Received: 22 June 2022 | Revised: 2 September 2022 | Accepted: 6 September 2022

DOI: 10.1002/icd.2372

THEORETICAL ARTICLE

WILEY

The dynamic functions of social cues during children's word learning

Crystal Lee | Casey Lew-Williams

Department of Psychology, Princeton University, Princeton, New Jersey, USA

Correspondence
Crystal Lee, Princeton University, Department of Psychology, Princeton, NJ 08540, USA.
Email: cl40@princeton.edu

Abstract

Children learn words in a social environment, facilitated in part by social cues from caregivers, such as eye-gaze and gesture. A common assumption is that social cues convey either perceptual or social information, depending on the age of the child. In this review of research on word learning and social cues during early childhood, we propose that (1) the functions of social cues are not categorically perceptual or social, and (2) social cues support word learning in four interdependent ways: by helping children to orient attention, extract relevant information, disambiguate referents, and understand others' referential intent. We conclude with specific recommendations for theory-building and suggest that the dynamic and complex functions of social cues need to be accounted for in any complete theory of word learning.

KEYWORDS
language development, social cues, word learning

1 | INTRODUCTION

As children learn language, they interact with other people. Their social environment contains a suite of social cues—provided by caregivers, siblings, and other children and adults—that support their emerging language abilities. For example, to learn new words, children are known to follow another's eye-gaze (Bannard & Tomasello, 2012; Booth et al., 2008; Gliga et al., 2012; Houston-Price et al., 2006; Paulus & Fikkert, 2014; Tenenbaum et al., 2014) or body movements and gestures (Booth et al., 2008; Rader & Zukow-Goldring, 2012) toward an object of joint

attention. In addition, a child's willingness to learn new words from others is influenced by whether they consider their interlocutor to be reliable or unreliable (Krogh-Jespersen & Echols, 2012; Scofield & Behrend, 2008; Sobel et al., 2011), and knowledgeable or ignorant (Brooker & Poulin-Dubois, 2013; Mangardich & Sabbagh, 2018; Sabbagh & Baldwin, 2001; Sabbagh & Shafman, 2009). In doing so, young children use cues from their interlocutors' behaviours to facilitate language learning.

Prior studies have exposed children to a range of social cues—namely pointing and eye-gaze—to understand how they support the learning of different types of words, including nouns, verbs, and adjectives. Experimental studies (Akhtar et al., 1996; Baldwin et al., 1996; Houston-Price et al., 2006; Striano et al., 2006; but see Scofield & Behrend, 2011) and longitudinal studies (Brooks & Meltzoff, 2008; Morales, Mundy, Delgado, Yale, Messinger, et al., 2000; Morales, Mundy, Delgado, Yale, Neal, & Schwartz, 2000; Tenenbaum et al., 2014) have determined that the presence of social cues promotes children's accuracy in identifying words during test phases that, for example, involve selecting or looking at a target referent. As an example, Morales, Mundy, Delgado, Yale, Messinger, et al. (2000) found that individual differences in infants' responses to initiations of joint attention by their caregivers between 6 and 18 months were significantly related to language outcomes at 24 months and 30 months of age. They also found that an aggregate measure of joint attention skills (between 6 and 18 months) predicted language outcomes at 30 months, even after accounting for language skills at 24 months. This suggested that children who possess better vocabulary skills at 24 months did not simply maintain their level of vocabulary in comparison to their peers at 30 months of age. Instead, the gap between higher- and lower-vocabulary individuals widened over time based on variation in social aspects of communication. Thus, individual differences in children's ability to engage with social cues seem to be predictive of their later language outcomes.

This leads us to ask: *why* do social cues facilitate word learning? What functions do social cues serve during the observation, processing, and learning of new words? There are three broad proposals that have aimed to answer these questions. First, in the socio-pragmatic theory of word learning, word learning requires children's recognition of language as a means to communicate, where social cues would support word learning by representing an interlocutor's communicative intent (Akhtar et al., 1996; Akhtar & Tomasello, 1998; Akhtar & Tomasello, 2000; Baldwin, 1993; Bruner, 1983; Nelson, 1988; Tomasello, 2000, 2001). In a seminal study, Baldwin et al. (1996) found that 18- to 20-month-old children were more likely to establish a word-object mapping when learning occurred when a speaker paid attention to the same object as the infant compared to when the speaker was not physically present. They concluded that the speaker's referential intent was an important component of children's learning. Tomasello et al. (1996) came to a similar conclusion using affect-related social cues. In their study, 18-month-old children played a finding game with an experimenter. The experimenter announced her intention to find a novel object using a nonce word (e.g., "Let's find the *gazzer*"). She picked up one nonce object, rejected it with disappointment, and gleefully picked up the target object. Children immediately mapped the nonce word to the latter. These studies argue that children are sensitive to the communicative nature of word learning and perceived these cues to be indicative of the referential intent of an interlocutor. Thus, the primary function of social cues is to provide *social information* or communicative information about a speaker's goal or intention (such as which object is being referred to in space).

A second proposal is based on associative mechanisms and has argued that social cues drive word learning via general cognitive processes, such as attention and memory (Colunga & Smith, 2005; Samuelson & Smith, 1998; Smith, 2000). This proposal prioritizes associative mechanisms over socio-pragmatic information, on the assumption that very young children would be unlikely to know how to interpret social information in a skillful way. What is likely happening, they argue, is that social cues provide information related to timing, spatial information, or simple associations rather than intent (Deák et al., 2014; Yu & Smith, 2012). For example, Samuelson and Smith (1998) argue against the socio-pragmatic conclusions drawn by Akhtar et al. (1996), suggesting that if learners' behaviours seem to indicate the recognition of referential intent, they may instead be behaviours that were "guided by the general effects of a contextual shift on memory and attention" (Samuelson & Smith, 1998, p. 94). This work emphasizes that word learning, particularly at younger ages, may be based mostly on general cognitive mechanisms rather than a recognition of the communicative intent of others.

