Gelade, 1980); (2) to guide and localize attention (Eriksen & Rohrbaugh, 1970; Posner, Snyder, & Davidson, 1980); and (3) to index memories (Richardson & Spivey, 2000). If these core processes play a critical role in word learning, then naming a novel object at a consistent location – rather than at varying ones – should promote the learning of that object name. Three experiments were designed to test a “spatial consistency” hypothesis in 16- to 18-month old infants: the spatial predictability of an object’s location fosters the learning of an object’s name and in so doing,

fosters the spatial and temporal transcendence of the bond between the name and the thing.

#### 1.1. Space, attention, and memory for objects

Posner (1980) characterized attention as a spatial spotlight that enhanced processing at a location. Treisman (1986, 1988) demonstrated that localized visual attention was central to binding visual features into a unified object representation. Although contemporary discussions have refined these ideas (and the role of objects in defining locations), the spatial nature of attention and its role in detecting, discriminating, and representing objects is well documented (e.g., Driver, 2001; Greenberg & Gmeindl, 2008; Scholl, 2001; Wright & Ward, 2008). Additional findings indicate the benefit of directing attention to a location via a top-down cue that generates an expectation as to where a relevant object will appear (see Gilbert & Sigman, 2007, for review). The result is this: an arrow that points to a location leads to faster reaction time to detect the visual target at that location in comparison to when the target’s location is uncued or invalid. However, a salient flicker of light near the target location – a cue that presumably draws attention to the location through low level and

\* Corresponding author. Tel.: +1 812 855 8256.

E-mail address: [vlbenite@indiana.edu](mailto:vlbenite@indiana.edu) (V.L. Benitez).

involuntary processes – does not lead to more rapid detection of the target (Gilbert & Sigman, 2007; Kerzel, Zarian, & Souto, 2009). The arrow is typically characterized as an “endogenous” cue that drives attention by expectancies, whereas the flicker is characterized as an “exogenous” or stimulus-driven cue (see Jonides, 1981; Posner, 1980). Such expectancy-based attention has been shown to result in enhanced discrimination as well as detection (Fischer & Whitney, 2009; Lange & Röder, 2006).

If the world presents consistent cues and if people are good at learning them, then the top-down cueing of attention to a location may play an important role in many cognitive tasks. Consistent with this idea, the evidence suggests that adults automatically encode the spatial location of visual events and do so with precision (e.g., Hasher & Zacks, 1979; Hommel, 2002; Nissen, 1985). Adults also appear to automatically learn predictive relations between irrelevant scene elements and the location of task relevant information and then use those background cues to more rapidly find target information in a cluttered array (Brockmole & Henderson, 2006; Chun & Jiang, 1998; Hollingworth, 2009). Finally, other work shows that learned cues that spatially direct attention are not just critical to the visual processing of objects but also in forming multimodal memories; more specifically, expectancy-based attention to a location has been shown to support the binding of sensory elements from different modalities, such as a sight and a sound (e.g., Fiebelkorn, Foxe, & Molholm, 2010; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010).

These findings implicate memories for the locations of attentionally-relevant events, and not just memories for the events themselves. Consistent with the more specific proposal of “spatially-indexed” memories are findings from studies in which adults were presented with an array of spatially organized information to study for later testing (Richardson & Kirkham, 2004; Richardson & Spivey, 2000). At test, while looking only at a blank screen, participants were asked questions about the previously presented information. The central finding is that when asked about specific information, adults consistently looked to the location where that information *had been* even though there was no longer any information there, a finding that is sometimes referred to as the “looking at nothing effect” (Richardson & Spivey, 2000). Although there are several accounts of this phenomenon that differ in their details (Ferreira, Apel, & Henderson, 2008; Richardson, Altmann, Spivey, & Hoover, 2009) all agree that there is a link between spatially directed eye-movements and memories for visual events. Critically, as Ferreira et al. (2008) point out, this “looking at nothing” phenomena also involves *words* as the cues that shift eye-gaze to the remembered locations, indicating that words that activate memories of objects also activate the recent location of those objects, a key idea for the present experiments. Ferreira et al. suggest that this word-localized attention effect reflects the fundamental nature of online word comprehension: words automatically direct looking to the location (or remembered location) of a mentioned object (see also, Altmann, 2004; Knoeferle & Crocker, 2007). In their review of this literature, Ferreira et al. conclude that attended objects and the locations of those objects form integrated memories,

such that the whole memory may be activated by any part—by the visual information, by linguistic input, and potentially by the direction of eye gaze itself.

In sum, there appears to be a spatial attention and memory system in which (1) localized visual attention fosters the binding of co-occurring events into a unified representation; (2) expectancy-based attention to a location supports deeper processing and the binding of multisensory events; (3) the locations of objects are parts of object memories; and (4) words that activate these object memories also activate location memories.

### 1.2. Developmental findings

There is limited direct evidence on these processes in infants and children but the evidence that does exist suggests that spatially localized attention plays a similar role to that observed in adults by late infancy. Quite young infants (3–4 months of age) register location information and use it to organize attention to objects, making anticipatory eye movements to the location at which an object will appear (Canfield & Haith, 1991; Haith, Hazan, & Goodman, 1988; Johnson, Posner, & Rothbart, 1991; Johnson & Tucker, 1996). The beneficial effects of expectancy-based cues in forming multi-sensory memories has been demonstrated in 8 month olds: specifically, infants were found to bind co-occurring sights and sounds only when the location of the multisensory event was cued by an expectancy-based signal but not when it was cued by salience (Wu & Kirkham, 2010). At 6 months, infants show the “looking at nothing” effect in simple object memory tasks (Richardson & Kirkham, 2004) and preschoolers, who can be tested in the same tasks as adults, show clear and adult-like spatial indexing of object memories (Martarelli & Mast, 2011). Also consistent with the idea of spatially indexed memories, one recent study found that 12-month-olds better understood adult talk about a not-present object when the referred-to object had been previously experienced in a constant and predictable spatial location rather than variable locations (Osina, Saylor, & Ganea, 2012). Finally, in a related study, spatial predictability has been shown to facilitate 16-month-olds’ linking of a name to an absent but previously experienced object (Samuelson, Smith, Perry, & Spencer, 2011). In sum, it seems likely that for infants and young children, as for adults, spatially localized attention supports forming and retrieving integrated and multisensory object memories. If this is so, then spatially localized attention might be expected to be a critical part of the mechanisms that support early object name learning: object names might be better learned when the referents are named at constant rather than varied locations.

### 1.3. Object name learning from the perspective of spatial attention and memory

For young children object name learning begins as a kind of on-line word comprehension task: That is, children learn their first words by hearing object names in the context of a scene with potential referents located within that scene. A large literature suggests that social (Baldwin,

1993) and linguistic cues (see Smith, Colunga, & Yoshida, 2010, for a review) play a key role in organizing visual attention to the intended referent and, moreover, that attention directed by such meaningful cues may be essential to the mapping of a name to an object (Waxman & Gelman, 2009). These facts are not usually thought of in terms of a role for localized spatial attention or expectancy-based attention in binding names to things (but see Smith et al., 2010; Wu & Kirkham, 2010). However, processes of spatial attention and memory as revealed in the adult literature could provide the mechanistic underpinning for these and other attention-related phenomena in early object name learning.

For example, object naming as typically experienced by young children contains many potential cues to support expectancy-based spatially-directed attention. Meaningful cueing events (points, looks, the object name, other words) generate attentional expectations that localize attention to specific objects in the scene, and in so doing may set up internal processes critical to binding the visual elements of the objects to each other and to the object name. Partners in a conversation (Clark & Wilkes-Gibbs, 1986; Hutchins, 1995) and perhaps also parents interacting with toddlers (Samuelson et al., 2011) often set up a spatial frame in which multiple referents are stably located with respect to one another. Establishing a spatially organized common ground, along with social cues that localize the referent, may matter to early object name learning because such spatial consistency supports location-based cognitive processes. These are the larger ideas that motivate the present study. However, the critical first question, one that has not been experimentally tested, is whether spatial consistency matters at all for the learning of object names by young children.

