

# Toddlers with Autism Spectrum Disorder Can Use Language to Update Their Expectations About the World

Allison Fitch<sup>1,3</sup>  · Annalisa Valadez<sup>1</sup> · Patricia A. Ganea<sup>2</sup> · Alice S. Carter<sup>1</sup> · Zsuzsa Kaldy<sup>1</sup>

© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

This study examined if two-year-olds with ASD can update mental representations on the basis of verbal input. In an eye-tracking study, toddlers with ASD and typically-developing nonverbal age-matched controls were exposed to visual or verbal information about a change in a recently encoded scene, followed by an outcome that was either congruent or incongruent with that information. Findings revealed that both groups looked longer at incongruent outcomes, regardless of information modality, and despite the fact that toddlers with ASD had significantly lower measured verbal abilities than TD toddlers. This demonstrates that, although there is heterogeneity on the individual level, young toddlers with ASD can succeed in updating their mental representations on the basis of verbal input in a low-demand task.

**Keywords** Representations · Updating · Receptive language · Language comprehension

## Introduction

The amount of information that any individual can learn about the world through direct observation is limited. We cannot, for example, directly observe the United States Civil War, or see molecules (without special equipment). To learn about such entities, we rely on verbal reports from other people who communicate this information to us; we then store that information as a set of mental representations that dictate our expectations about the world and guide our behavior. Verbal testimony also provides us with the means to update our existing mental representations to reflect the world's most current state. Failure to update mental representations can lead to problems in navigating the world, as demonstrated by young infants who produce perseverative errors in the classic A-not-B task (Diamond 1990; Piaget 1954).

✉ Allison Fitch  
allison.fitch001@umb.edu

<sup>1</sup> Department of Psychology, University of Massachusetts Boston, 100 William T. Morrissey Boulevard, Boston, MA, USA

<sup>2</sup> Eric Jackman Institute of Child Study, University of Toronto, Toronto, Canada

<sup>3</sup> Present Address: Speech, Language & Hearing Sciences Department, Boston University, Boston, USA

In typically developing (TD) toddlers, the ability to build and update representations based on verbal testimony has been demonstrated at an early age. Before the end of their second year, TD toddlers know that language can refer to absent entities (Ganea 2005; Saylor and Baldwin 2004), and can use language to generate expectations about the quantity and animacy of unseen objects (Ferguson et al. 2014; Xu et al. 2005). TD toddlers can also extract information from verbal testimony to update previously held expectations. After hearing about an object's feature change (a stuffed animal had become wet), toddlers reliably select the changed object, when asked for it by name, over an unchanged target object or a distracter (Galazka and Ganea 2014; Ganea et al. 2007). In a recent eyetracking study of verbal updating 16-month-olds showed a reliable violation-of-expectation (VoE) when they heard a false testimony (Ganea et al. 2016). In this study, infants were presented with a scene depicting a cat, bed, and dog, a screen was lowered, and infants heard that one of the animals had changed its location (e.g. *Now the cat is on the bed!*). Infants were then presented with an incongruent outcome in which the cat had stayed in place and the dog had moved instead, or a congruent one. Infants looked reliably longer at the incongruent outcome, suggesting that they had updated their expectations about the absent visual scene in accordance with the information provided in the testimony.

Using verbal testimony alone to update a representation is a complex interplay of linguistic processing and working memory (Ganea and Harris 2013). Once the testimony is parsed (no small task for a toddler with a developing lexicon), it is used as a retrieval cue for an existing mental representation, requiring a robust link between the representation and the linguistic input. That representation is then manipulated to incorporate new information in the absence of visual support. Although verbal updating emerges early in development, even TD toddlers show limits in their ability to update efficiently, depending on the level of conflict between their original representation and the new verbal input (Ganea and Harris 2010, 2013). For toddlers with autism spectrum disorder (ASD) who often present broad language delays (Lazenby et al. 2016; Luyster et al. 2008; Paul et al. 2008), verbal updating may pose a challenge even in low-level demand contexts. The first mental operation involved in verbal updating, understanding that the verbal input refers to absent referents, requires an understanding that words are referential. Evidence from infants at high-risk for ASD (Curtin and Vouloumanos 2013) as well as older, lower-functioning children with ASD (Hartley and Allen 2015b; Preissler 2008), suggests that children with ASD may lack such an understanding. This is in contrast to TD infants, who understand that speech is referential from early in life (Ganea and Saylor 2013; Vouloumanos et al. 2012, 2014). Recent research with children with ASD suggests that understanding of words as referential symbols, is related to language production, and that children with language delays may develop this understanding, just later than expected (Hartley et al. 2017). Thus, although to our knowledge no one has explicitly tested the acquisition of absent referent understanding in ASD, we would predict that this crucial first step to updating is likely to be affected.

These group level findings should be considered with caution, however, as the autism spectrum encompasses significant variability in language phenotypes (Kjelgaard and Tager-Flusberg 2001). For example, absent referent comprehension relies on having a robust link between the word form and the semantic meaning of the absent referent (Ganea and Saylor 2013). Findings on the semantic representations of children with ASD are mixed (see Arunachalam and Luyster 2016; Naigles and Tek 2017, for reviews). Some studies have demonstrated typical semantic representations in ASD, such as those examining children's ability to appropriately categorize familiar exemplars (Tager-Flusberg 1985; Ungerer and Sigman 1987), and those demonstrating similarities between TD children and children with ASD on vocabulary composition (Charman et al. 2003; Weismer et al. 2011; Luyster et al. 2007). Others have demonstrated atypical development of the linguistic biases associated with word-meaning links. For example, shape, which defines early semantic representations, does not constrain word meaning for young toddlers

with ASD (Hartley and Allen 2014; Tek et al. 2008). Additionally, children with poorer language skills rely on associationist links between words and pictures they are introduced with (Preissler 2008). Some have suggested that difficulties mapping word forms to semantic meaning in ASD are the result of disruptions in the cognitive systems that support word learning, such as attention and memory (Arunachalam and Luyster 2016; Luyster and Lord 2009). If true, this predicts that high-functioning children with ASD should have typical semantic networks. Regardless, we can infer that many children with ASD do develop absent referent comprehension as evidenced by the types of concepts involved in some common restricted interests. For example, physics-related concepts such as quantum mechanics, as well as time and schedules, and historical events are all widely-reported topics of restricted interest for individuals with ASD (Klin et al. 2007; South et al. 2005). These abstract entities are inherently unobservable, and thus always serve as absent referents in conversation.