A third approach has, in some sense, united these two proposals, suggesting that perceptual and social information work in tandem to support word learning. The Emergentist Coalition Model (ECM; Golinkoff & Hirsh-Pasek, 2006; Hollich et al., 2000) combines social-pragmatic and association-based theories and consists of three central tenets: (1) children are sensitive to perceptual, social, and linguistic cues, (2) the weights of these cues change over time, and (3) principles of word learning also change over time. This model suggests that children *first* consider perceptual information when learning new words, such as object salience. For example, an infant may map a novel word to an object that grabs their attention, suggesting that perceptual information provides the initial momentum to help the word learning process. This hypothesis is corroborated by the results of Pruden et al. (2006), where 10-month-olds were found to rely on perceptual cues over social cues to learn words. However, the weights of these cues change over time. As children develop, they deemphasize the weight of perceptual cues in favour of social cues such as eye-gaze or pointing. ECM suggests that these cues transform in a similar way: social cues initially function to provide low-level perceptual information, but ultimately change to provide abstract social information such as a speaker's goals or intentions during word learning.

What remains underspecified in current frameworks are the diverse and complex functions that social cues play during word learning. Socio-pragmatic and association-based theories, as well as ECM, have generally assumed that social cues provide *either* social *or* perceptual information at different points in developmental time. However, these theories do not fully consider the dynamics of interactions between caregivers and children. It may not be realistic to neatly categorize the functions of social cues as perceptual or social. Doing so hides the richness of social cues for elucidating word learning in the moment and for explaining children's learning outcomes. Instead, their functions may fall along a gradient between perceptual and social information.

In what follows, we examine existing literature on young monolingual children's exposure to (mostly concrete) words and social cues, and we find that the supportive role of social cues cannot be neatly discretized into a perceptual versus social division. Our review points to four distinct but interrelated functions of social cues in children's word learning, which vary based on the type of social cue and the age of the child. How do different functions support early word learning, separately and in combination? We present this review as a way of informing frameworks of word learning to accommodate the richness and diversity of natural social cues.

We first provide an overview of a range of social cues and their suggested functions during word learning. We include articles that primarily focus on noun learning by monolingual children, although verb- and adjective-learning studies are described at times. Then, we discuss the two social cues that have been studied thoroughly in previous research on word learning—eye-gaze and gesture—and we detail their possible functions. Following this overview, we discuss key directions for future research that are needed to create a complete theory of how children learn words within socially-rich contexts.

2 | WHAT ARE THE FUNCTIONS OF SOCIAL CUES?

The first finding from our literature review is that there are four distinct but interrelated functions that social cues may serve in word learning: orienting attention, extracting information, disambiguating referents, and indicating referential intent (Figure 1). These four functions were generated based on wording from many empirical articles described below, and/or from our interpretations of the functional roles played by social cues in the experiments. This way of categorizing different functions, although imperfect, builds upon the general assumption in prior literature that social cues provide only perceptual *or* social information. *Orienting attention* refers to a caregiver's cues that serve to orient a child's attention to an area in space. This is mostly a perceptual function that may operate separately from a caregiver's social desire to orient attention. *Extracting information* refers to the influence of a caregiver's cues on the type of information that a child detects and/or encodes from a visual scene, such as an object's features (e.g., Reid et al., 2004). *Disambiguating referents* refers to social cues from a caregiver that reduce ambiguity between a word and its possible referents or meanings. Lastly, *indicating referential intent* refers to a caregiver displaying social cues to convey their intent to communicate with the child.

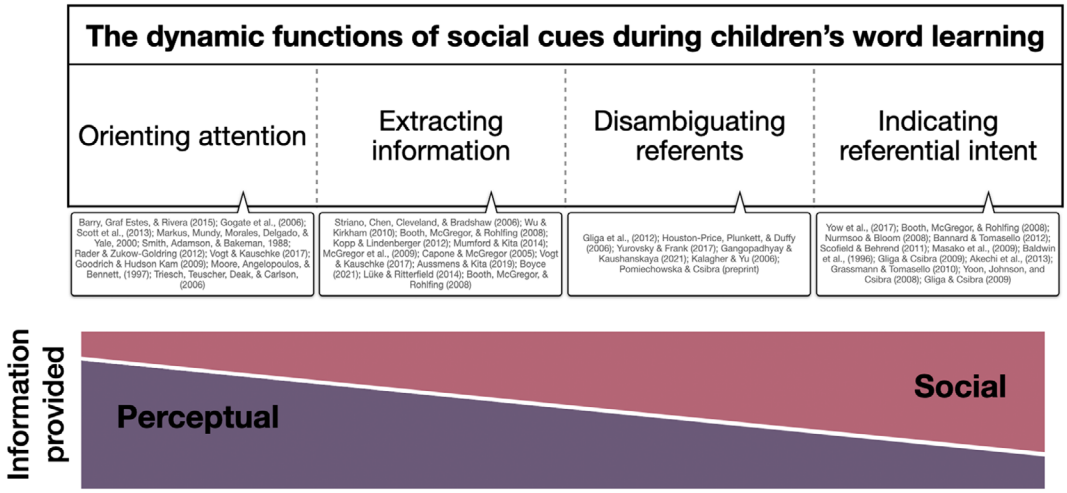

FIGURE 1 Key functions of eye-gaze and gestures reported in previous studies. The upper portion of the figure depicts the functions of social cues described in prior literature, with relevant studies cited underneath. The dashed lines between functions represent the fuzzy boundaries and interactiveness between functions. The lower portion of the figure depicts the type of information that social cues may provide across the four functions, demonstrating that each may provide ranging 'amounts' of perceptual and social information, but with notable overlap between the two.