#### 1.4. The hypotheses

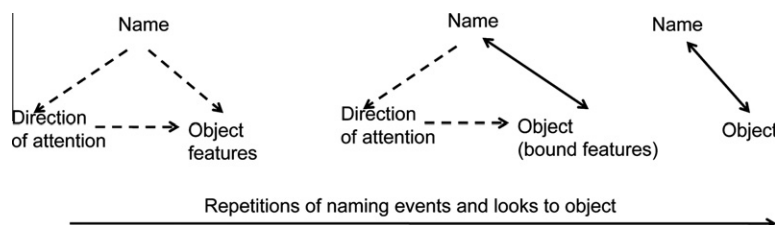
Fig. 1 illustrates a collective set of hypotheses about how spatial consistency matters to object name learning: (1) when an object is attended to, its location relative to the perceiver is stored with the object and when an object is named, the name is associated with both the spatial direction of attention and the object. (2) Over repeated naming events within a conversational context, these lead to integrated memories such that the name cues a direction of attention and the expectation of the specific referent; the direction of attention itself may also serve as a

cue for the memory of the object and the name. These expectancy-driven and localized attentional processes also lead to deeper processing of the object, the binding of its individual elements, and the multisensory binding of the name to the object. (3) In the end, this mutually reinforcing system builds a strong link between a name and an object that transcends the local context and thus is recoverable even when the object and name are presented in a new location. If these ideas are generally correct, then spatial predictability should support object name learning by young children.

The experiments that follow test this prediction and also other elements of the set of hypotheses illustrated in Fig. 1. To these ends, the experiments *do not* examine object name learning in a naturalistic context but rather in an experimental task created so that spatial consistency could be manipulated and expectations about where objects were located could be measured. Experiment 1 was designed to test the main hypothesis about spatial consistency: if spatial attention matters for object name learning, then naming individual objects at distinct and predictable locations rather than at varied locations should support the binding of the name to the object, and thus the formation of an object-name bond that ultimately is not tied to a single location. Experiment 2 was designed to examine the role of the object name as a cue to location. If expectancy-based attention – and not just salience-driven, exogenous attention – supports the binding of a name to a thing, then, within individual infants, the degree to which the name on its own generates an expectation about a location during the learning phase should be a strong predictor of whether the object name was learned when tested at a new location. Experiment 3 tests the idea of location as a memorial index through which new information is combined with information in memory: If the direction of spatial attention serves as a memory index that can connect current to previous experiences so as to build strong memory representations that integrate information over multiple experiences, then objects and names that are presented with attention directed to the same location – but at different times – should nonetheless be integrated into a single memory representation.

## 2. Experiment 1

Experiment 1 tests the hypothesis that spatial consistency across multiple naming events for a single object



**Fig. 1.** Hypothesis about how predictable locations may aid object name learning: an object-naming event involves the association of both the object features and a direction of attention with that name. After several repetitions of a name with an object at a specific location, the label may cue a direction of attention and enhance object processing, and the direction of attention may also cue the name and object features. Over time, these memories become integrated, and the result is a name being bound to an object.

supports the binding of the name to that thing by comparing two training conditions: In the Constant-Location condition three unique objects are repeatedly named – each with its own unique name and each at its own unique and distinct location. In the Varied-Location condition, the three objects are repeatedly named with their own individual names but at varied and overlapping locations. The experimental task was designed to be engaging to young children, to naturally orient attention to objects at specific locations through a highly salient and attention-grabbing event, and to allow for the association of object names to objects at constant or varied locations. For these purposes, we constructed the pop-up box shown in Fig. 2a. When activated by a lever, an object rapidly popped up at one of three locations. The event consisted of a popping sound and the rapid visual appearance of the object, both strong exogenous cues that naturally attract attention. In both the Constant- and Varied-Location conditions, infants were presented – via the pop-up box – with three unique objects, each with its own unique name. Each exposure to an object and its name consisted of three critical components: (1) the *pre-pop-up name cue* that an object was about to appear, which was the unique name of the object and thus – given some repetitions – could lead in both conditions to expectations about *what* should appear; (2) the *salient pop-up event* which should attract attention to the specific location and object; and (3) the *naming event* during which the object was pointed to and explicitly named by the experimenter. If all that matters to infants' learning an object name is the clear and explicit indication of the intended referent for that name and if the history of locations of the naming events do not matter, then infants in both conditions should perform comparably in learning the object names since the two conditions do not differ in the experimenter's explicit naming of the objects. The Constant- and Varied-Location conditions differ only in whether each unique object is presented and named at a constant location or a varied location across trials. When objects are presented at a constant location across trials, the name – said prior to the pop-up – cues not only the object that is about to appear but also its *location*. Thus, the two conditions differ only in the spatial predictability of the named object, in the possibility of expectancy-based spatial cueing, and in common spatial indexing of the repeated naming events for each individual object.

After the training of objects and their names via the pop-up box, the objects were removed from the pop-up box, presented in a novel spatial arrangement and location, and infants' knowledge of the object names was tested.

Experiment 1 thus directly tested the hypothesis that spatial predictability supports the formation of a bond between the name and the object, a bond that can become sufficiently robust so as to be independent of location.

The objects used in the experiment, as shown in Fig. 2b, were highly similar to each other in surface properties and colors. Infants' ability to attach such similar objects to unique names provides a measure of the degree to which infants were binding object features together to form robust representations of the individual objects. We assessed this by two kinds of test trials: Easy and Hard. On Easy test trials (Fig. 2c), when the infant was asked to indicate a named object, the infant only had to distinguish the target from two very different objects (which were never popped up during the learning phase). A right answer on these trials does not require a highly specific memory for the specific object features that were tied to a specific name. On the Hard test trials, the infant was asked to indicate the named target with the foils being the other two highly similar objects that had been associated with other names during the pop-up trials (Fig. 2b). Success on these trials required finer grained associations of object properties and the binding of these specific properties to specific names.

## 2.1. Method

### 2.1.1. Participants

Thirty-two infants (half female) ranging from 16.2 months of age to 18.2 months of age ( $M = 17.4$ ,  $SD = 0.59$ ) participated. All infants were recruited from a working and middle-class population in Bloomington, Indiana, and received a toy for participating in the experiment. Three infants were recruited but did not complete the task. Infants were randomly assigned to two training conditions: Constant- and Varied-Location.

### 2.1.2. Stimuli

The training stimuli consisted of three novel pseudo-words, *riff*, *dax*, *zup* and three novel, hand-made objects (Fig. 2b). The objects were unique in overall shape and texture, but were each composed of many colors and component shape features that overlapped. Two additional objects that shared none of these features were also used on the Easy test trials (Fig. 2c). The objects measured approximately  $7.6 \times 5 \times 10.2$  cm in size, had Velcro patches on the bottom, and could be fit into and attached to the pop-up platforms as described below. The names were randomly assigned to the three training objects. Finally, three familiar objects – a spoon, a cup, and a flower



**Fig. 2.** Pop up box (a) and the popped up objects (b) used in the experiments. The Hard test trials (b) consisted of choices among the three popped-up objects. The Easy test trials (c) consisted of the popped-up target and two novel objects not used in the training phase.

– were used for familiarization with the testing procedure (but were not used during the pop-up training).

### 2.1.3. Apparatus

A pop-up box was used to present the objects during training (Fig. 2a). This was a rectangular, hollow, wooden box (99.4 × 17 × 13.5 cm) that contained three openings (22 cm. apart) from which the objects popped up. Three springs were built inside the box and contained platforms where the objects could be attached with Velcro. There were three levers that controlled the pop-up of each location. When a lever was pushed, the lid of the respective location popped open and the platform on which the object was attached sprung up to reveal the object to the infant. After a trial, the lid was pushed down to hide the object again.

The three levers that controlled the pop-ups were located on the back of the box (which always faced the experimenter), with all three levers close together and just slightly to the right of center; in this way, the experimenter's hand movements could provide no reliable spatial cue as to the pop-up location. In addition, there were two openings on the back of the box (facing the experimenter) through which the experimenter could, unseen by the infant, move objects from one platform to another between trials. For both the Constant- and Varied-Location conditions, all objects were taken off their platforms between trials and replaced either to the same location (in the Constant-Location condition) or to a different location (in the Varied-Location condition).

### 2.1.4. Procedure and design

There were two pop-up training conditions, between subjects, that were identical with the exception as to whether the location of an object was the same each time it was presented or whether it varied. We first describe the procedure for the Constant-Location condition and then the differences that characterize the Varied-Location condition.