To conduct the second mental operation involved in verbal updating, the ability to manipulate the representation once it has been retrieved, one must be able to incorporate new information from verbal input into an existing representation. To our knowledge, no studies examine this ability per se, but studies examining the use of linguistic information to generate new representations in children with ASD have been conducted, and findings are mixed. For example, children with ASD are unimpaired in their use of mutual exclusivity in word learning, and are thus able to use known language to map a novel name to a novel object (de Marchena et al. 2011; Preissler and Carey 2005). They are also able to use linguistic knowledge to map events to novel verbs via syntactic bootstrapping, akin to typically-developing peers (Naigles et al. 2011; Swensen et al. 2007). However, children with ASD have enduring difficulties with other types of language-based inferences, such as categorical induction, where they do not extend categories to include new exemplars based on shared labels (Kelley et al. 2006).

It is worth noting that all of these tasks, that is, mutual exclusivity, syntactic bootstrapping, and categorical induction, require more inference than is necessary for verbal updating—oftentimes the testimony is explicit and no inference is required. However, these studies also provided participants with a supportive visual context, where candidate referents were visible throughout the task. Successes may not extend to updating expectations about absent referents.

In contrast to verbal updating, we expect that using visual information to update representations will be a relative strength for individuals with ASD. Many adults with ASD report that it is easier to understand a picture than a sentence (e.g. Grandin 1995). Indeed, enhanced performance has been observed on tasks such as Block Design, in which a pattern is recreated from square blocks, Embedded Figures, in which

a shape is located within a complex figure (Happé and Frith 2006), and Visual Search, in which a target is found among a set of distractors (Plaisted et al. 1998). This superiority is present early: 2-year-olds with ASD are better at a conjunctive feature visual search than age-matched TD toddlers (Kaldy et al. 2011). Further, one study recently demonstrated that compared to TD children, children with ASD were equally good at generating new expectations about a hidden object's location using visual information, a picture (Hartley and Allen 2015a). These findings point to a potential visual advantage in toddlers with ASD, who may be more efficient at learning from visual than verbal information.

Taken together, it is unclear how toddlers with ASD will perform in a task that requires using verbal input to update a representation that is visual in nature. We expect that their ability to update representations of the visual world using visual information is unimpaired, akin to other visual cognitive skills. However, it is not clear if, at this age, toddlers with ASD will (1) have a robust enough link between a word form and its referent to retrieve a representation of an absent referent, and (2) be able to use that linguistic information to update an existing visual representation.

This study tests if toddlers with ASD can use verbal information to update expectations about the visual world. We adapted Ganea et al. (2016) eye-tracking paradigm, with an added manipulation: the information needed to update was presented either verbally (as in Ganea et al. 2016) or visually. If toddlers with ASD are not able to use language to generate expectations at this age, they will not show a VoE when information is presented verbally. However, assuming there are no differences in visual updating between groups, they should still demonstrate this effect in the visual updating condition.

## Method

### Participants

Thirty-eight TD toddlers from the Greater Boston area (age range 15.0–22.0 months) participated in the study. Forty-seven toddlers, who were referred by local early intervention centers for concerns about ASD, were invited to participate immediately prior to a diagnostic evaluation for ASD (age range 18.0–36.0 months). Among the TD participants, five were excluded for failing an ASD screener (see "Inclusion Criteria" section), and seven were excluded due to fussiness or inattention. Among the ASD concerns group 11 did not meet criteria for ASD, and 14 were excluded due to fussiness or inattention. The final sample was comprised of 26 TD toddlers (13 females, mean age: 18.5, range 15.19–21.82) and 22 toddlers with ASD (1 female, mean age: 28;23, range 20;29–35;18); detailed participant characteristics are

**Table 1** Participant characteristics

	TD	ASD	p value
N	26	22	
No. of females	13	1	
Age in months (days)	18;15 (1;11)	28;23 (5;04)	<0.001
Ethnicity			
Hispanic or Latino	6	6	
Not Hispanic or Latino	20	16	
Race			
Asian	0	1	
African-American	0	5	
White	16	7	
Other/multiple/no response	6	9	
Mullen scales			
Visual reception	21.96 (3.24)	21.54 (4.68)	ns
Fine motor	26.96 (14.7)	21.00 (3.49)	0.055
Receptive language	19.73 (4.29)	12.86 (3.71)	<0.001
Expressive language	17.65 (1.98)	15.36 (4.80)	0.046
ELC	100.27 (14.75)	59.77 (8.62)	<0.001
ADOS-2			
Social affect	—	14.85 (4.15)	
Restricted rep. beh	—	5.26 (1.53)	
Total	—	20.23 (3.8)	
Calibrated severity score	—	8.0 (1.74)	

Standard deviations are reported in parentheses, following the means, where appropriate. All MSEL Subscale Scores are raw scores for matching purposes

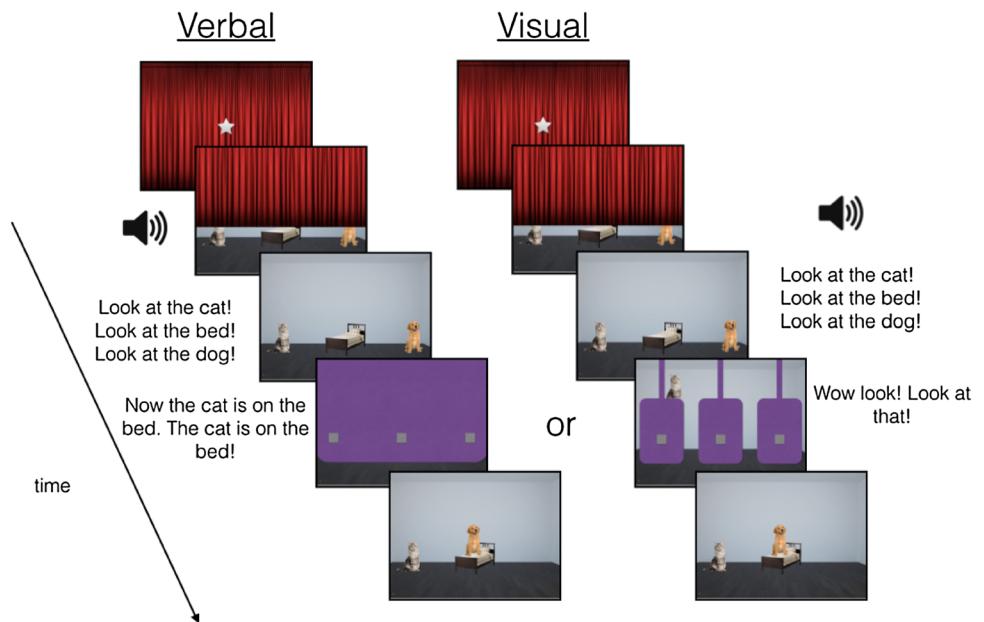
ELC early learning composite

in Table 1. All participants were assessed on the cognitive scales of the Mullen Scales for Early Learning (MSEL; Mullen 1995). Caregivers filled out a brief demographic questionnaire, as well as two autism screening tools: the Brief Infant Toddler Social Emotional Assessment (BITSEA; Briggs-Gowan et al. 2004; Kiss et al. 2017), and the Parent's Observation of Social Interaction (POSI; Smith et al. 2013). They were also asked to report their child's comprehension of critical words in the study (*cat*, *dog*, and *bed*). Importantly, the two groups did not differ on their nonverbal scores on the MSEL<sup>1</sup> (a proxy for nonverbal mental age), suggesting that any differences in group performance on the task are unlikely to be related to nonverbal components of updating in general, such as object recognition or working memory.