The second idea emerging from our review is that these four functions can provide a range of perceptual and social information, and to differing extents depending on the function. One could argue that orienting attention relies predominantly on perceptual information, at least initially, but with increasing influence from social context. Similarly, one could argue that inferring referential intent requires an aptitude for inferring abstract social information, but importantly, perceptual information cannot be divorced from this process. In addition, disambiguation could be categorized as perceptually- and/or socially-based, depending on context and interpretation. Finding a target referent among distractor referents could be the result of a perceptual mechanism, *or* from an implicit or explicit understanding that the referent was highlighted based on a caregiver's desire to communicate. Finally, a child's ability to extract relevant information from a scene may be due to the perceptual nature of a social cue, or by inferring a speaker's communicative intent in using it.

These functions are also likely to be at least somewhat interdependent on one another. Categorizing the functions of social cues in a binary way as perceptual *or* social is probably not an accurate way to think about their influence on learning. Instead, functions are likely to depend on each other, or the weights of perceptual and social information are likely to vary for different social cues. For example, extracting relevant information about objects, events, and/or people requires the ability to follow cues from others in a way that shifts attention, but also the ability to use a caregiver's attention to infer referential intent. It is also likely that how children extract information can be influenced by whether the referential intent of the caregiver is detected, such that information may be extracted better/faster when a child recognizes communicative intent. Similarly, a child's visual attention may be oriented faster when referential intent is detected. These examples suggest that the functions of social cues are inherently interactive, that the boundaries between each function are somewhat fuzzy, and that the functions may build upon each other to facilitate learning. Considering social cues as providing only perceptual *or* social information may mask the full richness of information offered by eye-gaze, gesture, and other social cues.

We also do not find a specific point in time where children transition from using social cues as a source of perceptual information to the same cues being a source of social information. While many studies do show that children follow the developmental trajectory hypothesized by ECM, there is sizable variability *within age points*. Across studies within the same developmental time period, social cues seem to play different functions in different experiments.

Our review indicates that there is more to be explained about the functions of social cues across children and across time. Therefore, in the sections that follow, we focus on eye-gaze (Section 3) and gestures (Section 4) and examine the four main ways in which they support children's word learning. Within the descriptions of each function, we provide a brief overview of relevant studies. Where pertinent, we include the ages of participants, which range from 6 months to 6 years across different studies, to provide a window into developmental change.

3 | THE FUNCTIONS OF EYE-GAZE AS A SOCIAL CUE

Gaze following refers to the ability to follow the eye-gaze of another during an interaction (Brooks & Meltzoff, 2005; Moore, 2008), and infants demonstrate this ability from early on in life. They begin paying attention to the eyes of others as early as 2–3 months of age (Caron et al., 1973; Maurer & Salapatek, 1976). This ability continues to develop in subsequent months (Vecera & Johnson, 1995) and stabilizes around 6–8 months of age (D'Entremont et al., 1997; Gredebäck et al., 2010). This early attentiveness to eyes is important as it is required for two important social skills: gaze following (in sighted children) and joint attention. Children use this skill to develop the ability to determine and share common reference points between themselves and a caregiver, which is also known as *joint attention* (Çetinçelik et al., 2021; Mundy & Newell, 2007; Tomasello, 1995). For the purposes of this review, we collapse 'gaze-following' and 'joint attention' because both involve similar behaviours (i.e., following the eye-movements of another) and are deeply intertwined as following eye-gaze is a crucial aspect of engaging in joint attention in sighted children (Çetinçelik et al., 2021).

The ability to follow eye-gaze has been linked with vocabulary development (Akhtar & Gernsbacher, 2007; Brooks & Meltzoff, 2005, 2008; Carpenter et al., 1998; Morales et al., 1998; Morales, Mundy, Delgado, Yale, Messinger, et al., 2000; Morales, Mundy, Delgado, Yale, Neal, & Schwartz, 2000; Tomasello & Farrar, 1986). Brooks and Meltzoff (2005) found in a longitudinal study that gaze-following abilities at 10–11 months were positively related to word and phrase comprehension, as well as to total gestures when measured again at 18 months, as measured by the MacArthur Child Development Inventory (CDI; Fenson et al., 1994). In a different study, positive relationships between gaze-following abilities and later productive vocabulary were also found (Brooks & Meltzoff, 2008). Given the link between gaze-following and language learning, what is the function of eye-gaze cues during moments when children have the potential to learn a new word? We turn to each of the four purported functions of eye-gaze cues in turn.

3.1 | Orienting attention

One way in which word learning outcomes can be influenced by social cues is through orienting attention, which at first pass seems categorizable as a perceptual function. Infants can follow the movements of the eyes or pupils (Farroni et al., 2004) and the direction of head motion (Corkum & Moore, 1998; D'Entremont, 2000; Moore & Corkum, 1998) to shift their attention to a specific object in a visual scene, and this variation in attention predicts subsequent word learning. Gogate et al. (2006) found that the extent to which 6- and 8-month-old infants gazed between their caregiver and the gazed-at-object positively predicted subsequent novel word learning. Moreover, individual differences in the number of parent–infant joint attention events in a fixed time window has been found to predict variation in early vocabulary size (Carpenter et al., 1998; Farrant & Zubrick, 2012; Markus, Mundy, Morales, Delgado, & Yale, 2000; Scott et al., 2013; Saxon, 1997; Smith et al., 1988; Tomasello & Farrar, 1986). Joint attention can be further analysed as cases that are caregiver-initiated or child-initiated. For instance, Scott et al. (2013) found that, compared to when mothers followed infants' attention to engage in joint attention, how often infants followed *mothers*' attention related to later vocabulary growth, suggesting that parent–infant joint attention events are indeed important in lexical development.