Prior to the pop-up task, the infant was given the three training objects and the two additional testing foils to manually and visually explore for approximately 2 min or until interest in those objects waned (this was done to limit attempts by the infant to reach to the objects during the series of pop-up trials and to familiarize infants with the novel foils). After familiarization with the objects, the pop-up box was placed on a table 68.5 cm in front of the infant. The infant sat on the parent's lap. Parents wore ear buds and were instructed to keep their eyes closed throughout the entire experiment so as not to interfere with the infant's behavior. Trials were presented in blocks of 3, such that each of the three unique objects was presented once and named at its location once before any object was repeated. Each trial, a single pop-up event, consisted of (1) the pre-pop-up name cue; (2) the salient pop-up event and (3) the explicit naming event. The pre-pop-up name cue event began with the experimenter saying, "Look! It's a *riff*." While speaking and during the whole pre-pop-up name cue phase of a trial, the experimenter looked directly into the subject's eyes and not at any pop-up location. After saying the name, the experimenter

waited approximately 1 s., and then popped up the named object. After the pop-up, the experimenter explicitly named the popped-up object, saying, "Look at the *riff*. See the *riff*" while looking at the object, at the infant, and back to the object, providing the social cues typical of an explicit naming event. The experimenter then pushed the object back into the box and proceeded to the next trial. The experimenter exposed infants to all three objects in this way for each block of three trials (one pop-up for each unique object). After every such block of three pop-ups, the experimenter handed the infant one of three familiar objects to examine (cup, flower, or spoon) so as to occupy their attention while the experimenter, from the back of the pop-up box (and thus not viewable by the infant) *re-moved and, in the Constant-Location condition, re-attached the objects in the same locations* – an activity that took 1–2 s. When this was accomplished, the experimenter requested the familiar object back from the infant, and then the next block of three pop-ups was presented. This procedure was repeated for six blocks of three trials each, creating a total of 18 pop-ups – six cueing events, six pop-ups and six explicit naming events for each object. The order of popped-up location (and thus the named object) within a trial was varied (by Latin Square) across trials. The 18 pop-up trials were immediately followed by the testing trials.

The procedure in the Varied-Location condition was identical to the Constant-Location condition except for the locations of the objects, which were changed between each block. By the end of the 18 trials, each object in the Varied-Location condition had been presented twice at each location. The order of popped-up location was varied (by Latin Square) across trials and the orders of pop-up locations within and across trials was matched – participant by participant – to those in the Constant-Location condition. Each unique object – both in the Constant- and Varied-Location conditions – was always cued with the same name and the popped up object was always explicitly named with the same unique name for each individual infant. For testing, the pop-up box was removed from the table and out of infants' sight.

### 2.1.5. Warm up trials

There were two warm-up trials with the familiar warm-up objects (flower, spoon, and cup) before testing with the trained objects. For these warm-up trials, the three known objects were presented in a shallow triangle-shaped array on a tray at the infant's mid point such that each object was 15.2 cm from the other two objects and all objects were roughly 12 cm from the midpoint of the infant's chest when the tray was pushed forward for the infant to respond. The test trial began with the tray close to the experimenter and away from the infant; with the tray so located, the experimenter looked directly into the infant's eyes and not at the objects and asked for one object by name, for example, "Where is the flower? Get the flower." Then the tray was moved forward on the table so that all objects were within the infant's reach. If the infant did not immediately select the named object on these warm-up trials, the infant was coached into doing so. There were two such warm up trials, each requesting a different object.

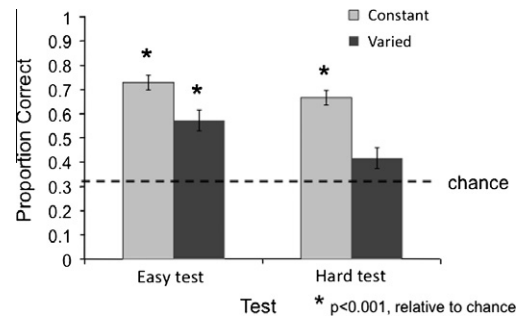
### 2.1.6. Testing trials

The testing trials were similarly structured to the warm up trials, except that there was no coaching. In addition, no feedback was given and the experimenter simply said “Thank you,” when the infant indicated or handed a toy over. There were two kinds of test trials, Easy and Hard. On the Easy test trials, the named target was presented with two very different foils that had not been used during the pop-up trials. On the Hard test trials, the named target was presented with the two other training exemplars as distractors; these test trials should be difficult if infants have trouble binding the complex features of the objects into unified representations. The spatial locations of the objects were randomly determined on each test trial and were not aligned in any way with their experienced spatial locations within the pop-up box. All three trained objects were tested twice in each of these two formats (yielding a total of six Easy and six Hard test trials). Test trials were randomly ordered with two constraints: Hard and Easy test trials alternated and the named target could not be the same on two successive test trials. The entire experiment lasted approximately 20 min.

The entire session was video-recorded. Choices during test were coded as correct or incorrect. Looking behavior during training was coded using MacShapa (Sanderson et al., 1994) and consisted of: (1) looking anticipations – that is, latencies to look to the popped up object measured from the pre-pop-up naming cue to the pop-up; (2) looking latencies – measured from the moment of pop-up to the first look to the popped-up object, and (3) looking duration to the popped up object during the explicit naming event, specifically, the total duration of looking at the popped-up object from pop-up to being pushed down out of view. One coder coded all data and then a second coder coded a randomly selected six participants. Agreement for first look to a location given the pre-pop-up name cue was .96 and for choice on test trials was .98. For the latency and duration measures, agreement was assessed by examining the timing differences between the two coders. More specifically, latencies and durations within 500 ms of each other were said to be in agreement. Agreement for looking latencies was 1.0 and for looking duration was .86. The correlation between the two coders’ judged latency was  $r = 0.85$  ( $p < 0.001$ ) and between the two coders’ judged durations was  $r = 0.98$  ( $p < 0.001$ ).

## 2.2. Results and discussion

A key prediction from the hypothesis set in Fig. 1 about the spatial nature of core memorial, attentional and cognitive processes is that spatial consistency should support learning. Thus, the central empirical question is whether infants learned the object-name mappings better in the Constant- than in the Varied-Location condition. Fig. 3 shows the main result: consistent with the prediction, infants showed better knowledge of the object-name mappings in the Constant- ( $M = 0.70$ ,  $SD = 0.16$ ) than Varied-Location ( $M = 0.50$ ,  $SD = 0.12$ ) condition. They showed this better knowledge even though tests of name comprehension were conducted in a new location and without spatial cues. The Easy trials ( $M = 0.65$ ,  $SD = 0.09$ ) were easier for



**Fig. 3.** Mean proportion correct (and standard error of the mean) for Easy and Hard test trials for the Constant- and Varied-Location conditions in Experiment 1. Asterisks denote significant differences ( $p < 0.001$ ) from chance (dashed line).

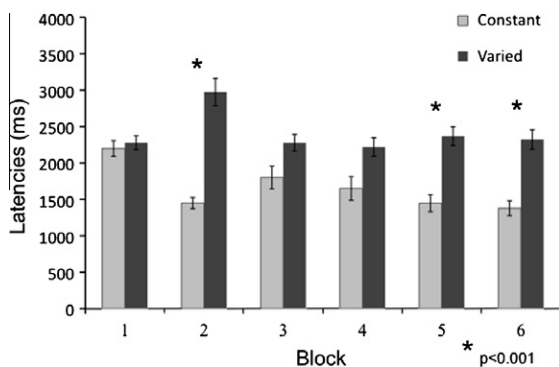
all infants than the Hard trials ( $M = 0.54$ ,  $SD = 0.16$ ), but infants in the Constant-Location condition out-performed infants in the Varied-Location condition on both Easy and Hard test trials, suggesting that predictable locations for the named objects during learning benefited both the learning of object properties and the binding of object names to individual objects. A (2) Condition by (2) Test-type ANOVA for a mixed design yielded main effects of Condition,  $F(1,30) = 15.38$ ,  $p < .0001$ , and Test,  $F(1,30) = 6.2$ ,  $p < .05$ ; the interaction did not approach significance. Learning was significantly above chance for both the Hard and Easy trials for infants in the Constant-location condition,  $t(15) = 5.91$  and  $t(15) = 3.33$ , respectively,  $p < 0.001$ , but above chance only on the Easy trials for infants in the Varied-Location condition,  $t(15) = 5.00$ ,  $p < 0.001$ .