English was the majority language spoken at home for all children, except for seven in the TD group and eight in the

<sup>1</sup> The marginally significant difference between groups on fine motor scores can be attributed to a strong, positively skewed distribution of scores on this scale, in which four TD participants scored extremely high (> 50).

**Fig. 1** Schematic of incongruent test trials in each condition



ASD group, who all came from bilingual homes (of which one of those languages was English). The bilingual children did not differ from monolingual children in the same group on the Receptive [ $t(46)=0.248, p=.81$ ] or Expressive Language scales [ $t(46)=1.26, p=.21$ ] of the MSEL. Children in the ASD group all received early intervention services in English. For the TD group, all participants met age-appropriate basal scores on the Receptive Language scale in English.

## Inclusion Criteria

Inclusion in the TD group required a negative screen for ASD on the BITSEA and POSI, and no history of developmental concerns. ASD diagnoses were confirmed using the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al. 2012) and a licensed psychologist's clinical impression. Those who did not meet criteria for an ASD diagnosis were excluded from analysis.

## Stimuli

Stimuli were animated stock images that included two pairs of agents (two cats and two dogs), and a bed that served as a central location for the animals to move to. The agents and location were set in a computer-generated room background that included three walls and a floor (Fig. 1).

## Apparatus and Procedure

Participants sat on their caregiver's lap in a dimly lit room, ~60 cm from the monitor of a Tobii T120 eyetracker, sampling at 60 Hz. Caregivers wore a visor that occluded their

vision from the screen. Participants were presented with three blocks of trials: first, a block of six familiarization trials, followed by two blocks of test trials, one in which new information was visual, and one in which new information was verbal. Each test block contained two test trials: one with a congruent outcome and one with an incongruent outcome, for a total of four test trials.<sup>2</sup> The order of trials within each test block was randomized. The order of test blocks was counterbalanced such that half of the participants received the visual condition first; the other half received the verbal condition first. A standard 5-point calibration was conducted before each block. Trials had a gaze-contingent onset: the trial began as soon as the participant made a fixation at a central attention-getter (spinning star), or after 30 s (this occurred only twice during testing), whichever came first.

In familiarization trials, stimuli were presented and labeled by a pre-recorded female voice through a central speaker. In the first trial, a curtain was raised to reveal two different cats. The cats were labeled in child-directed speech: *Look at the cats!* The cats then jumped twice and the curtain lowered again. This was repeated for two different dogs, as well as the bed (which jiggled, rather than jumped). This whole sequence was presented twice for a total of six trials, totaling 38 s.

Both types of test trials (Fig. 1) began with an *encoding phase* during which the curtain was raised to reveal a cat and a dog on either side of a bed at the centerpoint.

<sup>2</sup> Due to concerns about fussiness-related data loss, we reduced the number of trials from Ganea et al. (2016), in line with their analyses demonstrating the same effects when only looking at the first trial in each block as when averaging across the whole block.

The pre-recorded voice then labeled each stimulus from left to right. In the *occlusion phase* (8 s), an occluder was lowered to cover the stimuli, with three identical squares in the locations where each stimulus had been. As in Ganea et al. (2016), the Verbal Condition contained a single large occluder that left only the floor visible. During occlusion, participants heard the pre-recorded voice provide testimony about one of the animals moving to the bed (*Now the dog/cat is on the bed! The dog/cat is on the bed!*). In the Visual Condition, three separate, narrow occluders (each containing one square), left visible not only the floor, but the wall space around them. During occlusion, one of the animals jumped up from behind its occluder and disappeared behind the central occluder, with the path fully visible to the participant, while the pre-recorded voice said: *Wow, look! Look at that!* In the *outcome phase* (8 s), the occluder(s) was lifted to reveal one animal on the bed, and the other in its original position. In Congruent trials, the animal on the bed is congruent with information provided in the occlusion phase (either mentioned in the testimony or seen move). In Incongruent trials, the opposite animal is on the bed. The specific animal (which dog and cat), the side each animal was placed on, and which animal was mentioned in the testimony were counterbalanced across the experiment. The entire testing session lasted approximately 3 min, 7 s.

## Data Analysis

To analyze gaze patterns, we defined a 255 by 231 pixel rectangular area of interest (AOI) around the agent in the new location during the outcome period, as well as the entire screen during the outcome phase (overall outcome), as in Ganea et al. (2016). Total looking time to each AOI (defined by the cumulative looking time within the AOI during the outcome phase, in seconds) as well as the number of fixations made to the screen were analyzed across condition and outcome type. Fixations were determined using Tobii's I-VT fixation filter. Participants who did not provide sufficient data (attention to the screen for at least 1 s, the same criterion as used in Ganea et al. 2016) in all four test trials were excluded for fussiness or inattention (21 participants).

## Results

### Language Ability

On average, TD toddlers' MSEL T-scores (scores standardized for age) were within the normal range (Receptive:  $M = 49.34$ ,  $SD = 15.21$ , Expressive:  $M = 48.76$ ,  $SD = 5.54$ ) while ASD toddlers' scores were, on average, considered "Very Low" (Receptive:  $M = 21.22$ ,  $SD = 4.29$ , Expressive:  $M = 27.27$ ,  $SD = 6.42$ ). As a group, TD toddlers' raw

**Table 2** Comprehension of specific vocabulary items by number of participants

Item	TD comprehends (n = 26)	ASD comprehends (n = 22)
Cat	22 (84.6%)	10 (45.5%)
Dog	23 (88.4%)	10 (45.5%)
Bed	16 (61.5%)	9 (40.9%)

scores on the Receptive and Expressive scales of the MSEL were significantly higher than those of toddlers with ASD,  $t(46) = 5.872$ ,  $p < .001$ ,  $d = 1.71$ ,  $t(26.989)^3 = 2.090$ ,  $p = .046$ ,  $d = 0.62$  (respectively). Parent reports on the comprehension of the words *cat*, *dog*, and *bed* are reported in Table 2. Comprehension of both *cat* (Fisher's exact test,  $p = .006$ , two-sided) and *dog* (Fisher's exact test,  $p = .002$ , two-sided) were significantly associated with diagnostic group. Nearly all TD toddlers held both the words *cat* and *dog* in their receptive vocabulary, while less than half of the toddlers with ASD did. Most participants comprehended at least one agent word (either *cat* or *dog*). Only 13 participants (27% of the sample, 2 TD 11 ASD) reportedly did not understand either agent word. *Bed* was reportedly less well-known among all participants (41% of ASD toddlers and 62% of TD toddlers), and was not related to diagnostic group (Fisher's exact test,  $p = .25$ , two sided).