3.2 | Extracting information

Beyond directing and orienting attention to a stimulus in a visual scene, taking advantage of eye-gaze may enhance the information extracted from visual scenes. This could also be categorized as providing perceptual information, perhaps in a way that follows naturally from a redirection of attention. In a typical joint attention scenario, the eye-gaze cues provided by another speaker could enhance the processing of a gazed-at object by improving encoding or by highlighting relevant information about the object. This indirectly facilitates the mapping between a given label and object in the world (Çetinçelik et al., 2021).

Several studies have found that the presence of eye-gaze cues does facilitate extracting information from a visual scene (Hirotani et al., 2009; Hutman et al., 2016; Kopp & Lindenberger, 2011, 2012; Reid et al., 2004). For example, Reid et al. (2004) first established different ERP signals when an object was cued by another's eye-gaze or not; during a second viewing, 4-month-old infants showed an increased positive slow wave (PSW) for objects that were *not* cued by another's gaze, an indication for a memory representation of a partially encoded stimulus (Snyder, 2010). This result suggests that objects that were cued through eye-movements were more familiar than objects that were uncued. The subsequent viewing of the uncued object led to an updating of its mental representation. In comparison, objects that had been accompanied by eye-gaze toward them initially required less of an update, suggesting that information about the object had already been encoded by virtue of its initial presentation (i.e., gazing at the same object another has looked toward).

Eye-gaze cues can also facilitate information extraction even when situations are more cognitively challenging. For example, Striano et al. (2006) examined the effect of joint attention on object processing in 9- and 12-month-old infants. In a joint-attention condition, a speaker alternated gaze between a single toy and the infant. In an object-only condition, a speaker gazed at the toy and at the ceiling, but never the infant. At test, infants were presented with the toy they saw previously and a novel toy. The researchers found that only infants in the joint-attention condition preferred to look at the novel toy at 9 months. By 12 months, infants looked longer toward the novel toy in both conditions. This suggests that while 12-month-olds were able to process enough information about the object in either condition, encoding the object was significantly improved for the 9-month-olds by the presence of the speaker's eye-gaze. The authors suggest that infants' dependence on eye-gaze as an informative cue depends on the complexity of the learning task; in this experiment, encoding of the object may have been complex for 9-month-old infants but not for 12-month-old infants, such that they could rely on eye-gaze in the absence of bids for joint attention from the speaker.

Eye-gaze cues from another not only affects how much information is extracted from a visual scene, but also what *type* of information is extracted (Hoehl et al., 2012; Michel et al., 2019; Reid & Striano, 2005; Wu & Kirkham, 2010; Wu et al., 2011; Wahl et al., 2013, 2019;). Wu and Kirkham (2010) examined how infants' learning of multi-modal audio–visual events was affected when attention was cued to a portion of a screen from either eye-gaze cues or perceptual cues. They found that while social eye-gaze cues helped 4-month-olds learn where to look on a screen, those same social cues additionally helped 8-month-olds learn the specific multimodal events they were exposed to. Learning effects were not found with perceptual cues such as flashing central squares, despite reliably orienting attention to sides of the screen.

Some have speculated that the information abstracted from a situation can also be retained longer when eye-gaze from the speaker is present. Kopp and Lindenberger (2011, 2012) examined 4- and 9-month-olds' immediate recognition and long-term retention of novel objects depending on whether these objects were presented during high or low moments of joint attention (JA). Infants were exposed to a series of novel objects, and an experimenter either alternated her gaze between the object and speaker (high-JA) or only looked at the object (low-JA). The researchers found that presence of joint attention modulated 9-month-olds' recognition both immediately after exposure and after a one-week delay. In contrast, they found that 4-month-olds' recognition was evident *only* after a one-week delay, although they attribute this difference to experimental issues of ERP detection in young infants. Nevertheless, they argue that the presence of joint attention during learning scenarios may have facilitated elaborated encoding and storage, improving long-term retention.

3.3 | Disambiguating referents

Another proposed function of eye-gaze is to disambiguate referents in a visual scene. Children can readily use the eye-movements of another to disambiguate objects in a sensitive way: Hembacher et al. (2020) found that 3- and 4-year-old children used the gaze of another to identify a referent in different ways depending on the scenario. They looked to the experimenter to provide disambiguating information the most when two novel objects were presented, an intermediate amount when the object set included one novel and one familiar object, and the least when both objects were familiar. Likewise, in one of the experiments conducted by Graham et al. (2010), researchers sought to determine how 24-month-old children used eye-gaze to map novel labels to objects when presented with multiple sets of referents (both familiar and novel). They found that children were equally likely to map novel labels to the gazed-at object, regardless of whether the objects were familiar or novel. This suggests that young learners can indeed use eye-gaze to disambiguate a referent.

The ability to disambiguate visual referents through eye-gaze has also been shown to improve word learning (Houston-Price et al., 2006; Law et al., 2012). Law et al. (2012) tested 18-month-olds in a word-learning task that required them to follow a speaker's gaze to one of two novel objects and learn its novel label. Vocabulary measures were collected at 18, 24, and 30 months of age. They found that 18-month-old children were able to follow eye-gaze to disambiguate a visual referent and learn a novel label, and that performance on this disambiguation task was related to vocabulary both concurrently and longitudinally.