Overall these results support the general idea – as outlined in Fig. 1 – that core cognitive processes that use object location are relevant to the processes through which infants map names to things. Since the location is actually irrelevant to the meaning of most common nouns – a chair is a chair regardless of where it is – one might have expected that spatial consistency would be irrelevant to young children’s learning of the name. Moreover, since infants were tested for their knowledge of object names in a new location, one might have predicted that learning experiences with varied locations would highlight the relevant link between an object and a name while learning experiences with consistent locations—where the location as well as the particular object, was associated with the name—would not. However, this was not the case. Instead, and consistent with the proposals about the fundamentally spatial nature of memory and attention summarized in Fig. 1, spatial consistency across naming events support learning and the formation of a robust link between a name and a thing that is independent of location.

In order to better understand the role of spatial consistency and variability within the task, we also examined infants’ looking behavior during the pop-up phase. In the Constant-Location condition, the object name said *before* the pop-up, not only predicts what will appear, but also predicts the location of the next pop-up; that is, if the infant stores the location with the name, infants in this condition could anticipate that location at which the pop-up will occur. The pre-pop-up name cue in the Varied-Location

condition only predicts what will appear, and provides no information as to the location of the next pop-up. To measure infants' learning and use of the pre-pop-up name as a cue for *where* to look, we measured (1) *anticipations* of the pop-up location defined as the first look to a location that began after the name but prior to the pop-up, and (2) *latencies* to look defined as the time between pop-up and an initiated look. Participants' looking behavior showed no evidence of *anticipating* the location of the next pop-up in the Constant-Location condition. Infants in both conditions almost *never* looked to the location (or any location) after the pre-pop-up name cue but prior to pop-up. Instead, they turned their eyes to the popped-up location only *after* the abrupt visual appearance and associated popping sound. A turned head and/or shift in eye gaze are motor actions. Thus, the lack of an *anticipatory* response may not be surprising in that motor plans for spatially-directed actions are often prepared in anticipation of a target's appearance, but are not put into an action until triggered by the presentation of the stimulus target, a pattern particularly likely when there is a near constant temporal interval between cue and action target (e.g., Kaufman et al., 2010). Latencies of eye movements after the onset of a stimulus, however, are affected by pre-planning and expectation (see Van der Stigchel, Meeter, & Theeuwes, 2006).

Fig. 4 shows the mean latencies to look to the popped-up location for the six blocks of the training for the Constant- and Varied-Location conditions. A mixed analysis of variance for a (2) Condition by (6) Block yielded main effects of Condition,  $F(1,30) = 41.36$ ,  $p < 0.001$  and Block  $F(5,150) = 4.49$ ,  $p < 0.01$  and a significant interaction between Condition and Block,  $F(5,150) = 10.30$ ,  $p < 0.001$ . Infants in the Constant-Location condition had overall much faster looking latencies ( $M = 1654.47$  ms,  $SD = 348$  ms) than did infants in the Varied-Location condition ( $M = 2406.45$  ms,  $SD = 312$  ms). The faster latencies in the Constant-Location condition indicate that infants learned the association between the name and location and used that information to prepare attention to the location at which the object would appear (for previous studies that use faster latencies as evidence for learning in infants see Johnson, Amso, & Slemmer, 2003; Kirkham, Slemmer, Richardson & Johnson, 2007).



**Fig. 4.** Mean latencies to look at the popped up object (and standard error of the mean) for blocks 1–6 for the Constant- and Varied-Location conditions in Experiment 1. Asterisks denote significant differences between the two conditions (Bonferroni post hoc tests,  $p < 0.001$ ).

As is apparent in Fig. 4, group differences in latencies to look emerged at the *first repetition of each object's location* (that is, in Block 2). Post-hoc Bonferroni pairwise comparisons indicate significant differences between the two groups for Block 2,  $t(30) = 7.58$ ,  $p < 0.001$ , but not for Block 1,  $t(30) = 0.55$ ,  $p > 0.1$ . Infants in the Constant-Location condition were faster to look on the second presentation of each object ( $M = 1449.5$  ms,  $SD = 304.9$  ms) than they were on the first presentation ( $M = 2199.81$  ms,  $SD = 432.4$  ms). In contrast, infants in the Varied-Location condition were faster on the first set of three trials ( $M = 2279.15$  ms,  $SD = 378.53$  ms) than on the second set ( $M = 2972.77$  ms,  $SD = 743.77$  ms), suggesting interference in looking time resulted from the changing of locations. Thus, this pattern is consistent with both a benefit from the association of a name with a repeated location and interference when the location of attention varies. That is, because the order of pop-up locations varied across trials (and because the orders were yoked in the two conditions), the only basis for more rapid orienting to the sound of the popped-up object in the Constant condition, or the slower Orienting in the Varied condition, is the association between name and location formed on the first block of three trials. This pattern thus provides strong evidence for the proposal that infant memories for the object names include information about the direction of attention when the name is heard and that a heard name leads to the preparation of a look to the associated direction. Moreover, infants appear to register and store the direction of attention with the heard name on trial one; and after just this *one* experience, the association between name and location is sufficient to influence looking behavior on the next trials. This one-trial location learning is consistent with similarly rapid location learning in adults (Altmann, 2004; Altmann & Kamide, 2004; Spivey & Geng, 2001).

Finally, to ensure that overall attention to the explicit naming event was comparable for infants in the two conditions, we also measured the *total duration* of looks to the object during the explicit naming event, specifically from the time period between the pop-up of the named object and that object being pressed down into the pop-up box and out of sight. The expectation is that duration of attention to the naming event will not differ between the two conditions because – after the pop-up – the experimenter engaged the infant in the same way in both conditions: explicitly naming the object by looking to the infant, looking at the object, pointing to the object, and naming the object. A mixed analysis of variance for a (2) Condition by (6) Block yielded only a main effect of Block,  $F(5,150) = 24.48$ ,  $p < 0.001$ . Looking durations to the object during the explicit naming event decreased over the course of the training trials as the objects became more familiar for all infants, suggesting that the naming event and the popped-up objects became less interesting over the training trials; however, there were no differences between the groups in looking at the objects after pop-ups when the in-view objects were explicitly named in the same way in both conditions. Looking durations for the first and last block for the Constant-Location condition (first block:  $M = 1435.98$  ms,  $SD = 377.45$  ms; last block:  $M = 736.69$  ms,  $SD = 254.50$  ms) were similar to those from the Varied-Location condition

(first block:  $M = 1456.69$  ms,  $SD = 402.19$  ms; last block:  $M = 764.75$  ms,  $SD = 218$  ms). This lack of difference is critical to the hypothesis because it implies that both groups of infants attended to the explicit naming events and if these naming events – not the ability to predict the direction of spatial attention during these experienced events – were all that mattered, the infants in the two conditions should have performed comparably in the name comprehension test. This, however, was not the case.

If expectancy-based attention – being able to predict the location of the target object – supports object name learning, then one might expect, in the Constant-Location condition, a correlation across individual infants between latency to look to the cued pop-up and object name learning. However, the results from infants in the Constant-Location condition provide no support for this prediction. Correlations for latencies to look to the cued location after the pre-pop-up name cue and correct choices at test in the Constant-Location condition were calculated using the latencies for Block 1, for Block 6, for the average latencies over all trials, and for the difference score between first and last trials; for all measures,  $r < 0.14$ , ns. The decrease in latencies across trials for infants in the Constant-Location condition clearly show learning of a name-location association; by hypothesis the ability to predict the location to which attention is to be directed supports the binding of the name to the object. It could be that this hypothesis is correct but that latency to look is too variable within and across infants to provide a sensitive measure of this learning that discriminates at the individual participant level. Experiment 2 was designed to address this issue.

In sum, the results of Experiment 1 provide strong support for the hypothesis that spatial predictability fosters the learning of object names. The evidence also suggests that during the learning trials, infants associated names to the predictable locations of a named object since the latency to look decreases in the Constant-Location condition. These latency data also suggest that the learned links between the label and location set up rapidly as the decrease in latency was evident by the first repetition of the cueing label and the location of the salient pop-up event. By the hypotheses outlined in Fig. 1, there are (at least) two not mutually exclusive pathways through which such spatial constancy might contribute to learning: (1) through the cued localization of attention which may itself lead to deeper processing and thus better learning about the attended event and (2) through the spatial indexing of memories and thus the ability to combine information across different encounters with an object and its name. Experiment 2 was designed to examine the first possibility and Experiment 3 was designed to examine the second.

