



Contents lists available at ScienceDirect

# Journal of Experimental Child Psychology

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

## Sixteen-month-olds can use language to update their expectations about the visual world

Patricia A. Ganea<sup>a,\*</sup>, Allison Fitch<sup>b</sup>, Paul L. Harris<sup>c</sup>, Zsuzsa Kaldy<sup>b</sup>

<sup>a</sup> University of Toronto, Toronto, Ontario M5S 3G3, Canada

<sup>b</sup> University of Massachusetts Boston, Boston, MA 02125, USA

<sup>c</sup> Harvard University, Cambridge, MA 02138, USA

### ARTICLE INFO

Article history:  
Available online xxxx

Keywords:  
Mental representation  
Updating  
Absent reference  
Language  
Object location  
Visual world

### ABSTRACT

The capacity to use language to form new representations and to revise existing knowledge is a crucial aspect of human cognition. Here we examined whether infants can use language to adjust their representation of a recently encoded scene. Using an eye-tracking paradigm, we asked whether 16-month-old infants ( $N = 26$ ; mean age = 16;0 [months;days], range = 14;15–17;15) can use language about an occluded event to inform their expectation about what the world will look like when the occluder is removed. We compared looking time to outcome scenes that matched the language input with looking time to those that did not. Infants looked significantly longer at the event outcome when the outcome did not match the language input, suggesting that they generated an expectation of the outcome based on that input alone. This effect was unrelated to infants' vocabulary size. Thus, using language to adjust expectations about the visual world is present at an early developmental stage even when language skills are rudimentary.

© 2016 Elsevier Inc. All rights reserved.

### Introduction

Much of the information that we have about the world is based on testimony provided by other people. Thus, on the basis of other people's testimony, we can acquire knowledge about events that

\* Corresponding author.

E-mail address: [patricia.ganea@utoronto.ca](mailto:patricia.ganea@utoronto.ca) (P.A. Ganea).

<http://dx.doi.org/10.1016/j.jecp.2015.12.005>

0022-0965/© 2016 Elsevier Inc. All rights reserved.

we would otherwise know little or nothing about—for example, events in the distant past. In addition, however, other people's testimony can help us to update our knowledge of the current state of the world (Harris, 2012). Indeed, we often rely on such testimony to provide us with information about changes that have occurred with respect to people or objects that are already known to us. For example, via testimony, friends and family members often provide us with updates about their lives. They can tell us about changes in their love lives, their jobs, their kitchen appliances, and so forth. By implication, updating based on verbal testimony has a broad scope.

Developmental research has provided a wealth of information on infants' ability to update object representations on the basis of visual information (e.g., Feigenson & Yamaguchi, 2009; Koehlin, Dehaene, & Mehler, 1997; Uller, Carey, Huntley-Fenner, & Klatt, 1999; Wellman, Cross, & Bartsch, 1986; Wynn, 1992) such as when changes are being made to the number of objects in a scene or to an object's location in the scene. However, only recently have we begun to learn information about children's ability to update their knowledge of an object through language. Ganea and Harris (2010, 2013) showed that when toddlers aged 30 months were told that an object they had put in one container had been moved to a different container during their temporary absence, they searched for the object in the new container, not where they had originally put it. By implication, using the testimony provided by another person, they updated their own prior representation of the state of the world and searched correctly.

However, in those same studies, younger toddlers—aged approximately 24 months—were much less accurate in making such location updates. They often searched for the object on the basis of their earlier firsthand observation of its whereabouts—an error that did not occur in a control condition where they saw the object moved to a new location rather than learning about its movement via verbal testimony. Alternatively, they first searched in the original container and only subsequently went on to search at the object's new location. Thus, these younger infants were less likely than older infants to give priority to what they had been told.

One possible interpretation of this age change is that younger infants have difficulty in updating their representation of the world on the basis of verbal input. To the extent that a great deal of conversational input to young children is concerned with objects, events, and outcomes that are present and potentially observable within the immediate situation (Morford & Goldin-Meadow, 1997), such a restriction might not create any obvious cognitive difficulty. According to this hypothesis, we might expect children's updating ability to improve as their early language becomes increasingly displaced from the here and now.

However, despite the above findings, updating may emerge considerably earlier in development. There are several reasons for considering this alternative possibility. First, task demands may have underestimated the ability of younger toddlers in the studies of Ganea and Harris (2010, 2013). They may have understood what they were told and updated their representation of the object's location but failed to inhibit a prepotent response of searching in the place where they had last seen the object. Under this interpretation, infants can engage in language-based updating but have difficulty in using the updated representation of location to guide voluntary actions such as search. Similar gaps between the representation of location and the execution of a response have been well established in studies of infants' ability to search for a hidden object (Diamond, 1985).

Second, by the beginning of the second year of life, infants are adept at incorporating testimony into their expectations about new unseen referents. For example, infants are able to reason about the quantities of unseen objects and nonobvious functions of objects through information provided via language alone (Graham, Kilbreath, & Welder, 2004; Xu, Cote, & Baker, 2005).

Third, studies of infants' ability to engage in updating with respect to an object's properties—as opposed to its location—have yielded encouraging results. For example, in one study infants left a stuffed animal in a room and went next door. They were subsequently told that the stuffed animal had been accidentally made wet. When they went back to retrieve it, 22-month-olds picked out a wet version of the stuffed animal rather than a dry version (Ganea, Shutts, Spelke, & DeLoache, 2007). Furthermore, follow-up studies on property updating demonstrated that reducing task demands by strengthening the working memory representation of the to-be-updated object and increasing its familiarity promoted updating in even younger toddlers, with 19-month-olds succeeding in updating object properties (Galazka & Ganea, 2014).