At the same time, using eye-gaze as a disambiguating cue is not always successful. Frank et al. (2013) examined the efficacy of using mothers' eye-gaze to disambiguate a visual referent in mother–child dyads. They found that mothers' eye-gaze was noisy and did not fully disambiguate the referred-to object, suggesting that, in naturalistic situations, relying on an eye-gaze cue may not always lead to successful disambiguation. Furthermore, children do not always rely on eye-gaze cues for disambiguation when other sources of information are available. Graham et al. (2010) found that children are not likely to use eye-gaze as a disambiguating cue when they can rely on the "mutual exclusivity" assumption as another source of information. The mutual exclusivity assumption is the assumption that one object can have only one label (Markman, 1992; Merriman & Schuster, 1991). The question in the experiment by Graham et al. was whether children would link a novel label to a familiar object (as directed by eye-gaze) or a novel object (as directed by a mutual exclusivity assumption). Children were presented with sets of one novel object and two familiar objects, and on half of the trials, the experimenter asked for the novel object using a novel label. Critically, the experimenter's eye-gaze provided an inconsistent clue as she looked at a familiar object instead. They found that children mapped the novel label onto the novel object, suggesting the mutual exclusivity assumption overrode information provided by eye-gaze. Thus, eye-gaze as a disambiguating cue is not necessarily used when it contrasts with prior knowledge.

3.4 | Indicating referential intent

As mentioned previously, some scientists propose that the ability to follow gaze is evidence for infants' inferring another's referential intent, indicating a direct uptake of social information (Deligianni et al., 2011; Senju & Csibra, 2008; Striano & Reid, 2006; Tomasello, 2000; Tomasello et al., 1996; Tomasello & Barton, 1994). Senju and Csibra (2008) investigated whether 6-month-old infants would follow the eye-gaze of a speaker more often if the speaker presented direct eye contact to the participant, compared to an interesting perceptual cue (i.e., a moving flower) prior to the initiation of any eye movements. The authors found that infants who were presented with direct eye contact made more first eye movements to the gazed-at object, as well as made more saccades between the face of the speaker and the gazed-at object. In contrast, the perceptual attention-getter did not facilitate as much gaze-following behaviour as the ostensive cue (direct eye-contact). They interpret these results to suggest infants can indeed follow the gaze and attention of others, but require strong social, ostensive cues to do so. Perceptual cues or attention-getters alone were not enough to initiate gaze-following.

Beier and Carey (2014) attempted to figure out whether 12-month-old infants inferred perceptual or social information from an interaction by presenting them with a novel entity (a coloured blob) and examined whether infants would follow its movement (i.e., attentional focus). They found that infants followed the attentional focus of the blob when it responded contingently (e.g., when an experimenter looked at it and it lit up), which could be construed as infants inferring either perceptual or social information. However, crucially, infants demonstrated differentiated behaviours when the context did vs. did not suggest a social interaction. When an experimenter clapped two sticks together with a neutral facial expression and the blob lit up, infants did not follow the entity's focus. However, when an experimenter clapped her hands together and smiled warmly at the blob, and when it then lit up, the infant shifted their gaze. This suggests that infants attributed agency to the non-human object and inferred referential intent when the context seemed to communicate a social interaction. This study provides further evidence for children's ability to detect social information.

Studies examining gaze-following in children older than 12 months routinely find that the function of following eye-movements is to infer the referential intent of a speaker (Baldwin, 1991, 1996; Butler et al., 2000). For example, Baldwin (1991) taught 16–19-month-old infants novel names of objects when a speaker was either: focused on the same object as the infant or looked at an object different from the infant. Their experiment was designed to determine whether infants use nonverbal, social cues from another to infer word-object mappings or if they indiscriminately learned mappings regardless of where a speaker is looking. They found that children both learned the novel label when the participants' and speakers' attention coincided, as well as avoided mapping errors when the object of focus did not coincide with the speaker's view. This work suggests that children can use eye-gaze as a non-verbal cue that is relevant to the formation of novel label inference.

3.5 | Understanding the many functions of eye-gaze for word learning

The functions of eye-gaze as a cue to language learning vary widely. From very early ages of development, the eye movements of others serve to orient infants' attention. Beyond attention, eye-gaze has been shown to influence the information extracted from a visual scene, such as enhancing object processing or retaining different types of information. Moreover, eye-gaze can be used to indicate referential intent. The proposed transition from perceptual to social effects, as in the ECM model by Hollich et al. (2000), may depend on the age of individual learners. For instance, as infants mature, the information abstracted from eye movements may change from perceptual to social (e.g., from relying on gaze to guide their own attention vs. later using it to understand a speaker's intention).

At the same time, it is reasonable to question if this proposed developmental transition is an accurate characterization of how children extract perceptual and social information. We argue that this is not the case: in our overview, we find that eye-gaze can act in perceptually- or socially-based ways at the same time point in development. For example, some studies conclude that, under the right conditions, the social intent of a speaker can be inferred by infants younger than 12 months of age (see Senju & Csibra, 2008; Senju et al., 2008, for a review see Striano & Reid, 2006). In opposition, other studies with the same age of participants suggest that only perceptual information is conveyed (Brooks & Meltzoff, 2005; Corkum & Moore, 1998; Moore & Corkum, 1998). However, eye-gaze is likely to take on a richer social meaning as children age, based on the increasing diversity of social and communicative experiences that children engage in across time as well as the maturation of underlying neural and cognitive mechanisms. Furthermore, the different functions of eye-gaze can occur in some combination or along a gradient—and it is again likely that a single eye-gaze cue can serve an increasing number of functions as children develop. It does not seem to be the case that there is a developmental time point where children qualitatively change their usage of a cue from perceptual to social in service of word learning. Rather, children likely take advantage of social cues to support word learning in whatever way makes sense in the moment. For example, eye-gaze can be used to guide attention throughout development, but even from early in life, gaze may indicate referential intent only in certain contexts. This variation in functional roles is likely to be explained by other factors (see Section 5).