### 3. Experiment 2

In the Constant-Location condition of Experiment 1, the heard name prior to pop-up predicted the location of the to-be-seen object. According to the hypothesis, after the first pop-up and naming event for that object, the pre-pop-up name cue activates *expectations* about *where* an attention-relevant event will occur and by hypothesis

these expectations lead to more effective attention and learning. In the Constant-Location condition of Experiment 1, the condition in which participants showed superior object name learning, latency to look to the popped-up location declined over trials, suggesting that infants learned the predictive relation between the label and where to attend and used it to deploy attention. This finding is consistent with the expectancy-based cued-attention hypothesis. However, latencies to look by individual infants were not correlated with the learning of the object names. Experiment 2 was designed to provide a perhaps more sensitive test of infants' expectations on hearing the name prior to the pop-up event in the Constant-Location version of the task. To do this, we used probe trials: on these probe trials the experimenter said an object name, but did not trigger a pop-up. When no object popped up at the expected time, would the infant look to the expected location? Would individual infants' expectancy-based looking on these probe trials predict their success in learning the object names?

#### 3.1. Method

##### 3.1.1. Participants

Sixteen infants (half female) who had not participated in Experiment 1 participated in this experiment. Ages ranged from 16.2 months to 18.2 months with a mean age of 17.35 ( $SD = 0.59$ ). All participants were exposed to the Constant-Location condition. Infants were recruited from a working and middle-class population in Bloomington, Indiana, and received a toy for participating in the experiment. One infant was recruited but replaced because of failure to complete the task.

##### 3.1.2. Stimuli, design and procedures

All aspects of the experiment were identical to that of the Constant-Location condition in Experiment 1 with two exceptions. First, after the 18 training trials, there were three probe trials, one for each object and each location. On these trials, the experimenter said the object name while looking at the infant and then waited 5 s (saying nothing but looking at the infant) before going to the next probe trial. The first looks to a pop-up location were measured during this period. Second, and because the main question concerns the probe trial performance, we reduced the number of testing trials. Specifically, immediately after the probe trials, infants were tested using six Easy test trials as in Experiment 1. Thus the three central dependent measures are: (1) correct or incorrect choices during testing, (2) first look to a location after the naming event on probe trials, and (3) looking latencies to the popped up object during training as in Experiment 1. As in Experiment 1, the data were coded frame by frame from the videos by one coder and then a second coder coded a randomly selected three participants. Agreement for looking latencies (again, based on differences of 500 ms or less just as in Experiment 1) was .80, for first looks on anticipation trials was .98, and for choice on test trials was 1.0. The correlations between the two coders' judged latencies was  $r = 0.59$ ,  $p < 0.001$ .



### 3.2. Results and discussion

The key question for this experiment is whether there is a relation, within individual participants, between looking to locations on the probe trials and the learning of the three object names. On the Easy test trials used in this experiment, infants chose the named object reliably more often than expected by chance,  $M = 0.61$  ( $SD = 0.19$ ),  $t(15) = 6.00$ ,  $p < .001$ . For the Probe trials, the first looks were scored as correct if they were looks to the location associated with the name. The mean proportion of such first looks to the correct location when hearing the name was  $0.58$  ( $SD = 0.38$ ) which is reliably greater than  $0.33$ , the expected value if the first looks were equally distributed across the three locations without relation to the associated name,  $t(15) = 2.70$ ,  $p < .05$ . This result indicates that infants were learning an association between the name and location. As shown in Fig. 5, individual infants' performances revealed that looking to the correct location during the probe trials was strongly related to correct choices during testing,  $r = 0.77$ ,  $p < 0.01$ . The individual infants who better anticipated the correct location on the probe trials were the infants who also were more correct in choosing the named object at test, a test that contained no spatial cues as to the name-object correspondence.

As in Experiment 1, the infants showed decreased latencies to look at the popped-up object across Blocks 1–6,  $F(5,75) = 6.19$ ,  $p < 0.01$ . Looking latencies towards the popped-up object were shorter for the last block ( $M = 1425.88$  ms,  $SD = 361.55$  ms) than for the first block ( $M = 2153.88$  ms,  $SD = 574.48$  ms). Again, comparisons of Block 1 with Block 2 ( $M = 1425.88$  ms,  $SD = 504$  ms) showed declines in latencies at the first repetition of an object's location (Bonferroni post hoc test,  $t(15) = 4.31$ ,  $p < 0.01$ ).

In sum, a learned name that is associated with attention to an object also provides a top-down guide as to *where to look*, an internal expectation of the location of the to-be-attended naming event. These internal expectations about where to look given the name predict children's knowledge of the object names even when tested in a new location and spatial arrangement. This result provides support for the hypothesis that *expectations* about looks *that are prior*

to the explicit naming event itself support the binding of the name to the thing.

### 4. Experiment 3

Experiment 3 was designed with three goals. First, in Experiments 1 and 2, in the Constant-Location conditions, there is a three-way temporal co-occurrence between name, location, and object. The hypothesis set in Fig. 1, and the results of Experiments 1 and 2, imply that name-location and object-location associations support the formation of name-object mappings that then do not depend on a common location. Accordingly, a first goal of Experiment 3 was to provide a strong test of this claim by temporally separating these two hypothesized contributing associations – name-to-location and object-to-location. That is, in Experiment 3, infants were presented with objects at locations (but no name) in the training phase, and then with a name at a location (but no object). Can they use separately formed object-location and name-location associations to bind the object and name?

The second goal was to more directly examine the proposal of the spatial indexing of memory. Experiments 1 and 2 provide direct evidence for spatial cueing in very young children and for the hypothesis that cued spatial attention supports the binding of an object name to object representations. However, the results are also consistent with the proposal that location information is stored with attended objects and with the object name. If this is so, then the direction of attention might also serve as an index for memories, activating the object that has been experienced at that location. This spatial indexing function of location information seems potentially important to the combining of information in a current experience with those in memory and thus in aggregating information across individual learning events. The purpose of Experiment 3 was to test this proposition. If infants index memories within a task context by the spatial location of those events and if they bind together multimodal events separated in time to build strong unified representations, then infants should be able to bind a name that is heard at one time but indexed in memory by a particular location to an object that is seen at another time at the same location.

The third goal of Experiment 3 was to specifically link the present phenomena with previous findings that children can link a remembered object to a heard name through the spatial direction of attention (Baldwin, 1993; Osina et al., 2012; Samuelson et al., 2011). To do this, we borrowed the logic of the Baldwin (1993) and Samuelson et al. (2011) design in temporally segregating the object information and the naming information but – as they did in their studies – providing a way to align names and objects via the location. Experiment 3 thus also provides a critical conceptual replication of those earlier findings which involved social cues, objects hidden and found in buckets, and an overall very different task context. Baldwin interpreted her results in terms of children's ability to infer the referential intent of the speaker and did not consider the role of the spatial alignment of the naming events and experienced objects; Samuelson et al. offered

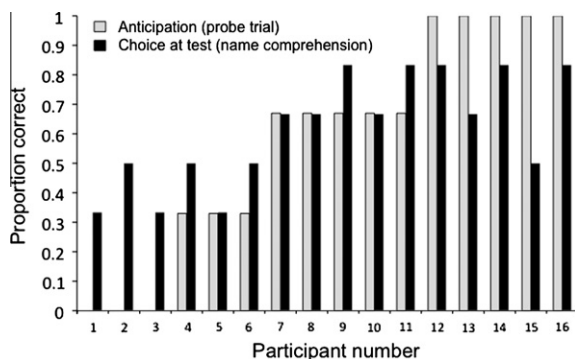


Fig. 5. Proportion correct on probe trials and on test trials for each participant in Experiment 2. Participants are ordered by performance on the probe trials testing anticipations of object locations given the cue.

and tested an alternative interpretation of the Baldwin findings in terms of spatial alignment. Relating the present findings to these earlier ones is critical to a programmatic and mechanistic understanding of how spatial attention and memory support word learning, as the novel triggered pop-up of the present task could engage distinct mechanisms from those found in more naturalistic social discourse.