Taken together, these considerations suggest that even younger infants—provided that they can understand a simple assertion about object movement—might update their expectations about the location of an object based on what they are told. The current violation-of-expectation study tested this prediction using a computer presentation. The violation-of expectation paradigm used here was an adaptation of the blank-screen paradigm, which was originally designed to record eye movements as adult participants listen to a sentence when the screen is blank (Altmann & Kamide, 2004). This paradigm allows one to measure patterns of eye movements when the visual scene is absent (e.g., whether infants look to locations of mentioned absent objects) and to measure whether looks to the scene when visible are influenced by prior language. Accordingly, in this study we wanted to explore infants' looking patterns when the visual scene was absent at the time of the linguistic input and also when the visual scene was available again.

Sixteen-month-old infants were first familiarized with two pairs of common animals (two dogs and two cats) on a computer screen. Next, they were presented with a scene that included a target location—for example, a bed—and two different familiar animals, always one dog and one cat, located on either side of the target location. Infants heard a speaker call their attention to each animal and to the target location. Then children watched as a curtain went down to hide the scene and the speaker described an invisible action of one of the two animals—for example, “Now the dog goes to the table. The dog is on the table.” The curtain was then raised to reveal a scene that was either congruent with the speaker's testimony (i.e., the dog on the table and the cat still in its original position) or incongruent with her testimony (i.e., the cat on the table and the dog still in its original position). If infants successfully updated their representation of the scene based on what they were told while the curtain was lowered, they should be surprised by the incongruent scene but not by the congruent scene. We tested infants' expectations by comparing their looking time patterns to the incongruent versus congruent scene.

## Method

### *Participants*

The participants were 41 healthy, full-term, 16-month-old infants (age range = 14;15–17;15 [months;days]). Of these participants, 11 did not complete the required two blocks of trials as a result of fussiness and were excluded from further analysis. In addition, 3 infants were excluded due to experimenter error, and 1 infant was excluded due to equipment failure. The final sample consisted of 26 infants (12 female and 14 male; average age = 16;0,  $SD = 27.2$  days). Of this final sample, 16 infants were Caucasian (1 of them Hispanic), 3 were African American, 2 were biracial, and 5 families chose not to report their race and ethnicity. Parents of participants were recruited from the greater Boston area in the northeastern United States based on data from state birth records. They received \$20 and a small gift for participation. All participants were from dominantly (>75%) English-speaking households, with receptive vocabularies ranging from 11 to 350 words ( $M = 153.54$ ,  $SD = 75.51$ ) and productive vocabularies ranging from 0 to 66 words ( $M = 26.83$ ,  $SD = 20.81$ ) as measured by the MacArthur–Bates Communication Development Inventory (MCDI). Statistics on MCDI items pertinent to the study's design are displayed in Table 1.

### *Apparatus*

We used a Tobii T120 eye-tracker running Tobii Studio 3.0 software (Tobii Technology, Stockholm, Sweden) sampling at 60 Hz to measure eye movement patterns. Eye position was determined by averaging data from both eyes. When data could be acquired from only one eye (e.g., when the infant's face was momentarily not centered on the screen), they were not used. Fixations were defined using Tobii Studio's standard built-in fixation filter (Tobii Fixation Filter), which automatically interpolates segments of missing data that are shorter than 100 ms. Participants sat on their parents' laps approximately 60 cm away from the built-in 17-inch monitor in a dimly lit testing room. Parents' eyes were shielded by a visor, by modified sunglasses with occluded lenses, or by parents closing their eyes.

Please cite this article in press as: Ganea, P. A., et al. Sixteen-month-olds can use language to update their expectations about the visual world. *Journal of Experimental Child Psychology* (2016), <http://dx.doi.org/10.1016/j.jecp.2015.12.005>

**Table 1**

Summary statistics of MCDI results for names of agents and objects used in experiment.

Item	Understands	Understands and says	Does not understand
Dog	10	12	2
Cat	14	7	3
Table	0	10	14
Bed	18	2	4

Note. Values in table are numbers of participants.