Relatedly, some studies cannot definitively determine the function of an eye-gaze presented in an experiment. Often, studies do not find evidence for eye-gaze functioning as conveying perceptual information over social information. Consider Striano et al. (2006). In this study, 12-month-olds looked toward a novel object in both a joint-attention and object-only condition, while 9-month-olds looked toward the novel object reliably more than chance only in the joint-attention condition. What function did joint attention serve for 12-month-olds in this experiment? The authors conclude that while 9-month-olds benefitted from joint attention to enhance object processing in a cognitively challenging scenario, 12-month-olds used domain-general mechanisms to encode object characteristics regardless of experimental condition. However, it remains possible that the heightened performance of the 12-month-olds is the result of inferring referential intent from following the gaze of the speaker in both conditions, as opposed to using eye-movements for simply orienting attention to a spatial location. The experimental setup may have facilitated object processing regardless of whether it was establishing joint attention or simply following the eye-movements of another. This alternative explanation remains plausible, as other research has argued that 12-month-olds can determine the referential intent of another (Csibra & Volein, 2008; Okumura et al., 2013; Senju & Csibra, 2008; Striano & Reid, 2006). For example, Okumura et al. (2013) found that 12-month-olds anticipated the appearance of objects on a screen when the objects were previously referred to by a human agent, compared to a robot's gaze.

In the opposite way, consider Butler et al. (2000). Fourteen- and 18-month-old infants were engaged in joint-attention events with an experimenter who looked toward two visual events (one on each side) in varying scenarios. These scenarios included: when an experimenter's view was unobstructed, when their view was obstructed by opaque screens, and when their view was obstructed by opaque screens but included windows to the targets. Eighteen-month-olds turned in accordance with the experimenter's gaze much more when the experimenter's view was unobstructed, suggesting that they understood the referential nature of eye-gaze. Fourteen-month-olds, however, were much more variable—they turned less in accordance with the experimenter in comparison to the 18-month-olds, regardless of whether their view was obstructed. This variation would seem to contradict the inferences concluded by research such as Okumura et al. (2013) in that 14-month-old infants may have been able to follow the gaze of another, but they did not understand the referential intent behind this movement. It could be the case that the elaborate set-up of the experiment might have presented too difficult a situation for the 14-month-olds, but nevertheless, we cannot infer that the 14-month-olds are detecting referential intent.

In summary, the functions of eye-gaze cues during word learning present a much less defined picture compared to the black-and-white dichotomy between perceptual and social information that has sometimes been assumed. While it is fairly clear that the functions of eye-gaze cues vary by the age of the child, there is not a clear transition from inferring perceptual to social information, as all age points, including infancy, may infer both perceptual and social information. Furthermore, it is difficult to determine whether a social cue is providing only perceptual or only social information. A more conservative conclusion is that eye-gaze cues can convey a combination of perceptual and social information at the same time. Therefore, other factors may be at play that influence the function of social cues during word learning.

4 | THE FUNCTIONS OF GESTURE AS A SOCIAL CUE

Gestures—hand movements that co-occur with speech—are prevalent in a sighted child's visual input throughout development (Goldin-Meadow & Singer, 2003; Vigliocco et al., 2019). While deictic gestures (e.g., pointing) are the most prevalent in a child's environment, representational gestures provided by caregivers are also present (e.g., holding up two fingers in a V-shape to represent a rabbit's ears) (Iverson et al., 1999; Özçalışkan & Goldin-Meadow, 2005; Rowe et al., 2008). (Conventional gestures—that is, gestures that are used by a particular culture or community, such as waving a hand to say 'hi'—are certainly present in children's communicative environments, but we are not aware of research that suggests they facilitate word learning). Children exploit information conveyed by

gestures beginning at an early age. For example, between 9 and 12 months of age, infants can reliably follow pointing gestures to nearby objects (Bates, 1976). At about 14–15 months of age, children demonstrate the ability to follow pointing gestures to objects further away or outside of their visual field (Delgado et al., 2002; Desrochers et al., 1995). These gestures are ultimately incorporated as important sources of information during word learning scenarios. For example, deictic gestures that indicate a referent in space are often related to more accurate word learning (Booth et al., 2008; Brooks & Meltzoff, 2008; Carpenter et al., 1998; Gogate et al., 2000; Kalagher & Yu, 2006; Harris et al., 1995; Vogt & Kusschke, 2017). Representational gestures too are related to improvements in word learning (Boyce, 2021; Goodwyn et al., 2000; Lüke & Ritterfeld, 2014; McGregor, 2008). But exactly how do gestures benefit word learning?

4.1 | Orienting attention

One function of gestures is to orient attention to a specific location in space. Thoermer and Sodian (2001) examined whether infants at 10 and 12 months could follow the pointing gesture of an adult over multiple trials. They found that most infants at 10 months of age did not reliably follow gestures, while they did at 12 months of age. But in a visual dishabituation task, they also examined whether infants interpreted pointing as goal-directed or referential. To do so, an experimenter presented infants with two objects. They habituated infants to movement of the experimenter, where she looked and pointed to one of the objects. They then switched the locations of the objects and examined whether infants' looking behaviours differentiated depending on whether they saw the experimenter follow a new path (with points to the same object) or point to a new object (but following the same spatial path). If infants interpreted pointing behaviour as indicating referential intent, they would exhibit longer looking times when dishabituated to the experimenter pointing to a new object compared to a new path. If they were simply following spatial motion, they would show longer looking times when they saw the experimenter point to a new path compared to a new object. Researchers found no significant difference between the conditions in 10-month-old infants, but 12-month-olds looked longer at changes in path, suggesting that pointing behaviour was not interpreted as indicating the referential intent of the experimenter but rather allowed infants to follow the spatial location of the object. They thus conclude that at 12 months, infants were more likely to follow perceptual information conveyed by pointing. Other studies have also found that deictic gestures are used to orient attention (Vigliocco et al., 2019) and, as a result, enhance word learning (Deák et al., 2000; Deák et al., 2008; Deák et al., 2018; Rader & Zukow-Goldring, 2012).