In sum, Experiment 3 was designed to examine the necessary ingredients to such spatial indexing – that memories of visual objects include the location at which they were experienced and that a common location can be used to bind together currently experienced information with object memories. To this end, we repeated the Constant-Location procedure of Experiments 1 and 2 with two critical changes. First, for the 18 training trials, *no names were provided* – not as cues and not after pop-up. Thus from these trials, infants could only learn the association between objects and their locations. Second, after these 18 trials, the experimenter pointed to one covered location and provided the name. Testing followed immediately. If directed attention to the location activates the memory of the object that had been consistently seen there, then this single naming event with no object in view could be enough for infants to bind the heard name to the remembered object. In addition, Experiments 1 and 2 provided hints that infants might be encoding location information from the first trial, suggesting that perhaps *repeated* exposure to an object at a unique location may not be necessary for spatial indexing. Thus, a small sample of infants were also tested in a One-Block condition in which infants were presented with each object *once* at a unique location. If one exposure is all that is necessary for infants to encode the spatial location of objects – and if this one exposure is sufficient to form a name-object link when the names and objects are separated in time, then infants should also be able to map the name and object – through the common location – with just one exposure to the object at that location.

#### 4.1. Method

##### 4.1.1. Participants

Twenty-eight infants (half female) who had not participated in Experiment 1 or 2 participated in this experiment: Eighteen in the Main Condition and 12 in the One-Block condition. Ages for these infants ranged from 16.2 months to 18 months with a mean age of 17.42 ( $SD = 0.58$ ). All infants were recruited from a working and middle-class population in Bloomington, Indiana, and received a toy for participating in the experiment. Two infants did not complete the procedure and were replaced.

##### 4.1.2. Stimuli, design and procedure

The procedure in the Main Condition was similar to the Constant-Location condition in Experiments 1 and 2 except that there was no cue that predicted the next pop-up location and there was no naming event after the pop-up of the object. Each of the three unique objects were presented at their own constant location on all trials and each pop-up event was structured as follows: Prior to pop-up, the experimenter simply said, “Look!” then waited about 1 s,

and then popped the object up. After 18 trials (organized, as in Experiment 1 in the Constant-Location condition, into six blocks of three presentations in which each object was experienced once per block at its unique location) infants were then presented with *one* naming event. With no objects in view (all pressed down), the experimenter pointed to one pop-up location and said, “Look! It’s a blicket. There is a blicket in there.” The location (and thus target object) was randomly chosen for each participant. Following the training, Hard test trials as in Experiment 1 were used to test infants’ ability to pick out the “blicket” – the specific object associated with the named location. We did not include the Easy test trials because repeated testing on one object in these easy trials might have taught the word to the participants during the test trials alone. Testing trials alternated between Hard tests trials and Filler test trials using familiar objects (cup, flower, or spoon) for a total of six tests (three tests for the target, three for filler items). Choices were scored from video recordings, with 25% of the data scored by two independent coders with 100% agreement. The procedure in the One-Block condition consisted of just the first block of pop-ups, each object presented – with no naming – just *once* at its unique location. Infants proceeded immediately from this first block to the one naming event as described above, and other than the number of pop-up trials, the procedure was identical to that in the Main condition.

#### 4.2. Results and discussion

If infants in the Main Condition stored the object and its properties with its repeated location of experience and if attention to this location during a single naming event – with the object not present – was sufficient to activate location-indexed memories of the object, then the infants should have been able to bind the heard name to the *memory* of the object and its properties and then later choose that object from a set of similar and familiar objects when tested with the name. The results support this prediction: The mean proportion for correct choices of the named object at test in the Main Condition was 0.65 ( $SD = 0.29$ ) which is significantly greater than chance (0.33),  $t(17) = 4.64$ ,  $p < 0.001$ . These results suggest that infants stored and retrieved object representations by location.

The results in the One-Block condition suggest that multiple encounters with a spatially constant object may be necessary for young children to build object memories that are strong enough for infants to connect an object and name experienced at separate times through a common location; infants’ performance in this condition did not differ from chance, ( $M = 0.25$ ,  $SD = 0.35$ ,  $t(11) = -0.79$ ,  $p > 0.1$ ).

The latency effects of Experiments 1 and 2 suggest that name-location associations were formed in a single trial and were sufficient for a name heard prior to pop-up to lead to more rapid orientation to that pop-up. A one-trial effect in this experiment would require that infants form strong enough associations between a seen object and a location in one encounter, allowing them to link the *memory* of that object and all its properties to a heard word when attention was directed to that location. The results

indicate that the association was not sufficiently strong enough to allow infants to link a name and an object that was decoupled from location and correctly respond in the new location at test. Apparently, for infants to use location as an index to bind a name to a thing that do not co-occur with each other, multiple *spatially consistent* encounters with objects are needed. Thus the results provide evidence for the spatial indexing of object memories and for the use of a spatial direction of attention for combining new information with memories of recent experiences.

## 5. General discussion

Contemporary theories of early object name learning focus on the role of knowledge in organizing how children map names to things: knowledge about the structure of categories (Colunga & Smith, 2005; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002), about the structure of language (Waxman, 1998; Waxman & Booth, 2001; Woodward, L., & Ellen, 1998), about the referential nature of words (Bloom, 2000; Waxman & Gelman, 2009), and about the intentions of social partners (Baldwin, 1993; Tomasello, 1992). The extant evidence makes it quite clear that young children do have substantial knowledge in all these domains and that this knowledge guides word learning (Bloom, 2000; Smith et al., 2002, 2010; Waxman, 1998; Woodward & Markman, 1998). However, from the perspective of what children might know about how object names map to objects, the present findings might seem surprising. The meanings of basic level nouns, object names, do not depend on location: a cup is a “cup” be it in the sink or on the table. If very young children know this about object names, then one might have expected that the constant or varying location of a named object would not matter for the binding of the name to the object. We suspect that very young children do “know” that object names transcend space and time, but this does not mean that space and time do not matter for learning. Learning an object name and binding a name to a thing requires the real time working of memory and attention, and these core processes are foundationally spatial.

A large literature has shown that localized attention is itself important for forming coherent object representations (e.g., Johnson, Hollingworth, & Luck, 2008; Treisman, 1986), that location information is stored as part of object representations (Hollingworth & Rasmussen, 2010; Kahneman, Treisman, & Gibbs, 1992; Treisman & Zhang, 2006; Wood, 2011), and that top-down cues that direct attention to a specific location enhance processing and the binding of multimodal experiences into a unified memory (Talsma et al., 2010). The present experiments provide clear evidence of these processes in 1½ year olds, suggesting that they operate in these infants much as they do in adults and that they play a role in the binding of names to things. The results suggest that successful object name learning depends on localized attention, cued spatial attention, and spatial indexing, and have potentially important implications for understanding how disruptions in the integrated functioning of these processes may contribute to delays and atypical patterns of language development.

The present experiments – a first step with respect to a mechanistic understanding of the role of spatial attention in early word learning – were designed around the key prediction that spatial predictability should enhance the binding of objects to their names. This prediction, though supported in the object pop-up task used here, might seem at odds with a world – and everyday learning environment – in which objects and people move about. In the discussion that follows, we first consider the present findings relative to processes of spatial attention more generally and then we consider the specific question about their possible role in everyday, and more naturalistic, word learning.

### 5.1. Spatial attention and learning

The results strongly suggest that not all forms of attention are the same, nor that they have the same consequences for cognition and learning. Infants in both the Constant- and the Varied-Location conditions attended to the explicit naming event, in the sense of looking at the object when it was named. But infants in the Constant-Location condition, who had experiences that could build expectations about *where* the relevant information would appear and who could prepare spatially directed attention before the event, learned the object names much more robustly than infants in the Varied condition, who could not predict *where* the next naming event would be. That is, in Experiment 1, infants in both the Constant and Varied groups were cued by the name of the to-be-seen object prior to its sudden appearance. For infants in the Constant-Location condition this cue predicted both where to look and what would be seen. For infants in the Varied-Location condition, the cue predicted only what would be seen. This result thus suggests that the key cognitive difference is the expectation of *where* to direct attention. In Experiment 2, all infants heard the name as a predictive cue to a constant location for each unique object and the results show that infants who could direct attention to the right location given the naming cue (as measured on probe trials) were the ones who best learned the object names, a result that has implications for expectations – and prior planning of where to direct attention – as critical to binding the name to the object.