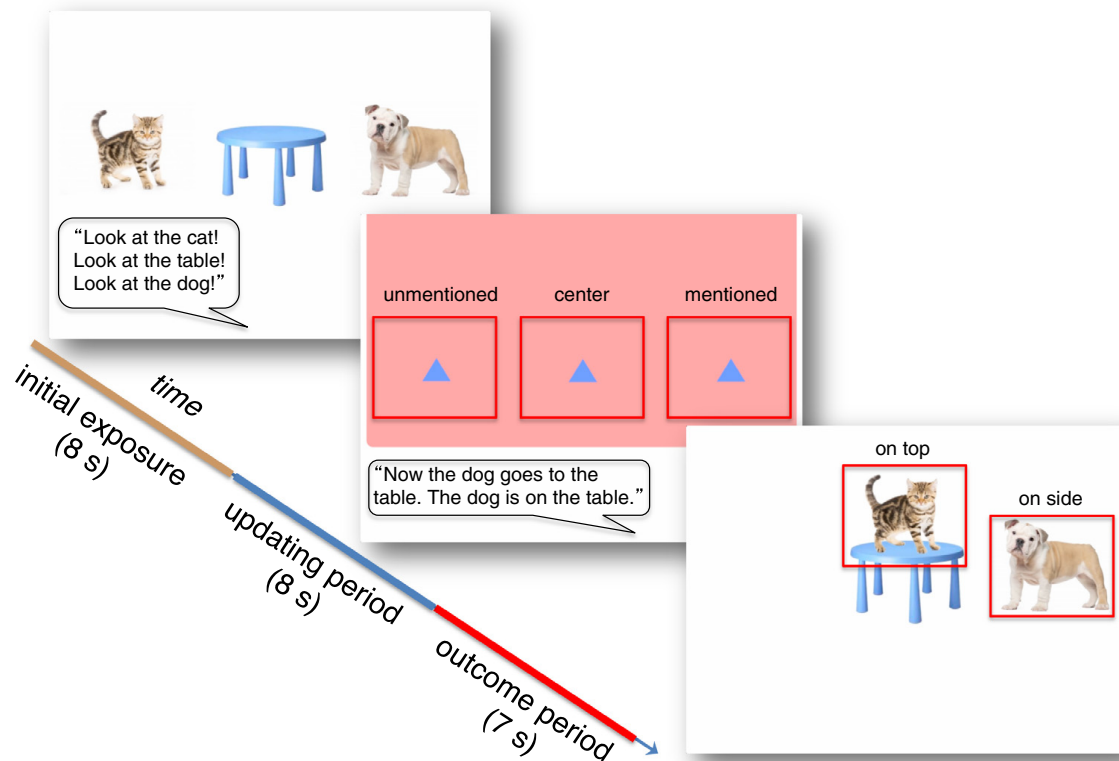
### Stimuli and procedure

We employed a within-participants design with two blocks of test trials: Congruent versus Incongruent trials. Each block consisted of 4 familiarization trials followed by 6 test trials. Participants saw both blocks of trials, with a short break in between for feeding, changing, or a brief walk down the hallway. (Breaks varied in length based on the individual infants' needs but were typically between 5 and 10 min.) The order of the two blocks was counterbalanced across participants. The blocked design was chosen to maximize the likelihood of being able to observe a difference between the two conditions. At the beginning of each block, each participant was calibrated using a standard 5-point calibration procedure. Each block took approximately 2.5 min. Agents (cats and dogs) and locations (table and bed) were quasi-randomly drawn from the same two sets of cats and dogs and locations (table and bed) in each trial.

In *familiarization* trials, a curtain was raised to reveal either two cats or two dogs (counterbalanced in order of presentation) for 4 s. Right after the curtain went up, a female voice (prerecorded by a native English speaker in child-directed speech) said, "Look at the cats!" or "Look at the dogs!" *Test* trials consisted of three phases: an initial exposure, an updating period, and an outcome period. In the initial exposure, a curtain was raised to reveal an object in the middle of the stage (either a bed or a table) and a cat and a dog on each side of the object (side counterbalanced and order quasi-randomized; see Fig. 1). The female voice said, "Look at the cat! Look at the table/bed! Look at the dog!" This initial presentation of the scene, along with the recording, lasted for 8 s. Then a curtain lowered, occluding the scene, displaying only three identical triangles as placeholders for the positions of the animals/object. These triangles, which remained static throughout the test phase, served as anchor points for the referents and allowed participants to access each referent's spatial location during speech processing, as has been demonstrated with adults (Altmann & Kamide, 2004; Ferreira, Apel, & Henderson, 2008). Following the descent of the curtain, the female voice said, "Now the dog goes to the table. The dog is on the table." The curtain remained in the lowered position while the testimony was given for 8 s (henceforth the updating period). Then the curtain was raised again to reveal either a scene that was congruent with the statement during the updating period (the dog was on the table, whereas the cat remained in its original position; *Congruent* condition) or incongruent with the statement (the cat was on the table, whereas the dog remained in its original position; *Incongruent* condition). This outcome scene was presented for 7 s (henceforth the outcome period), and then the curtain was lowered and the trial ended. A sample video of a participant during the Incongruent condition is provided in the online [Supplementary material](#) to illustrate the procedure.

### Data analysis

To analyze gaze patterns, we defined two rectangular (327 by 268 pixels) areas of interest (AOIs) around the two animals (the cat and the dog; see Fig. 1) during the outcome period and defined one rectangular (327 by 268 pixels) AOI around each triangle placeholder during the updating period. AOIs were defined a priori to contain the entire animal during the outcome period plus some additional space around them to allow for small calibration errors. The AOIs for the triangle placeholders during the updating period were larger, such that the adjacent AOIs filled the space between them without any gaps. Total looking time to each AOI (defined by the cumulative duration of visits within an AOI) as a proportion of looking to the entire screen was analyzed for both the updating and outcome



**Fig. 1.** Schematic of a test trial (Incongruent condition). Red rectangles and tags above them indicate AOIs; they were not visible to participants. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

periods (see Results). Trials in which a 1-s cumulative looking time threshold was not met were excluded from analysis (42 of 312 trials); participants who did not meet this threshold in at least 3 of 6 trials in both blocks were likewise excluded (11 participants, listed under excluded participants). The 1-s threshold was chosen as a conservative estimate of the minimum length of time needed for infants to process each object and its corresponding location. No other data filtering was applied.

## Results

For each participant, a proportion of total looking time during the outcome period of each condition was calculated by dividing the total looking time to each AOI (e.g., the animal on top of the table/bed) by the total looking time to the whole screen and then averaging across the 6 trials. Proportion of looking time was calculated only for valid trials (at least 1 s of cumulative looking to the screen during the outcome period). All participants completed at least 3 of the 6 trials per block ( $M = 5.19$  trials in the Congruent condition and  $M = 5.11$  trials in the Incongruent condition), and the number of trials completed did not significantly differ between conditions,  $t(25) = 0.278$ ,  $p = .78$ ,  $d = 0.05$ . Mean looking times during the outcome and updating periods are presented in Table 2.