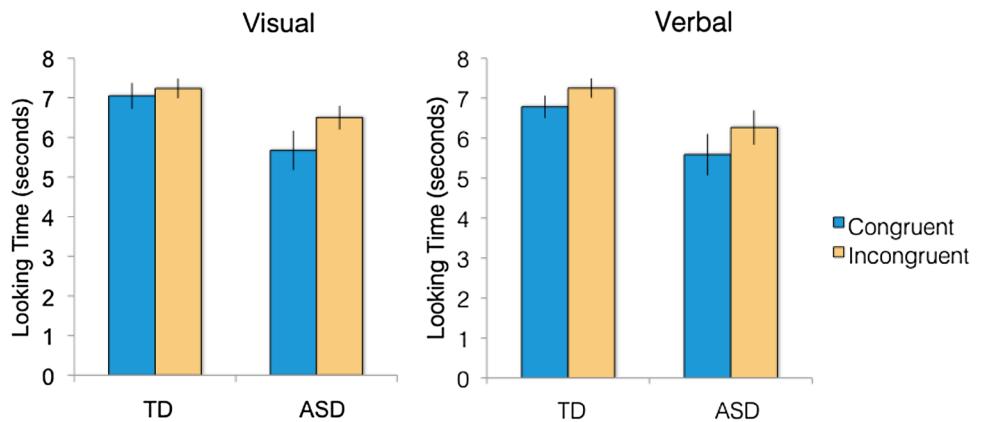
### Eyetracking Analysis

For each participant, three variables of interest were calculated: total looking time to the screen and the number of fixations made to the screen during the outcome phase, and the proportion of looking time to the animal in the changed location (i.e. the animal on the bed). The proportion variable was calculated by dividing looking time to the animal in the new location by total looking to the outcome on each trial. Because data did not conform to the normal distribution, analyses were repeated after performing a logarithmic transformation (see Csibra et al. 2016). Results did not differ from the untransformed results reported here.

To identify a VoE effect, we conducted a 2 (diagnosis: TD vs. ASD)  $\times$  2 (condition: visual vs. verbal)  $\times$  2 (congruency: congruent vs. incongruent) ANOVA on looking time to the overall outcome (Fig. 2). Findings showed a significant main effect of diagnostic group,  $F(1,46) = 12.23$ ,  $p = .001$ ,  $\eta_p^2 = 0.21$ , with the ASD group having shorter looking times ( $M = 6.01$  s,  $SD = 1.06$ ) than the TD group ( $M = 7.08$  s,  $SD = 1.06$ ). There was also a significant main effect of congruency,  $F(1,46) = 6.31$ ,  $p = .016$ ,  $\eta_p^2 = 0.12$ ,

<sup>3</sup> Adjusted df in accordance with a significant Levene's test.

**Fig. 2** Looking time to outcome by group, condition, and congruency. Error bars represent standard error



where participants looked consistently longer at the incongruent ( $M=6.81$  s,  $SD=1.01$ ) than the congruent outcome ( $M=6.27$  s,  $SD=1.49$ ), regardless of condition or group. There was no main effect of condition,  $F(1,46)=0.262$ ,  $p=.611$ , or significant interactions between diagnosis and either condition or congruency.

This finding held when the same analysis was limited to just the AOI of the animal in the changed location (in other words, was the longer looking just to the screen in general, or was it specific to the animal that changed location?). Findings again demonstrated a main effect of diagnostic group,  $F(1,46)=9.56$ ,  $p=.003$ ,  $\eta_p^2=0.17$ , with the ASD group having overall shorter looking times to the animal that changed location ( $M=2.81$ ,  $SD=0.91$ ) than the TD group ( $M=3.62$ ,  $SD=0.91$ ). The main effect of congruency was also significant,  $F(1,46)=4.59$ ,  $p=.038$ ,  $\eta_p^2=0.09$ , with both groups looking longer to the animal that changed location in the incongruent outcome ( $M=3.40$ ,  $SD=1.04$ ), than the congruent outcome ( $M=3.02$ ,  $SD=1.15$ ). Here, the main effect of modality was marginally significant  $F(1,46)=3.92$ ,  $p=.054$ ,  $\eta_p^2=0.08$ . Participants looked at the animal that changed location marginally less in the visual condition ( $M=3.023$ ,  $SD=1.04$ ), than the verbal condition ( $M=3.42$ ,  $SD=1.30$ ), regardless of outcome.

To examine if participants searched the scene more when presented with incongruent than congruent outcomes, we conducted a 2 (diagnosis)  $\times$  2 (condition)  $\times$  2 (congruency) ANOVA on number of fixations to the outcome. Findings again showed significant main effects of diagnosis,  $F(1,46)=4.59$ ,  $p=.037$ ,  $\eta_p^2=0.091$ , where the TD toddlers ( $M=13.01$ ,  $SD=2.67$ ) made more fixations to the outcome overall than the toddlers with ASD ( $M=11.35$ ,  $SD=2.67$ ), and congruency,  $F(1,46)=5.271$ ,  $p=.026$ ,  $\eta_p^2=0.103$ , where participants made more fixations to the incongruent ( $M=12.72$ ,  $SD=2.88$ ), relative to the congruent outcome ( $M=11.64$ ,  $SD=3.24$ ). There was no significant main effect of visual versus verbal condition,  $F(1,46)=1.617$ ,  $p=.21$ . There was a marginally significant interaction between condition and congruency,  $F(1,46)=2.85$ ,  $p=.09$ ,  $\eta_p^2=0.06$ ,

which suggests this effect was driven by a large difference between the congruent ( $M=10.72$ ,  $SD=4.32$ ) and incongruent trials ( $M=12.85$ ,  $SD=4.33$ ) in the verbal condition,  $t(47)=-2.788$ ,  $p=.008$ , and not by the differences in the visual condition [congruent:  $M=12.55$ ,  $SD=4.23$ , incongruent:  $M=12.59$ ,  $SD=4.19$ ;  $t(47)=-0.026$ ,  $p=.98$ ].