Only one study examined whether representational gestures facilitate word learning by orienting attention. Goodrich and Kam (2009) examined how 2–4-year-old children and adults used co-speech representational gestures to infer novel intransitive verb meanings. They taught participants four novel intransitive verbs either with representational gestures matching the action described by the verb, with uninformative gestures (e.g., pointing at the child), or with speech alone. Both children and adults were better able to infer the novel meanings when speech was paired with representational gestures, compared to when learning occurred with uninformative gestures or speech alone. The researchers conclude that the function of representational gestures, at least in the realm of verb learning, is mainly to direct and sustain a child's attention to a given scene.

4.2 | Extracting information

By comparison, representational gestures have more often been thought to support information extraction because they are usually schematic in nature, allowing gestures to convey what is critical in a visual scene (Aussems & Kita, 2019; de Ruiter, 2000; Kita, 2000; Kita et al., 2017; Vigliocco et al., 2019). Aussems and Kita (2019) explored how representational gestures impacted what information 3-year-olds encoded from observing various events. They found that participants who saw actors moving in unusual ways, along with gestures that represented that

movement, recalled these events better than control conditions. This suggested that seeing representational gestures while encoding events schematically highlighted critical aspects of the scene, which improved later recall.

Exposure to representational gestures not only impacts what information is extracted but also facilitates word learning (Boyce, 2021; Capone & McGregor, 2005; Lüke & Ritterfeld, 2014; McGregor et al., 2009; Mumford & Kita, 2014; Vogt & Kusschke, 2017). Capone and McGregor (2005) taught six novel labels to 2-year-olds and varied whether the labels were accompanied by representational gestures or not. They found that the toddlers retrieved the novel labels more accurately both immediately after learning and after a four-day delay when exposure occurred with the gestures. The inclusion of the representational gestures appeared to reinforce the semantic content of the novel label: the gestures may have drawn attention to the critical semantic information of the word learning problem, allowing for improved retrieval and recall.

This increased attention to relevant information also improves children's generalization of novel meanings. Mumford and Kita (2014) found that 3-year-olds used representational gestures seen during novel verb learning to guide their generalizations of those meanings. When generalizing verbs to novel scenes that could have change-of-state or manner meaning, children demonstrated a stronger manner generalization when they had been exposed to manner gestures in comparison to either end-of-state gestures or no gestures. In this way, the representational gestures appeared to help children identify which information in a visual scene was relevant or critical to the meaning of the novel verb, which facilitated learning and generalization of the verb's meaning.

4.3 | Disambiguating referents

A third main function of gestures is to disambiguate speech in context. Studies have found that deictic gestures are relied on by children to disambiguate ambiguous scenarios (Grassmann & Tomasello, 2010; Kalagher & Yu, 2006; Pomiechowska & Csibra, 2020, preprint). Grassmann and Tomasello (2010) pit mutual exclusivity and socio-pragmatic conventions against each other by presenting 2- and 4-year-olds with a novel and familiar object. In their critical condition, an experimenter pointed to a known object (e.g., cup) and requested a novel label (e.g., the *modi*). They sought to determine whether participants would prefer to follow social cues (e.g., choosing the object that was pointed to) or adhere to a mutual exclusivity bias and choose the novel object (since they already knew the word *cup* at this age). They found both age groups were more likely to provide the pointed-to object in response to the experimenter's request. This indicated that children were more likely to use gestures as a source of information, compared to lexical information like mutual exclusivity, to disambiguate the immediate environment.

This study mirrors observational research (O'Neill et al., 2005; Özçalışkan & Goldin-Meadow, 2005; Vigliocco et al., 2019). O'Neill et al. (2005) examined the gestures that English-speaking mothers used in two contexts—a counting task and a free-play session—when interacting with their 20-month-old infants. They found that mothers overwhelmingly used deictic gestures like pointing when interacting with their child (97% of gestures in the free play session, and 91% of gestures in the counting task). These gestures were then used most often to disambiguate referents in the environment. Indeed, all 12 mothers in this task used deictic gestures most often to convey disambiguating speech acts. In addition, this work found that representational gestures were used sparingly in free-play contexts and completely unused in the counting task. Similarly, a study conducted by Vigliocco et al. (2019) examined caregivers' use of gestures in a free-play session with their 2-year-olds. They found that deictic gestures were primarily used to point out target objects that were close in space to other objects.

Although representational gestures are common in natural interactions, our literature review found a few studies suggesting that they assist in the disambiguation of referents. In the free-play study by Vigliocco et al. (2019), representational gestures were found to be helpful for disambiguation in contexts when target referents were absent. Other work has shown that children can use representational gestures to disambiguate referents and learn novel verbs (Goodrich & Kam, 2009), but the argument in this work was that representational gestures function to heighten and direct attention to a certain aspect of a visual scene.