### Outcome period

To identify a violation-of-expectation (VoE) effect, we conducted a 2 (Condition)  $\times$  2 (Order of Block Presentation) mixed design analysis of variance (ANOVA) on the proportion of looking time to the outcome AOI (the animal in the outcome location, i.e., on top of the table/bed). Data are shown



**Table 2**

Mean looking times (in seconds) to test trials by phase and condition.

Condition	Encoding phase	Testimony phase	Outcome phase
Congruent	5.89 (1.75)	5.1 (1.85)	5.79 (1.10)
Incongruent	5.03 (1.67)	5.19 (1.35)	5.63 (1.57)

Note. Standard deviations are in parentheses.

in Fig. 2. Results indicated a main effect of condition,  $F(1, 24) = 44.16, p < .001, \eta_p^2 = .648$ , where participants showed a significantly longer looking time in the Incongruent condition ( $M = .68, SD = .12$ ) than in the Congruent condition ( $M = .46, SD = .11$ ). There was no significant main effect of order,  $F(1, 24) = 2.02, p = .17, \eta_p^2 = .078$ , or interaction between condition and order,  $F(1, 24) = 0.02, p = .89, \eta_p^2 = .001$ . Proportion of looking time to the outcome AOI (the animal that was on top of the target location) was also significantly above chance in the Incongruent condition,  $t(25) = 7.48, p < .001, d = 1.41$ , and at chance in the Congruent condition,  $t(25) = -1.56, p = .13, d = 0.32$ . Finally, participants were categorized into those who demonstrated a VoE effect (i.e., those who looked longer in the Incongruent condition than in the Congruent condition) and those who did not. Of the 26 participants, 23 demonstrated a VoE effect. A chi-square goodness-of-fit test determined that this is significantly different from chance,  $\chi^2(1) = 15.385, p < .001$ .

To explore the robustness of the VoE effect observed above, we repeated the 2 (Condition)  $\times$  2 (Order of Block Presentation) mixed design ANOVA on proportion of looking time to the outcome AOI on just the *first* trial in each block.<sup>1</sup> The effect was present from the first trial; the main effect of condition was significant,  $F(1, 24) = 5.143, p = .033, \eta_p^2 = .176$ ; that is, participants looked longer at the outcome AOI in the Incongruent condition ( $M = .554, SD = .17$ ) than in the Congruent condition ( $M = .446, SD = .20$ ). No significant main effect of order was found,  $F(1, 24) = 0.53, p = .473, \eta_p^2 = .022$ , and condition and order did not interact,  $F(1, 24) = 0.11, ns, \eta_p^2 = .005$ .

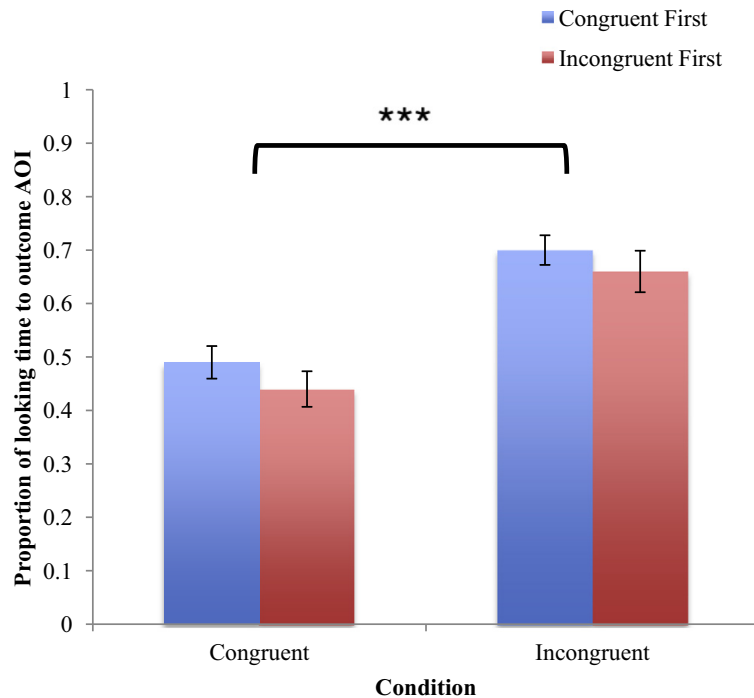
Of the 24 participants with completed MCDIs, the overwhelming majority reportedly understood the words “dog”, “cat”, and “bed”, but only 10 were reported to understand the word “table” (see Table 1). To test whether the VoE effect was dependent on the destination of the location change (i.e., only those who understood “table” showed a VoE for trials where a table was presented), trials in which the table was the destination were analyzed separately. We performed a 2 (Condition)  $\times$  2 (Understands “Table” vs. Does Not Understand “Table”) mixed design ANOVA on proportion of looking time for these trials only. The trend from the previous analyses was present (albeit with a smaller effect) in this subset of our data; the main effect of condition was marginally significant,  $F(1, 22) = 4.125, p = .055, \eta_p^2 = .158$ , in that the proportion of looking to the outcome AOI was larger in the Incongruent condition ( $M = .56, SD = .13$ ) than in the Congruent condition ( $M = .48, SD = .13$ ). There was no main effect of understanding the word “table”,  $F(1, 22) = 0.319, p = .578, \eta_p^2 = .014$ , and this factor did not interact with condition,  $F(1, 22) = 1.352, p = .257, \eta_p^2 = .058$ .