We also explored the possibility that participants might allocate their attention around the screen differently across conditions (as was the case in Ganea et al. 2016). We conducted a third ANOVA on proportion of looking time to the animal in the changed location relative to the entire screen. For this analysis, there was neither a main effect of diagnosis,  $F(1,46)=2.18$ ,  $p=.15$ , nor congruency,  $F(1,46)=0.968$ ,  $p=.33$ . However, there was a significant main effect of condition,  $F(1,46)=6.14$ ,  $p=.017$ ,  $\eta_p^2=0.12$ , where all participants looked longer to the animal in the changed location relative to the entire screen when the information was verbal ( $M=0.52$ ,  $SD=0.16$ ) than when it was visual ( $M=0.45$ ,  $SD=0.14$ ).

We went on to examine the relationship between developmental factors and task performance by conducting a series of correlations. First, we explored possible relations between age, overall cognitive ability, verbal MSEL scores, and the magnitude of the verbal VoE effect (looking time in verbal incongruent minus looking time in verbal congruent outcomes). We expected that children with better language skills would process the verbal testimony better, and would show larger VoE effects. No significant relations were found between age ( $r=.11$ ,  $p=.46$ ), MSEL ELC scores ( $r=-.033$ ,  $p=.82$ ), receptive language ( $r=-.004$ ,  $p=.98$ ), or expressive language ( $r=.225$ ,  $p=.13$ ), and the verbal VoE magnitude. When run separately on each group, these relations were all nonsignificant (see Tables 3, 4). We also examined the relations between task performance and autism-specific characteristics. We found no significant correlations between VoE magnitude and scores on the ASD problems subscale of the BITSEA ( $r=.075$ ,  $p=.62$ ), nor the ASD competence subscale ( $r=.066$ ,  $p=.66$ ). We also did not find a significant relationship between VoE magnitude and ADOS calibrated

**Table 3** Correlation table describing relationships between VoE and developmental factors for the TD group

	Verbal VoE	Visual VoE	Age	MSEL ELC	MSEL RL	MSEL EL	BITSEA-Prob	BITSEA-Comp
Verbal VoE	–	–	–	–	–	–	–	–
Visual VoE	– 0.067	–	–	–	–	–	–	–
Age	– 0.016	– 0.334	–	–	–	–	–	–
MSEL ELC	0.082	0.119	– 0.012	–	–	–	–	–
MSEL RL	0.043	0.096	0.134	0.875***	–	–	–	–
MSEL EL	0.192	– 0.073	0.46*	0.654***	0.633**	–	–	–
BITSEA-Prob	– 0.004	0.466*	0.036	0.002	0.105	0.109	–	–
BITSEA-Comp	0.094	– 0.108	– 0.083	0.486*	0.424*	0.131	– 0.054	–

RL receptive language, EL expressive language, BITSEA-Prob ASD Problems score, BITSEA-Comp ASD Competence score

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

**Table 4** Correlation table describing relationships between VoE and developmental factors for the ASD group

	Verbal VoE	Visual VoE	Age	MSEL ELC	MSEL RL	MSEL EL	ADOS-SA	ADOS-RRB
Verbal VoE	–	–	–	–	–	–	–	–
Visual VoE	0.132	–	–	–	–	–	–	–
Age	0.136	0.496*	–	–	–	–	–	–
MSEL ELC	– 0.013	– 0.225	– 0.193	–	–	–	–	–
MSEL RL	0.055	0.18	0.354	0.443*	–	–	–	–
MSEL EL	0.282	0.202	0.671**	0.223	0.57**	–	–	–
ADOS-SA	0.042	– 0.25	– 0.284	– 0.344	– 0.477*	– 0.608**	–	–
ADOS-RRB	0.306	0.179	0.285	-0.004	0.076	0.221	0.534*	–

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

severity scores in either the Social Affect ( $r = .042$ ,  $p = .85$ ) or Restricted and Repetitive Behaviors (RRB) subscales ( $r = .306$ ,  $p = .17$ ) in the ASD group.

We also examined the role of word-specific knowledge, that is, whether the participant reportedly comprehended either agent word or not, on looking time. To do this, we conducted a 2 (diagnosis)  $\times$  2 (congruency)  $\times$  2 (knowing at least one agent) ANOVA on looking time in the verbal condition only (see Fig. 3). Here, only the interaction between the last two factors was marginally significant,  $F(1,44) = 3.38$ ,  $p = .073$ ,  $\eta_p^2 = 0.07$ . Participants who reportedly knew at least one agent (*cat* or *dog*) looked longer in the incongruent outcome ( $M = 7.3$ ,  $SD = 1.04$ ) than the congruent outcome [ $M = 6.54$ ,  $SD = 1.71$ ;  $t(34) = -2.88$ ,  $p = .007$ ], while participants who reportedly did not know either agent word looked for similar amounts of time in the congruent ( $M = 5.40$ ,  $SD = 2.59$ ) and incongruent ( $M = 5.45$ ,  $SD = 2.35$ ) conditions,  $t(12) = -0.077$ ,  $p = .94$ .

Given the lack of VoE in the verbal condition by the participants who did not know either agent word, we wondered if this was due to word-specific knowledge or difficulties with updating representations in general. To examine this, we tested whether or not this group of toddlers demonstrated

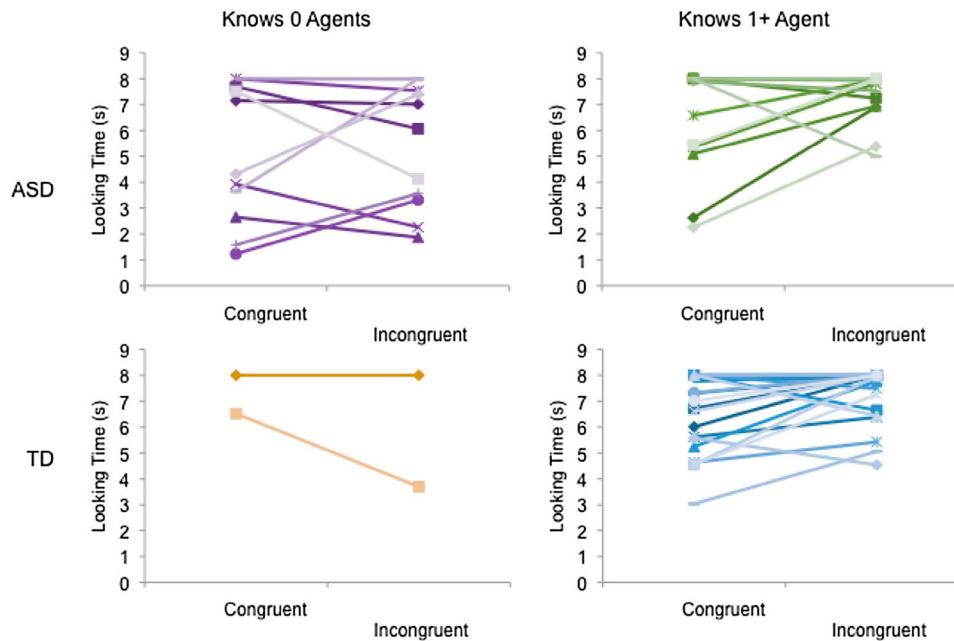
a VoE in the visual modality. Findings demonstrated that like the verbal condition, participants did not look significantly longer in the incongruent outcome than the congruent outcome  $t(12) = -0.921$ ,  $p = .38$ . Note that this group did not display poorer visual reception on the MSEL than the participants who knew an agent word,  $t(46) = 1.25$ ,  $p = .22$ .