What is the mechanism through which expectations about where to direct attention benefit learning? As outlined in Fig. 1, there are several relevant pathways. Pertinent to these pathways is the distinction between endogenous and exogenous attention. Attention that is captured versus attention that is directed by expectancy (that is, by a learned cue) appear to involve fundamentally different systems that engage different neural networks (Landau, Esterman, Robertson, Bentin, & Prinzmetal, 2007) and to compete for control of the information that is processed (Berger, Henik, & Rafal, 2005). Critically, salience driven versus expectancy driven attention also have different consequences for how the attended information is processed (Breitmeyer & Ganz, 1976; Lennie, 1980). The pop-up event in the present experiment is highly salient. No infant ever failed to look at a pop-up event, even on the later trials. If in infants, attentional capture leads to stimulus processing that is less conducive to learning

about the object (Smith & Yu, *in press*; Wu & Kirkham, 2010), then rapid, startling presentations, such as the pop-ups in the present experiment, may be very poor contexts in general for learning about objects or object names. Spatial predictability in this context may have been important because it generated an expectancy-based cue that engaged the higher-level attentional networks that supported the higher-level learning processes needed to learn an object name.

It is also possible that expectations enhance bottom-up processing, leading to more *localized* and less diffuse attention. If attention is (at least in some ways, see Driver, 2001) like a spotlight that binds elements within that spotlight and enhances their processing, then a critical question is whether that spotlight is sharply defined and narrow, or diffuse and broad. One recent study (Farzin, Rivera, & Whitney, 2010) used the phenomenon of crowding, the reduced ability to identify an object as a result of surrounding clutter (Pelli & Tillman, 2008), to measure the spatial resolution of the “spotlight” in 6- to 15-month olds infants. The effects of crowding on object identification is not due to acuity *per se*, but rather is believed to reflect an inability to segregate and thus appropriately bind the features of different objects (Pelli, Cavanagh, Desimone, Tjan, & Treisman, 2007; Treisman, 2006). Spatial resolution with respect to crowding decreases from the fovea to the periphery (Pelli et al., 2007). Farzin et al. found that the spatial resolution of perception in 15-month-olds was *only half that of adults*, with object detection impaired when the neighboring object was as near as 3°. Evidence in the adult literature (e.g., Doshier & Lu, 2000; Shiu & Pashler, 1994) suggests that cued spatial attention is particularly effective in cluttered fields with many distractors and in noise. If expectancy-based attention narrows or sharpens the border of the attentional spotlight for children, then spatially predictive cues may be especially important for younger children who appear to have broader or more diffuse attentional fields.

Also relevant to the ideas illustrated in Fig. 1 is the spatial indexing of memories. Because location information is stored with object representations, location can also serve as an index to object representations in working memory and thus within a task context (Astle, Nobre, & Scerif, 2010; Ballard, Hayhoe, Pook, & Rao, 1997; Chun & Jiang, 1998; Xu & Chun, 2006). Within any learning task in which there are multiple encounters with multiple objects, the learner – in order to learn from and aggregate over those separate encounters – must connect the immediate instance of an object with its representation in working memory. Experiment 3 provides clear evidence that 1½-year-old infants can use consistent locations to connect an immediate experience to a previous one, and more specifically to bind a *memory* for an object to a currently heard word. This finding adds further support to the two recent demonstrations of the role of spatial consistency in young children’s retrieval of remembered objects (Osina et al., 2012; Samuelson et al., 2011). Those two studies, as is the present one, were specifically about *words* and object memories. As Fereirra et al. point out in their discussion of the “looking at nothing” phenomenon, this fact may be deeply related to the fundamental nature of online word comprehension, words direct looking to the mentioned ob-

ject, and when that object is not present, to the location that is stored with the object representation (see also, Knoeferle & Crocker, 2007). Looking to the location is an overt *behavioral* manifestation of referring: words literally direct attention to the physical location and remembered location of objects in the world.

The present results do not clearly distinguish these pathways – some may be more essential and others more supporting – but the results do provide a strong demonstration that space matters to infant learning. The experimental paradigm provides a clear framework for pursuing the independent contributions of these pathways.

## 5.2. Spatial attention and word learning

The pop-up task, though highly engaging and designed to enable the manipulation and measurement of cued spatial attention, is not at all like the contexts in which young children usually learn words; objects do not spring up suddenly in one location and then disappear. Instead, young children commonly learn object names in free-flowing interactions with objects in the context of a mature social partner (e.g., Lucariello & Nelson, 1986; Masur, Flynn, & Eichorst, 2005). Thus, one might ask: Is spatial consistency – and cued spatial attention – at all relevant to object learning in naturalistic learning contexts? Samuelson et al. (2011) and Osina et al. (2012) provide correlational evidence that it is. In the Samuelson et al. study, parents were asked to teach 1½-year-old infants two object names at a time as they engaged in play with their infants. Samuelson et al. found that parents established common ground by naturally segregating the objects on different sides of the table. Although they often held each object at center, and sometimes stacked them on top of each other, over the course of play, the parent typically held one object with one hand and the other object with the other hand, and perhaps more critically, when they put an object down to rest so as to engage in play with another object, they tended to consistently put unique objects to rest in unique locations. Thus, when the parent or child returned attention to that object, they looked and manually retrieved it from a certain direction. Across multiple interactions with an object, the child saw the object brought into play from a location which may have led to expectancy-based attention that supported binding the name to the object. Consistent with this idea, Samuelson et al. found that parents who more consistently located individual objects had infants who more effectively learned the object names. This idea that spatial consistency of referents plays a role in everyday language use also fits how speakers consistently gesture to different regions of space to indicate meaningful contrasts in a conversation (McNeill & Pedelty, 1995; Tuite, 1993).

In the present study, we used spatial consistency as a design property of the task through which to experimentally examine the role of cued spatial attention through spatial consistency. However, top-down and expectancy-based cues that spatially direct attention do not require a *constant* location, just a predictable one. The original work on endogenous cueing used the direction of an arrow (see Jonides, 1981; Posner, 1980) – which sometimes pointed in one direction and sometimes in another without any

directional association to the specific target. In these experiments, the arrow serves a top-down signal as to where the task relevant information will appear on that trial, not as a cue to a specific object at a specific location. The mechanistic underpinnings of arrow cues that direct attention by symbolic meaning and learned cues that direct attention to a specific location as in the present study may be functionally the same: enabling the perceiver to pre-orient attention to the location. Critically, everyday language is replete with cues that direct attention in the same way that arrows in the cued-attention experiments do: head turns, points and gestures, and the direction of eye gaze by the speaker, all of which have been shown to be critical – perhaps necessary – to children’s learning of object names, provide information about the location of relevant information (Ahktar & Tomasello, 2000; Pruden, Hirsh-Pasek, & Golinkoff, 2006). Typically the relevance of these cues to object name learning is discussed in terms of the *referential* nature of names and the young child’s ability to infer the referential intentions of the speaker (Baldwin, 1993; Bloom, 2000; Tomasello, 1992). However, gestures may also play a role in recruiting expectancy-based attentional networks that may be essential to binding names to things. Indeed, one might conjecture that the macro-level phenomenon of “referring” has its micro-level function in the expectancy-based localization of attention.

## 6. Conclusion

The spatial predictability of an object across multiple naming events supports better learning of the object name. Like adults, very young children store location information as part of their object representations. Like adults, they rapidly – in one trial – learn links between words and expected locations. And, as in adults, location itself serves an index to memory, enabling the learner to combine temporally separate but co-located information. These fundamental processes have not been systematically studied in infants and young children, and thus there are many open questions about their development and their exact nature. However, the present results suggest the urgency of a better understanding of these processes in young children, as they may be critical to object name learning and to the role of coherent attention (and social cues to attention) in that learning.

## Acknowledgements

We thank Amy Richards, Lauren Baker, David Samuels, Char Wozniak and Kelly Dakarian for assistance in the data collection and coding. We thank Dana Schuller for the design of the stimuli and conducting a pilot study with the pop-up box as part of her master’s thesis. We also thank the anonymous reviewers for their very helpful comments. The research was supported by National Institutes of Child Health and Development, R21HD068475.