### Updating period

To examine whether the differences in looking time during the outcome period were reflected in gaze patterns when the predictive expectation may have been generated, we also examined looking time differences during the updating (screen down) period. We hypothesized that if predictive expectations occurred, participants would display longer looking to the central AOI as well as the AOI of the mentioned animal when those referents were stated in the testimony, reflecting adult-like speech processing. Fig. 3 plots a time course of the proportion of looks to each AOI by condition during this period.

We conducted a 2 (Condition)  $\times$  2 (Order of Block Presentation)  $\times$  3 (AOI) mixed design ANOVA on the proportion of total looking. Data are shown in Fig. 4. Results indicated a main effect of AOI,  $F(2, 48) = 35.59, p < .001, \eta_p^2 = .597$ . Pairwise comparisons (Bonferroni adjusted for multiple comparisons)

<sup>1</sup> We also analyzed whether looking times changed over time within a block and found no significant trends.



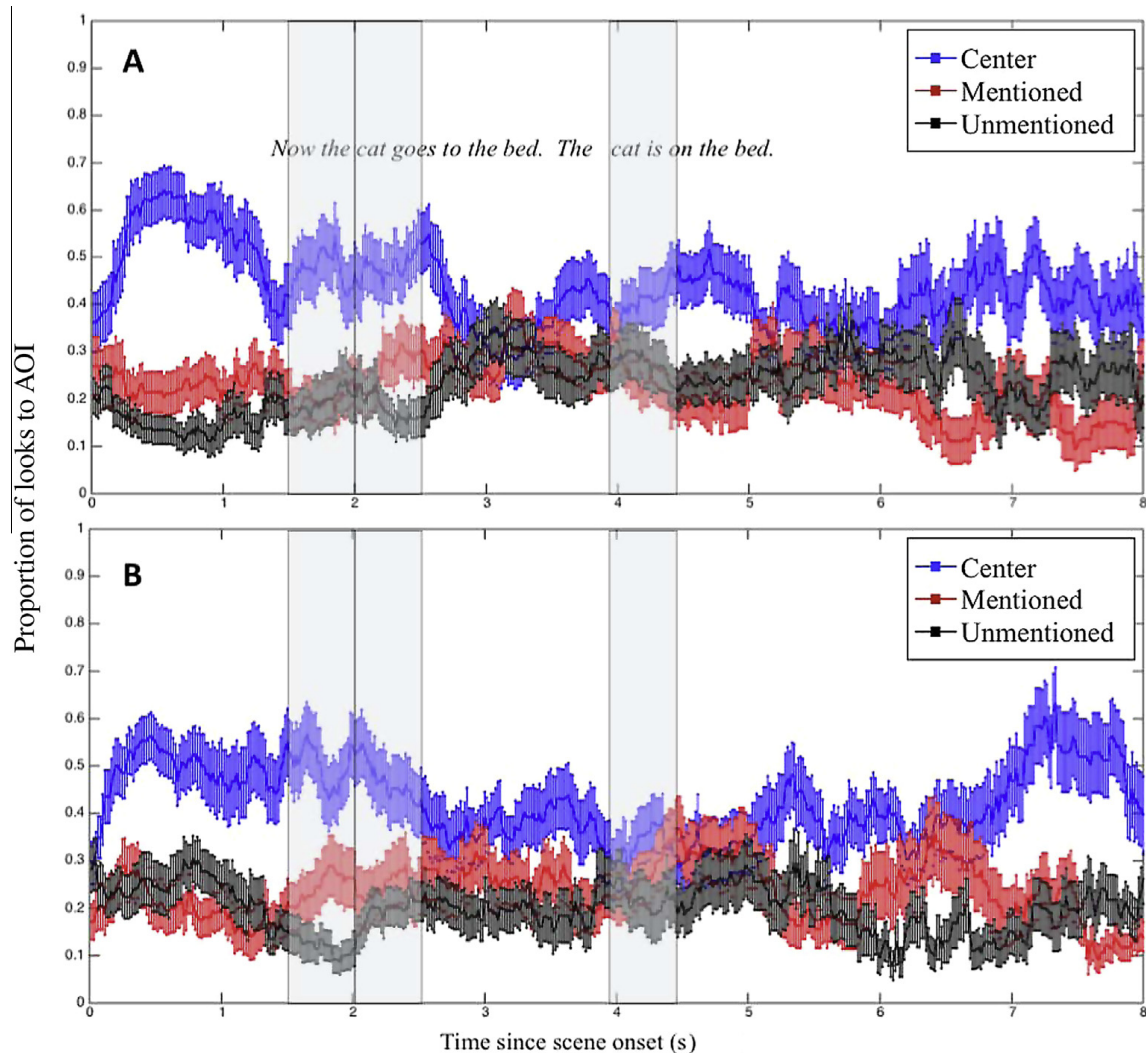
**Fig. 2.** Proportions of looking time to outcome AOI (animal in the outcome location) during the outcome period as a function of condition and block order. Error bars display standard errors of the means. \*\*\* $p < 0.001$ .

demonstrated that this effect was driven by looking to the center AOI ( $M = .421$ ,  $SD = .12$ ), which was significantly longer than looking to both the mentioned AOI ( $M = .202$ ,  $SD = .076$ ,  $p < .001$ ) and the unmentioned AOI ( $M = .188$ ,  $SD = .081$ ,  $p < .001$ ). There were no significant differences between the looking times to the mentioned and unmentioned AOIs ( $p > .99$ ).