Because the male to female ratio was much higher in the ASD group than the TD group, we conducted exploratory analyses on the role of gender on VoE magnitude in the TD group alone. A 2 (gender)  $\times$  2 (condition) ANOVA yielded no main effects of gender or condition, and no interaction between them ( $p > 0.45$ ).

## Discussion

In our dynamic world mental representations must be continually generated and updated. The current study shows that on average, toddlers with ASD can demonstrate this skill in a low-demand task using both visual and verbal information. When faced with a scene that was incongruent with the information they had gained during a brief occlusion, as a group, toddlers with ASD (just like TD controls) looked

**Fig. 3** Individual patterns of looking time to congruent and incongruent outcomes in the Verbal condition by group and word knowledge. Individual lines represent a single participant



longer and searched the scene more than when they saw a congruent outcome, regardless of whether that information was gained through verbal or visual means. To the best of our knowledge, this is the first demonstration that verbal updating skills are present in any young toddlers with ASD.

Relative to the TD group, our measures of language ability in the ASD group demonstrated clear delays, which makes the similarity in their pattern of task performance rather striking. The average MSEL receptive language score in the ASD group was classified as "Very Low," and all participants with ASD fell more than one SD below their age-appropriate standard score. We hypothesized that these difficulties with linguistic processing might make it difficult for children with ASD to use language to update mental representations. On the contrary, these delays were unrelated to their success in our task at the group level. Only toddlers who reportedly did not comprehend the word for at least one agent in the testimony (*cat* or *dog*), performed poorer in the verbal condition than those who did comprehend at least one agent word. Although this was true for toddlers in both diagnostic groups, it is worth noting that when queried on the specific items *cat*, *dog*, and *bed*, several caregivers of toddlers with ASD indicated that their child did not understand any of those words, and that toddlers with ASD comprised the majority of the group who did not comprehend the words referring to the agents (11 ASD vs. two TD). This underscores the heterogeneity associated with ASD: all children in our ASD group demonstrated language delays on the MSEL, but the specifics of those delays (e.g. vocabulary size and composition) likely varied. We suspect that toddlers with ASD (or any language-delayed toddlers) who have extremely low language skills may not be able to

perform verbal updating (because verbal updating depends on knowing at least some of the relevant words). Our data suggest that further research on the role of these individual differences on representational updating deserves further exploration. The group of participants who did not know at least one agent word also did not demonstrate a VoE in the visual updating condition, suggesting that they may have broader difficulties with updating, regardless of modality. Second, the analysis on the verbal condition did not meet criterion for statistical significance, despite the trend. This may be the result of low statistical power (only 13 participants did not know an agent word) and/or imprecise parental reports of vocabulary knowledge.

Although not statistically significant, the trend that the group of participants who comprehended at least one agent showed a VoE in the verbal condition while the other group did not sheds light on the strategies participants may be using to perform verbal updating in this paradigm. On the one hand, the task may be solved by processing the entire sentence (*Now the cat is on the bed!*) and detecting a violation when the dog is on the bed, or by solely processing the agent of the movement (e.g. the cat). Because the incongruity is created by a difference in this agent, the toddlers may have generated a less specific expectation ("something will happen to the cat"), which was then violated when the only change that happened involved the other agent. If this is the case, then detecting the violation may be an all-or-nothing phenomenon that hinges on the knowledge of either *cat* or *dog*. The words *cat*, and especially *dog* are some of the earliest words in TD children's vocabularies (Frank et al. 2017). Once those words enter a child's lexicon, the remaining vocabulary size will still vary. Thus, we may not

expect a relationship between VoE and the absolute size of the vocabulary, but instead a relationship between VoE and whether or not that vocabulary includes *cat* or *dog*. Findings from Ganea et al. (2016) are in line with this interpretation: there was also no relationship between receptive vocabulary size and VoE in that study. A more explicit test of the role of vocabulary knowledge on verbal updating could involve testing even younger children who understand *cat* and *dog* but fewer other words to comprise their total vocabulary (e.g. 12-month-old TD children).

It is important to note that some toddlers with ASD reportedly did not comprehend either of the agent words and still showed a VoE. This raises the possibility that the ASD group's language level was underestimated, at least in part. Research on tools used for language assessment in toddlers with ASD (parent report and MSEL) reveals that on the whole, these assessments are well-correlated (Luyster et al. 2008). However, the degree of agreement between expressive and receptive language scales varies, and production studies have identified that specific items can be underestimated by parent report (Jyotishi et al. 2016; Luyster et al. 2008). In response, some language researchers have suggested eyetracking as a potentially better measure to capture vocabulary knowledge in children with ASD (Brady et al. 2014). At least one other study has demonstrated relatively good performance on eyetracking measures of language comprehension among toddlers at risk for ASD (Chita-Tegmark et al. 2015).

Although we did not see differences in VoE between the ASD and TD groups, one difference we did find between the groups was that the ASD group did not look as long at the outcome scenes overall, nor did they search the scene as much as the TD group. This may be due to the fact that the children in the ASD group were significantly older than the TD group (mean: 28.6 vs. 18.5 months), and novelty-based looking time tends to decrease with age (Courage et al. 2006). It is also possible that the differences exhibited in looking time and fixation count are related to diagnostic status per se. Children with ASD are better and faster in visual search tasks than TD peers (Kaldy et al. 2011; Plaisted et al. 1998), and thus may be able to process the scene with fewer looks. Finally, children with ASD in our sample may have been more distractible than TD children. Regardless, this between-groups difference in looking time was present in both conditions, and did not influence their pattern of VoE.