## References

Ahktar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. Woodward, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Becoming a*

- word learning: A debate on lexical acquisition* (pp. 115–135). New York: Oxford University Press.
- Altmann, G. T. (2004). Language-mediated eye movements in the absence of a visual world: The ‘blank screen paradigm’. *Cognition*, 93(2), 79–87.
- Altmann, G. T., & Kamide, J. (2004). Now you see it, now you don’t: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferrerira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). New York: Psychology Press.
- Astle, D. E., Nobre, A. C., & Scerif, G. (2010). Subliminally presented and stored objects capture spatial attention. *The Journal of Neuroscience*, 30(10), 3567–3571.
- Baldwin, D. A. (1993). Early referential understanding: Infants’ ability to recognize referential acts for what they are. *Developmental Psychology*, 29(5), 832–843.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20(4), 723–767.
- Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology: General*, 134(2), 207–221.
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: The MIT Press.
- Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, 83(1), 1–36.
- Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition*, 13(1), 99–108.
- Canfield, R. L., & Haith, M. M. (1991). Active expectations in 2- and 3-month-old infants: Complex event sequences. *Developmental Psychology*, 27, 198–208.
- Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, 36(1), 28–71.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, 112(2), 347–382.
- Dosher, B. A., & Lu, Z.-L. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Research*, 40(10–12), 1269–1292.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, 92(1), 53–78.
- Eriksen, C. W., & Rohrbaugh, J. W. (1970). Some factors determining efficiency of selective attention. *American Journal of Psychology*, 83(3), 330–342.
- Farzin, F., Rivera, S. M., & Whitney, D. (2010). Spatial resolution of conscious visual perception in infants. *Psychological Science*, 21(10), 1502–1509. <http://dx.doi.org/10.1177/0956797610382787>.
- Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Science*, 12(11), 405–410.
- Fiebelkorn, I. C., Foxe, J. J., & Molholm, S. (2010). Dual mechanisms for the cross-sensory spread of attention: How much do learned associations matter? *Cerebral Cortex*, 20(1), 109–120.
- Fischer, J., & Whitney, D. (2009). Attention narrows position tuning of population responses in V1. *Current Biology*, 19(16), 1356–1361.
- Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron*, 54(5), 677–696.
- Greenberg, A. S., & Gmeindl, L. (2008). Strategic control of attention to objects and locations. *The Journal of Neuroscience*, 28(3), 564–565.
- 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(2), 467–479.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108(3), 356–388.
- Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, 17(1/2), 273–291.
- Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 36(3), 543–564.
- Hommel, B. (2002). Responding to object files: Automatic integration of spatial information revealed by stimulus-response compatibility effects. *Quarterly Journal of Experimental Psychology*, 55A(2), 567–580.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT.
- Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye tracking

- paradigm. *Proceedings of the National Academy of Sciences (USA)*, 100, 10568–10573.
- Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 34(1), 41–55.
- 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.
- Johnson, M. H., & Tucker, L. A. (1996). The development and temporal dynamics of spatial orienting in infants. *The Journal of Experimental Child Psychology*, 63(1), 171–188.
- Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kaufman, M., Churchland, M. M., Santhanam, G., Yu, B. M., Afshar, A., & Ryu, S. I. (2010). Roles of monkey premotor neuron classes in movement preparation and execution. *Journal of Neurophysiology*, 104(2), 799–810.
- Kerzel, D., Zarian, L., & Souto, D. (2009). Involuntary cueing effects on accuracy measures: Stimulus and task dependence. *Journal of Vision*, 9(11), 1–16.
- Kirkham, N. Z., Slemmer, J. A., Richardson, D. C., & Johnson, S. P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development*, 78, 1559–1571.
- Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, 57(4), 519–543.
- Landau, A. N., Esterman, M., Robertson, L. C., Bentin, S., & Prinzmetal, W. (2007). Different effects of voluntary and involuntary attention on EEG activity in the gamma band. *The Journal of Neuroscience*, 27(44), 11986–11990.
- Lange, K., & Röder, B. (2006). Orienting attention to points in time improves stimulus processing both within and across modalities. *Journal of Cognitive Neuroscience*, 18(5), 715–729.
- Lennie, P. (1980). Parallel visual pathways: A review. *Vision Research*, 20(7), 561–594.
- Lucariello, J., & Nelson, K. (1986). Context effects on lexical specificity in maternal and child discourse. *Journal of Child Language*, 13, 507–522.
- Martarelli, C. S., & Mast, F. W. (2011). Preschool children's eye-movements during pictorial recall. *British Journal of Developmental Psychology*, 29, 1–12. <http://dx.doi.org/10.1348/026151010X495844>.
- Masur, E. F., Flynn, V., & Eichorst, D. (2005). Maternal responsive and directive behaviours and utterances as predictors of children's lexical development. *Journal of Child Language*, 32, 63–91.
- McNeill, D., & Pedelty, L. (1995). Right brain and gesture. In K. Emmorey & J. S. Reilly (Eds.), *Gesture, sign, and space*. Hillsdale, NJ: Erlbaum.
- Nissen, M. J. (1985). Accessing features and objects: Is location special? In M. Posner & O. Marin (Eds.), *Attention and performance XI: Mechanisms of attention* (pp. 205–220). Hillsdale, NJ: Erlbaum.
- Osina, M. A., Saylor, M. M., & Ganea, P. A. (2012). When familiar is not better: 12-month-old infants respond to talk about absent objects. *Developmental Psychology*. <http://dx.doi.org/10.1037/a0027903>.
- Pelli, D. G., Cavanagh, P., Desimone, R., Tjan, B., & Treisman, A. (2007). Crowding: Including illusory conjunctions, surround suppression, and attention. *Journal of Vision* 7(2), 1–1. [www.csa.com](http://www.csa.com).
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129–1135. <http://dx.doi.org/10.1038/nn.2187>.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3–25.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109(2), 160–174.
- Pruden, S. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). The social dimension in language development: A rich history and a new frontier. In P. Marshall & N. Fox (Eds.), *The development of social engagement: Neurobiological perspectives* (pp. 115–152). New York: Oxford University Press.
- Richardson, D. C., Altmann, G. T. M., Spivey, M. J., & Hoover, M. A. (2009). Much ado about eye movements to nothing: A response to Ferreira et al.: Taking a new look at looking at nothing. *Trends in Cognitive Science*, 13(6), 235–236.
- Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, 133(1), 46–62.
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, 76(3), 269–295.
- Samuelson, L., Smith, L. B., Perry, L., & Spencer, J. P. (2011). Grounding word learning in space. *PLoS One*, 6(12), e28095. <http://dx.doi.org/10.1371/journal.pone.0028095>.
- Sanderson, P. M., Scott, J. J. P., Johnston, T., Mainzer, J., Wantanbe, L. M., & James, J. M. (1994). MacSHAPA and the enterprise of Exploratory Sequential Data Analysis (ESDA). *International Journal of Human-Computer Studies*, 41(5), 633–681.
- Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, 80, 1–46.
- Shiu, L., & Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology: Human Perception and Performance*, 20(5), 1037–1054.
- Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science*, 34(7), 1287–1314.
- Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, 13, 13–19.
- Smith, L. B. & Yu, C. (in press). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language, Learning and Development*.
- Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235–241.
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Science*, 14(9), 400–410.
- Tomasello, M. (1992). The social bases of language acquisition. *Social Development*, 1(1), 67–87.
- Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, 255(5), 114–125.
- Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 40A(2), 201–237.
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition Special Issue: Visual Search and Attention*, 14(4–8), 411–443. <http://dx.doi.org/10.1080/13506280500195250>.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
- Treisman, A., & Zhang, W. W. (2006). Location and binding in visual working memory. *Memory & Cognition*, 34(8), 1704–1719.
- Tuitt, K. (1993). The production of gesture. *Semiotica*, 93(1/2), 83–105.
- Van der Stigchel, S., Meeter, M., & Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neuroscience and Biobehavioral Reviews*, 30(5), 666–679.
- Waxman, S. R. (1998). Linking object categorization and naming: Early expectations and the shaping role of language. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 38, pp. 249–291). San Diego, CA: Academic Press.
- Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, 43, 217–242.
- Waxman, S., & Gelman, S. (2009). Early word learning entails reference not merely associations. *Trends in Cognitive Science*, 13(6), 258–263.
- Wood, J. N. (2011). When do spatial and visual working memory interact? *Attention, Perception, & Psychophysics*, 73(2), 420–439.
- Woodward, Amanda, L., Markman, Ellen, M. (1998). Early word learning. In W. Damon, D. Kuh, & R. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (Vol. 2, pp. 371–420). New York, NY: John Wiley & Sons Inc.
- Wright, R. D., & Ward, L. M. (2008). *Orienting of attention*. Oxford: Oxford University Press.
- Wu, R., & Kirkham, N. Z. (2010). No two cues are alike. Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*, 107(2), 118–136.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440, 91–95.