This analysis also yielded a significant main effect of condition,  $F(1, 24) = 9.26$ ,  $p = .006$ ,  $\eta_p^2 = .278$ , in that participants spent a slightly larger portion of their total looking time within the three AOIs in the Congruent condition (for any given AOI:  $M = .284$ ,  $SD = .035$ ) than in the Incongruent condition (for any given AOI:  $M = .257$ ,  $SD = .035$ ). Arguably, infants in the Incongruent condition became more aroused over time because they were presented with a conflict between sentences and outcomes on 6 successive test trials. Such a repeated violation of expectation may have led them to look around randomly during the updating (screen down) period. There was no significant main effect of order,  $F(1, 24) = 0.007$ ,  $p = .94$ ,  $\eta_p^2 < .001$ , and no interactions.

Although participants looked overall at the center AOI more than at the other two AOIs during the updating period, we investigated the possibility that infants' gaze behavior may have changed right around the time when they heard the agent being mentioned during the testimony. To do so, we compared the proportion of looking to the mentioned agent during three 500-ms time windows—pre-mention (500 ms before the first mention), post-first mention (500 ms after the first mention), and post-second mention (500 ms after the second mention)—averaged across trials within a block. We conducted a 2 (Condition)  $\times$  3 (Time Window) repeated measures ANOVA on the proportion of total looking to the mentioned AOI. Results indicated that there was no significant main effect of condition,  $F(1, 19) = 0.225$ ,  $p = .64$ ,  $\eta_p^2 = .012$ , or time window,  $F(2, 38) = 0.926$ ,  $p = .41$ ,  $\eta_p^2 = .046$ . There was also no interaction between condition and time window,  $F(2, 38) = 0.942$ ,  $p = .39$ ,  $\eta_p^2 = .047$ . In the Incongruent condition, the average proportion of time spent on the mentioned agent was .257 (.215) during the pre-mention period, .286 (.314) during the first post-mention period, and .269 (.255) during the second post-mention period. In the Congruent condition, these values were .289 (.256), .262 (.258), and .184 (.168), respectively.

Finally, we went on to explore possible developmental contributions to the VoE effect observed during the outcome period. To do this, we examined relationships between the magnitude of VoE



**Fig. 3.** Proportion of looks to each AOI over time during the updating period in the Congruent condition (A) and the Incongruent condition (B). Testimony and 500-ms time windows (pre-mention, post-first mention, and post-second mention) are overlaid.

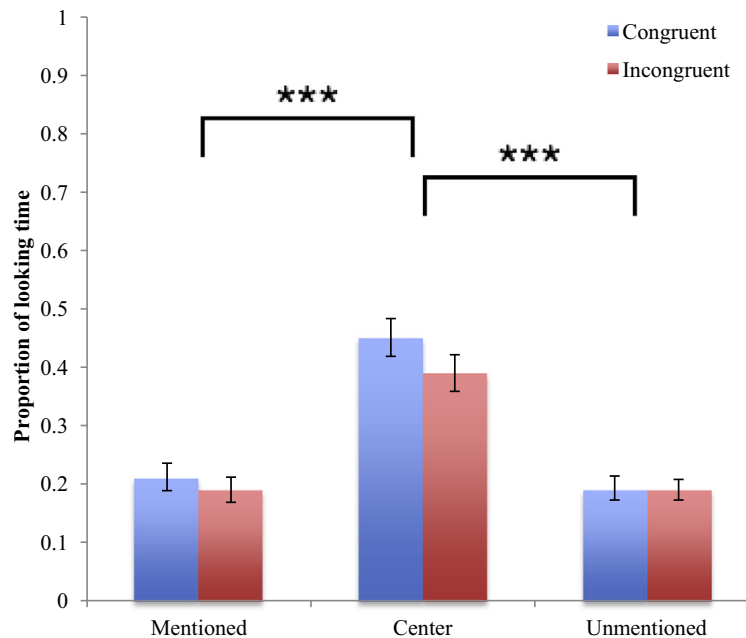
(the difference in proportion of looking to the event AOI between the Incongruent and Congruent conditions) and age (in days) as well as receptive vocabulary size on the MCDI.<sup>2</sup> The correlations between VoE effect magnitude and age (Pearson's  $r = -.06$ ,  $p = .78$ ) and receptive vocabulary ( $r = -.27$ ,  $p = .21$ ) were not significant.

## Discussion

The capacity to use language to acquire knowledge is a fundamental aspect of human cognition. Here we examined an important aspect of that knowledge acquisition process, namely the ability to update one's representation of the world on the basis of verbal testimony rather than direct observation. Thus, we asked whether infants can update their representation of a recently encoded visual scene based on a verbal description of a change in that scene. The visual scene contained two referent animals with an inanimate object located between them. When the visual scene was occluded, infants

<sup>2</sup> Two children were not included in the MCDI analyses. The parent of one of these children chose not to fill out the questionnaire; another parent misunderstood the instructions, yielding an extreme score.





**Fig. 4.** Proportion of looking to each AOI during the testimony period as a function of condition. Error bars display standard errors of the means. \*\*\*  $p < 0.001$ .

were told about the change in the scene, specifically a change in the location of one of the referent animals. When the visual scene was accessible again, infants looked longer at the event outcome (the animal on top of the inanimate object), relative to the other animal on the screen, when this outcome did not match the state of affairs described by the prior language. When the event outcome was congruent with the prior language, looking at the outcome event was at chance level. These results indicate that infants used the language input to generate an expectation of the described outcome and that they were surprised when the outcome did not match that input. This effect was not related to infants' age or vocabulary size.