Here, we did not see improved performance in the visual modality over the verbal modality in the ASD group. Both groups were able to learn about an occluded scene from verbal input, suggesting that children with ASD may have no problem learning from simple verbal testimony. The only difference between visual and verbal conditions emerged in how the toddlers allocated their attention to the screen: as a proportion of their total looking time, participants spent

slightly more time on the agent in the new location in the verbal, relative to the visual condition. Because similar differences between conditions were not present in any of our other measures and the effect was equally present in both groups, it would be premature to speculate about the interpretation of this finding.

Verbal information is typically transmitted in a social context. The task presented here has much of that social context stripped away: information is given unidirectionally, from a disembodied voice recording. While this is similar to some ways in which information is transmitted through media (e.g. television), it does not necessarily capture the rich context in which an informant may communicate new information face-to-face. Indeed, studies examining TD children's evaluation of testimony as reliable demonstrate differential behavior when the speaker is present versus absent (Jaswal et al. 2010; Koenig and Echols 2003). It is therefore possible that a disparity between individuals with TD and ASD would arise in a task with higher social demands.

Finally, it is important to note that this study is not a direct replication of Ganea et al. (2016), and there are some differences between that study's findings and our own. Namely, here we found VoE effects in the overall looking time measure, but not in the proportion of looking time measure. Ganea and colleagues found the opposite in their sample of TD toddlers. Because there are several methodological differences between that study and the current one, there are many plausible explanations for the asymmetry in findings. For example, when adapting the eyetracking paradigm, we set the encoding and outcome scenes in a more naturalistic room background, while the background of the agents in Ganea et al. (2016) was solid white. Given that the image background is somewhat more interesting to look at, it is not surprising that a larger proportion of the total looking time was used to examine the background in the current study. Additionally, in the current study only four test trials were presented compared to the prior study's 12 trials. Overall looking time in that study may have been subject to greater fatigue effects than seen here. Indeed, on average, looking time was much longer among the TD toddlers in this study ( $M=7.08$  s) than in Ganea et al. (2016) ( $M=5.25$  s). This overall greater looking time may also explain the difference in effect sizes between this study ( $\eta_p^2=0.09$  in the congruency comparison) and Ganea and colleagues' study ( $\eta_p^2=0.65$ ). Because the outcome scene was only presented for 8 s, looking time in the current study may have been subjected to ceiling effects. Indeed, participants in our study looked for the full 8 s outcome period in 31% of the 192 test trials.

This study is the first to examine visual or verbal updating abilities in toddlers with ASD. We have demonstrated that even with significant language delays, many of the toddlers with ASD in our sample were able to make inferences

based on verbal input and update their expectations about the world, as long as the task demands are low. Our study adds to the existing literature that suggests that while the language comprehension abilities of young children with ASD are heterogeneous, they may be intact on a variety of measures, and potentially be underestimated through traditional assessments. This work also raises several interesting questions for future studies about the linguistic and social demands associated with ASD toddlers' ability to use verbal input to update mental representations of the visual world.

**Acknowledgments** We thank the Autism, Behavior, and Child Development Lab Assessment Team with their help with clinical assessments as well as the families that participated in this study. Portions of this manuscript were presented at the International Meeting for Autism Research 2017.

**Author Contributions** AF conceived of the study, participated in its design, conducted the statistical analysis, and drafted the manuscript; AV participated in the design and coordination of the study and conducted assessment measures; PAG conceived of the study and participated in the design; ASC participated in the coordination of the study and helped conduct statistical analysis; ZK conceived of the study, participated in its design and helped to draft the manuscript. All authors read and approved the final manuscript.

**Funding** This project was supported by a Seed Grant from the Simons Foundation under the auspices of the Simons Center for the Social Brain at MIT (#319294) to Z. Kaldy, and US Department of Health and Human Services, Health Resources and Services Administration (HRSA) Grant #R40MC26195 to A. S. Carter.

## Compliance with Ethical Standards

**Conflict of interest** All authors declare that they have no conflict of interest.

**Ethical Approval** All procedures performed this study were in accordance with the ethical standards of the Institutional Review Board of the University of Massachusetts Boston and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

## References

- Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research*, 9(8), 810–828. <https://doi.org/10.1002/aur.1590>.
- Brady, N. C., Anderson, C. J., Hahn, L. J., Obermeier, S. M., & Kapa, L. L. (2014). Eye tracking as a measure of receptive vocabulary in children with autism spectrum disorders. *Augmentative and Alternative Communication*, 30(2), 147–159. <https://doi.org/10.3109/07434618.2014.904923>.
- Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The brief infant-toddler social and emotional assessment: Screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology*, 29(2), 143–155. <https://doi.org/10.1093/jpepsy/jsh017>.
- Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, 30(1), 213–236.
- Chita-Tegmark, M., Arunachalam, S., Nelson, C. A., & Tager-Flusberg, H. (2015). Eye-tracking measurements of language processing: Developmental differences in children at high risk for ASD. *Journal of Autism and Developmental Disorders*, 45(10), 3327–3338. <https://doi.org/10.1007/s10803-015-2495-5>.
- Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development*, 77(3), 680–695. <https://doi.org/10.1111/j.1467-8624.2006.00897.X>.
- Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, 52(4), 521–536. <https://doi.org/10.1037/dev000083>.
- Curtin, S., & Vouloumanos, A. (2013). Speech preference is associated with autistic-like behavior in 18-month-olds at risk for Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 43(9), 2114–2120. <https://doi.org/10.1007/s10803-013-1759-1>.
- de Marchena, A., Eigsti, I. M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, 119(1), 96–113. <https://doi.org/10.1016/j.cognition.2010.12.011>.
- Diamond, A. (1990). The development and neural bases of memory functions as indexed by the AB and delayed response tasks in human infants and infant monkeys. *Annals of the New York Academy of Sciences*, 608, 267–317.
- Ferguson, B., Graf, E., & Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: Evidence from 15-and 19-month-olds. *Cognition*, 131(1), 139–146. <https://doi.org/10.1016/j.cognition.2013.12.014>.
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 44(3), 677–694. <https://doi.org/10.1017/S0305000916000209>.
- 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. <https://doi.org/10.1016/j.jecp.2013.12.002>.
- Ganea, P. A. (2005). Contextual factors affect absent reference comprehension in 14-month-olds. *Child Development*, 76(5), 989–998. <https://doi.org/10.1111/j.1467-8624.2005.00892.X>.
- Ganea, P. A., Fitch, A., Harris, P. L., & Kaldy, Z. (2016). Sixteen-month-olds can use language to update their expectations about the visual world. *Journal of Experimental Child Psychology*, 151, 65–76. <https://doi.org/10.1016/j.jecp.2015.12.005>.
- 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(2), 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(1), 89–101.
- Ganea, P. A., & Saylor, M. M. (2013). Talking about the near and dear: Infants' comprehension of displaced speech. *Developmental Psychology*, 49(7), 1299–1307. <https://doi.org/10.1037/a0030086>.
- 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(8), 734–739. <https://doi.org/10.1111/j.1467-9280.2007.01968.X>.
- Giserman Kiss, I., Feldman, M. S., Sheldrick, R. C., & Carter, A. S. (2017). Developing autism screening criteria for the Brief Infant Toddler Social Emotional Assessment (BITSEA). *Journal of*