These findings add to an increasing body of research showing that from the earliest stages of language development, children have the capacity to treat language as a source of information about changes in the world that they have not witnessed firsthand. They can use verbal testimony to update their representation of the world.

During the first half of their second year of life, infants understand that words can refer to absent objects (Ganea, 2005; Ganea & Saylor, 2013a; Osina, Saylor, & Ganea, 2013, 2014; Saylor & Baldwin, 2004) and that words can refer not only to particular objects but also to general kinds (Gelman, Hollander, Star, & Heyman, 2000; Graham, Booth, & Waxman, 2012; Waxman & Gelman, 2010). By 18 months of age, infants can also use language about a change in the properties or location of an absent object to subsequently re-identify or search for the object (Galazka & Ganea, 2014; Ganea & Harris, 2013). Nevertheless, if an object was originally encoded in a specific location (in the drawer rather than elsewhere in the room), it is not until approximately 30 months of age that infants show evidence of updating by searching systematically in the new location rather than in the old location (Ganea & Harris, 2010, 2013).

The current research shows that even at 16 months of age, infants can use language to form expectations about an object's current location in a visual scene. Infants' looking times to the event outcome were significantly longer when the visual scene did not match the language they received about the scene in its absence than when it did. Infants' responses at test were not simply due to having observed any change in the new visual scene relative to the already encoded visual scene because the visual scene presented at test was changed from the familiarization phase in both the Congruent and Incongruent conditions. More specifically, the scene always showed one of the two animals having moved from the side onto the object at the center. Thus, the language received in the absence of

the scene influenced infants' expectation of the outcome in the scene (with one of the referent animals on top of the object rather than in its original location).

The language used in this study predicted a relation between two of the referent objects in the scene ("Now the dog goes to the table. The dog is on the table") and, as a result, a change in the location of one of the referents (the dog). As children heard the sentence, did they mentally represent the dog's prior location in the previously visible scene and then update their representation of the dog's location based on the language input? Alternatively, did infants create a new representation of the scene (of a dog on the table) at the time they heard the language input with no updating of the prior representation involved? In both cases, infants would have been surprised when the visual input did not match the language they heard when the occluder was lowered.

One reason to prefer the first alternative is that studies with adults using the blank screen paradigm have shown that when a scene is absent, adults look where a mentioned object had been visible and update the representation of the scene as the language is being processed (Altmann & Kamide, 2004). The data during the updating period in our study (when the scene was no longer visible) show that infants looked at the location where the table or bed had been visible significantly longer than at the prior locations of the animals, but the proportion of looks to the mentioned agent was not higher than that to the unmentioned agent overall (Fig. 4), and our time course analysis demonstrated that it did not significantly increase around the time of the mentions (Fig. 3). This implies that infants did not process the testimony in a fully adult-like way; that is, unlike adults, they did not look at the mentioned animal's placeholder while the scene was occluded. Yet this is not surprising in light of the evidence that toddlers do not display adult-like gaze behavior during sentence processing until they are approximately 8 months older and have a sizable (>225 words) productive vocabulary (Mani & Huettig, 2012). The largest productive vocabulary among the infants in our study was a mere 66 words.

Studies on infants' understanding of references to absent objects show that, beginning at 12 months of age, infants show comprehension of references to absent objects by looking, pointing, or searching at the object's likely location on hearing its name mentioned (Ganea, 2005; Ganea & Saylor, 2013b; Osina et al., 2013, 2014; Saylor & Baldwin, 2004). Infants are less likely to engage in these communicative behaviors toward absent referents if they hear an unrelated name (Ganea, 2005). These results make it unlikely that the infants in the current study interpreted the sentence afresh rather than as referring to the visual scene that they had seen only a few seconds prior to the language input. Nevertheless, further research with infants is needed to demonstrate whether the incremental language processing skills shown in adults emerge at the earliest stages of language development. Even though children may activate a representation of the absent referent (for a review of the evidence, see Ganea & Saylor, 2013a), it is not clear that at this age the timing of this activation happens incrementally as the language is being processed.

Another important question regarding infants' processing of absent reference concerns the nature of the information extracted from the language that infants heard when the scene was not visible. Research with older infants shows that by 24 months of age infants can use semantic information that is language derived (e.g., whether the verb is transitive or intransitive) to determine the specific type of event that a sentence describes (Arunachalam & Waxman, 2010). The language we used here specified the spatial relation between the animal and its destination (e.g., that the dog was on the table). However, it was not possible to determine exactly what information was extracted from this language and, therefore, how specific infants' representation of the relation was.

As a baseline, it is clear that infants extracted the agent of the change (either cat or dog) from the testimony because that formed the critical distinction between congruent and incongruent testimony. However, we might expect the same pattern of results if the agent was *all* that the participants had extracted from the testimony. In other words, if infants extracted only the dog from the sentence, they may have expected any salient change involving the dog to be a possible outcome once the occluder was lifted; this would yield longer looking when the location change involved the cat instead. Likewise, we do not know whether children would have been surprised, and thus looked longer, if the scene outcome showed the dog "under" the table rather than "on" the table. Additional experiments are already under way in our lab varying the amount of information participants are able to extract from the testimony as well as the degree to which the outcome scene is congruent with the testimony.

These will help to determine whether infants' semantic knowledge and processing of semantic links between words in a sentence constrains their representation of the outcome event.

Our findings, together with those of others (Gliga & Csibra, 2009; Graham et al., 2004; Vouloumanos, Onishi, & Pogue, 2012; Xu et al., 2005), show that from the early stages of language development infants take language to provide relevant information about entities in the world even when those entities are not visible. More specifically, we have shown here that infants as young as 16 months have the capacity to update their expectations about the location of an object based on information they hear in the absence of the object. This finding suggests that older children's difficulty in relying on language to search for an object in a new location in prior updating studies (Ganea & Harris, 2010, 2013) may be due to difficulty in using an updated representation to guide their search for the object, which is related to working memory limitations shown to play a role in children's ability to update (Ozdemir, Gallant, & Ganea, 2015). It is also possible that the level of conflict here between the old location and the new location of the target animal was not as strong as in previous research (Ganea & Harris, 2010, 2013) given that the initial location of the target animal in the current paradigm was not highlighted in any special way. Further research is needed to probe the sophistication of children's ability to process linguistic information in the absence of a visual scene and to update their mental representation of such scenes based on new input.

### Acknowledgments

We thank Marisa Biondi, Annalisa Groth, and other members of the UMass Boston Baby Lab for their help with data collection. The research presented here and the preparation of this manuscript were supported by a Seed Grant from the Simons Foundation under the auspices of the Simons Center for the Social Brain at MIT (Grant 319294) to Zsuzsa Kaldy and an Insight Grant from the Social Sciences and Humanities Council in Canada to Patricia Ganea.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jecp.2015.12.005>.

### References

- Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). New York: Psychology Press.
- Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*, 442–446.
- Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on A-not-B. *Child Development*, *56*, 868–883.
- Feigenson, L., & Yamaguchi, M. (2009). Limits on infants' ability to dynamically update object representations. *Infancy*, *14*, 244–262.
- Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Science*, *12*, 405–410.
- Galazka, M. A., & Ganea, P. A. (2014). The role of representational strength in verbal updating: Evidence from 19- and 24-month-olds. *Journal of Experimental Child Psychology*, *121*, 156–168.
- Ganea, P. A. (2005). Contextual factors affect absent reference comprehension in 14-month-olds. *Child Development*, *76*, 989–998.
- Ganea, P. A., & Harris, P. L. (2010). Not doing what you are told: Early perseverative errors in updating mental representations via language. *Child Development*, *81*, 457–463.
- Ganea, P. A., & Harris, P. L. (2013). Early limits on the verbal updating of an object's location. *Journal of Experimental Child Psychology*, *114*, 89–101.
- Ganea, P. A., & Saylor, M. M. (2013a). Representational constraints on language development: Thinking and learning about absent things. *Child Development Perspectives*, *7*, 227–231.
- Ganea, P. A., & Saylor, M. M. (2013b). Talking about the near and dear: Infants' comprehension of displaced speech. *Developmental Psychology*, *49*, 1299–1307.
- Ganea, P. A., Shutts, K., Spelke, E. S., & DeLoache, J. S. (2007). Thinking of things unseen: Infants' use of language to update mental representations. *Psychological Science*, *18*, 734–739.
- Gelman, S. A., Hollander, M., Star, J., & Heyman, G. D. (2000). The role of language in the construction of kinds. In D. Medin (Ed.), *Psychology of learning and motivation* (Vol. 39, pp. 201–263). San Diego: Academic Press.
- Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*, 347–353.

Please cite this article in press as: Ganea, P. A., et al. Sixteen-month-olds can use language to update their expectations about the visual world. *Journal of Experimental Child Psychology* (2016), <http://dx.doi.org/10.1016/j.jecp.2015.12.005>

- Graham, S. A., Booth, A. E., & Waxman, S. R. (2012). Words are not features of objects: Only consistently applied nouns guide 4-year-olds? Inferences about object categories. *Language Learning and Development*, 8, 1–11.
- Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development*, 75, 409–427.
- Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Belknap Press/Harvard University Press.
- Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numerical transformations in five-month-old infants. *Mathematical Cognition*, 3, 89–104.
- Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 843–847.
- Morford, J. P., & Goldin-Meadow, S. (1997). From here and now to there and then: The development of displaced reference in homesign and English. *Child Development*, 68, 420–435.
- Osina, M., Saylor, M. M., & Ganea, P. A. (2013). When familiar is not better: 12-month-olds respond to talk about absent objects. *Developmental Psychology*, 49, 138–145.
- Osina, M. A., Saylor, M. M., & Ganea, P. A. (2014). Object locations, identity, and absent reference understanding at 12 months. *Infancy*, 19, 65–81.
- Ozdemir, B., Gallant, H., & Ganea, P. A. (2015). Do working memory and inhibition explain toddlers' ability to verbally update mental representations? In: Poster presented at the 2015 biennial meeting of the Society for Research in Child Development, Philadelphia, PA.
- Saylor, M. M., & Baldwin, D. A. (2004). Discussing those not present: Comprehension of references to absent caregivers. *Journal of Child Language*, 31, 537–560.
- Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge. *Cognitive Development*, 14, 1–36.
- Vouloumanos, A., Onishi, K. H., & Pogue, A. (2012). Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 12933–12937.
- Waxman, S. R., & Gelman, S. A. (2010). Different kinds of concepts and different kinds of words: What words do for human cognition. In D. Mareschal, P. Quinn, & S. Lea (Eds.), *The making of human concepts* (pp. 101–130). Oxford, UK: Oxford University Press.
- Wellman, H., Cross, D., & Bartsch, K. (1986). Infant search and object permanence: A meta-analysis of the A-not-B error. *Monographs of the Society for Research in Child Development*, 51 [3, Serial No. 214].
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749–750.
- Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, 16, 372–377.