- Autism and Developmental Disorders.* <https://doi.org/10.1007/s10803-017-3044-1>.
- Grandin, T. (1995). How people with autism think. In E. Schopler & G. B. Mesibov (Eds.), *Learning and Cognition in Autism* (pp. 137–156). New York: Plenum Press.
- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5–25.
- Hartley, C., & Allen, M. L. (2014). Brief report: Generalisation of word-picture relations in children with autism and typically developing children. *J Autism Dev Disord*, 44(8), 2064–2071. <https://doi.org/10.1007/s10803-014-2074-1>.
- Hartley, C., & Allen, M. L. (2015a). Iconicity influences how effectively minimally verbal children with autism and ability-matched typically developing children use pictures as symbols in a search task. *Autism*, 19(5), 570–579. <https://doi.org/10.1177/1362361314536634>.
- Hartley, C., & Allen, M. L. (2015b). Symbolic understanding of pictures in low-functioning children with autism: The effects of iconicity and naming. *Journal of Autism and Developmental Disorders*, 45(1), 15–30. <https://doi.org/10.1007/s10803-013-2007-4>.
- Hartley, C., Trainer, A., & Allen, M. L. (2017). Investigating the relationship between language and picture understanding in children with autism spectrum disorder. *Autism*. <https://doi.org/10.1177/1362361317729613>.
- Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, 21(10), 1541–1547. <https://doi.org/10.1177/0956797610383438>.
- Jyotishi, M., Fein, D. A., & Naigles, L. R. (2016). “Didn’t I just say that?” Comparing parent report and spontaneous speech as indicators of grammatical development. *Research On Developmental Disabilities*, 61, 32–43. <https://doi.org/10.1016/j.ridd.2016.12.013>.
- Kaldy, Z., Kraper, C., Carter, A. S., & Blaser, E. (2011). Toddlers with autism spectrum disorder are more successful at visual search than typically developing toddlers. *Developmental Science*, 14(5), 980–988. <https://doi.org/10.1111/j.1467-7687.2011.01053.x>.
- Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders*, 36(6), 807–828.
- Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16(2–3), 287–308. <https://doi.org/10.1080/01690960042000058>.
- Klin, A., Danovitch, J. H., Merz, A. B., & Volkmar, F. R. (2007). Circumscribed interests in higher functioning individuals with autism spectrum disorders: An exploratory study. *Research and Practice for Persons with Severe Disabilities*, 32(2), 89–100.
- Koenig, M. A., & Echols, C. H. (2003). Infants’ understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, 87(3), 179–208.
- Lazenby, D. C., Sideridis, G. D., Huntington, N., Prante, M., Dale, P. S., Curtin, S., ... Tager-Flusberg, H. (2016). Language differences at 12 months in infants who develop autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 46(3), 899–909. <https://doi.org/10.1007/s10803-015-2632-1>.
- Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule* (2nd ed.). Los Angeles: Western Psychological Corporation.
- Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38(8), 1426–1438. <https://doi.org/10.1007/s10803-007-0510-1>.
- Luyster, R. J., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language*, 34(3), 623–654.
- Luyster, R. J., & Lord, C. (2009). Word learning in children with autism spectrum disorders. *Developmental Psychology*, 45(6), 1774–1786. <https://doi.org/10.1037/a0016223>.
- Mullen, E. (1995). *Mullen scales of early learning*. Circle Pines: American Guidance Service, Inc.
- Naigles, L. R., Kelty, E., Jaffery, R., & Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. *Autism Research*, 4(6), 422–437. <https://doi.org/10.1002/aur.223>.
- Naigles, L. R., & Tek, S. (2017). ‘Form is easy, meaning is hard’ revisited: (re) Characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *Wiley Interdisciplinary Reviews: Cognitive Science*. <https://doi.org/10.1002/wcs.1438>.
- Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: A two year follow-up. *Autism Research*, 1(2), 97–107. <https://doi.org/10.1002/aur.12>.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
- Plaisted, K., O’Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 39(5), 777–783. <https://doi.org/10.1017/S002196309800213>.
- Preissler, M. A. (2008). Associative learning of pictures and words by low-functioning children with autism. *Autism* 12(3), 231–248. <https://doi.org/10.1177/1362361307088753>.
- Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, 97(1), B13–B23. <https://doi.org/10.1016/j.cognition.2005.01.008>.
- Saylor, M. M., & Baldwin, D. A. (2004). Discussing those not present: Comprehension of references to absent caregivers. *Journal of Child Language*, 31(3), 537–560. <https://doi.org/10.1017/S030500904006282>.
- Smith, N. J., Sheldrick, R. C., & Perrin, E. C. (2013). An abbreviated screening instrument for autism spectrum disorders. *Infant Mental Health Journal*, 34(2), 149–155. <https://doi.org/10.1002/imhj.21356>.
- South, M., Ozonoff, S., & McMahon, W. M. (2005). Repetitive behavior profiles in Asperger syndrome and high-functioning autism. *Journal of Autism and Developmental Disorders*, 35(2), 145–158. <https://doi.org/10.1007/s10803-005-1992-3>.
- Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child Development*, 78(2), 542–557. <https://doi.org/10.1111/j.1467-8624.2007.01022.x>.
- Tager-Flusberg, H. (1985). Basic level and superordinate level categorization by autistic, mentally retarded, and normal children. *Journal of Experimental Child Psychology*, 40(3), 450–469.
- Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, 1(4), 208–222. <https://doi.org/10.1002/Aur.38>.
- Ungerer, J. A., & Sigman, M. (1987). Categorization skills and receptive language development in autistic children. *Journal of Autism and Developmental Disorders*, 17(1), 3–16.
- Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, 17(6), 872–879. <https://doi.org/10.1111/desc.12170>.
- 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(32), 12933–12937. <https://doi.org/10.1073/pnas.1121057109>.
- Weismer, S.E., Gernsbacher, M. A., Stronach, S., Karasinski, C., Eernisse, E. R., Venker, C. E., & Sindberg, H. (2011). Lexical and grammatical skills in toddlers on the autism spectrum compared to late talking toddlers. *Journal of Autism and Developmental Disorders*, 41(8), 1065–1075. <https://doi.org/10.1007/s10803-010-1134-4>.
- Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, 16(5), 372–377. <https://doi.org/10.1111/J.0956-7976.2005.01543.X>.