4.4 | Indicating referential intent

Finally, gestures have been shown to inform children of speakers' referential intent. Deictic gestures have often been included to indicate referential intent in communicative scenarios (Akechi et al., 2013; Behne et al., 2012; Booth et al., 2008; Csibra & Volein, 2008; Gliga & Csibra, 2009; Paulus & Fikkert, 2014; Sodian & Thoermer, 2004; Yoon et al., 2008). Gliga and Csibra (2009) investigated whether 13-month-olds appreciate deictic gestures as communicative. In one condition, they presented infants with objects located on the left and right sides behind a barrier. An experimenter pointed to and named one of the hidden objects (e.g., duck) and measured infants' looking times in response to raising the barrier and finding the object either on the indicated or the unexpected location. They found that infants looked longer when the object was presented in the unexpected location, compared to the indicated location, suggesting that they believed the experimenter to have referred to a unique object at a specific location. Furthermore, the researchers argued that a simple associative view cannot explain their results. If infants were following only the pointing motion, different looking behaviours would not be predicted for an unexpected object. In addition, if infants were only using the verbal label to anticipate the hidden object, there should not have been different looking behaviour because the expected object was always present. They argue that the only way to explain infants' looking behaviour is that the infants believed the pointing gesture and labelling to refer to a combination of object and location.

4.5 | Understanding the many functions of gesture for word learning

Taken together, prior studies point to a number of ways in which gestures facilitate word learning (see also Cook, 2018), including orienting attention, extracting information, disambiguating referents, and indicating referential intent.

As in our review of research on eye-gaze, we can ask the question: Do gestures serve a perceptually- or socially-based function, or some combination of the two? Is there a clear developmental trend that captures the many documented ways that children exploit gestural information? Similar to our conclusion about the function of eye-gaze, we do not find a clear answer. Instead, we find an overlap of functions that could be categorized as perceptually- or socially-based within the same time frame. Even at the age of 12 months and younger, researchers have concluded that infants use gestures to both orient attention and infer referential intent. This suggests that gestures do not simply serve one stable function, and there may be no inflection point in development where social cues change functions in a qualitative way.

Like eye-gaze, we also find that there are no clear, distinct boundaries between the different types of functions. Instead, there are four interrelated functions that seem to guide word learning. Booth et al. (2008) sought to determine how 2-year-olds use social cues to learn new words in a novel word-learning task. Words were taught to children in one of the following conditions: eye-gaze cues alone, gaze + point, gaze + touch, and gaze + manipulation (movement). Booth et al. reasoned that if orienting attention was the primary function of social cues, accuracy should increase as cues increased in directness (e.g., touching a referent is more unambiguous than pointing). However, they found that the only difference in learning was found between the gaze alone and gaze + point conditions. No increases in performance were observed in learning conditions that provided increasingly direct social cues. Instead, children simply seemed to attend longer during the condition that involved both gaze and touch. And notably, time spent looking at the target object accounted for small but significant amounts of test accuracy variance. What this study suggests is that information presented in multiple modalities or multiple cue types increases attention, which in turn explains some (though not all) of children's success in word learning.

Booth et al. put forth two (not mutually-exclusive) possibilities: one explanation is that those social cues *vary in the degrees to which they indicate referential intent*, and that a pointing gesture may sufficiently indicate intent for 2-year-olds. A second explanation is that social cues heighten attention to a learning scenario, and that gestures

enable deeper encoding or more elaborated associations between labels and referents. What do we make of these explanations? If we are to consider these explanations as serving social versus perceptual functions, respectively, we would be unable to determine which explanation more accurately explains how children used the social cue presented to learn the novel words. The first explanation seems to support a socially-based perspective on social cues, and the second, a perceptually-based one. The interpretations of Goodrich and Kam (2009) demonstrate how difficult it is to parse the boundaries between each proposed function. They examined how representational gestures might inform verb learning for 2–4-year-olds and find that children can use representational gestures to disambiguate the meaning of a novel verb. Is the function of the gestures presented in this task to disambiguate meanings? Not necessarily, as the authors assume that the gestures served to "direct a listener's attention to a given scene" (p. 86). This supports the idea that it is often unclear *why* a given social cue facilitates learning. To improve our understanding of why word learning is enhanced by the presence of social cues, we need to grapple with the idea that there are (at least) four different functions of social cues without distinct boundaries between them.

These studies indicate that the functions of gestures should not be categorized in a binary way as perceptually- vs. socially-based. Their functions instead both co-occur and vary depending on a number of contextual factors. Moreover, different types of gestures are likely to serve different functions: several prior studies have concluded that deictic gestures function primarily to direct attention as well as to disambiguate visual referents, representational gestures more often facilitate information extraction, and both are likely to convey referential intent. This interplay between gesture type and function has yet to be considered in a systematic examination of word learning in early childhood.

5 | IMPLICATIONS FOR BUILDING A COMPLETE THEORY OF WORD LEARNING IN CONTEXT

Spanning our review of developmental research on eye-gaze and gesture, we can make two main conclusions. First, the functions of social cues during children's word learning do not involve a clear distinction between perceptual and social functions. That is, existing research does not support the idea of a purely associative or purely social mechanism to explain the role of social cues in word learning, and a number of studies have examined the overlap between these mechanisms (Houston-Price et al., 2006; Moore et al., 1999; Pruden et al., 2006; Yu & Ballard, 2007; Yurovsky & Frank, 2017). Second, we have seen that social cues can serve a variety of functions that include orienting attention, extracting information, disambiguating referents, and indicating referential intent. But importantly, it is not straightforward to pinpoint where a function exists along this perceptual–social dimension. For example, a given cue does not seem to just direct attention or disambiguate a visual referent without also influencing what information is extracted—or vice-versa. Plus, in some cases, indicating referential intent may not even be possible without also directing attention. There are also interactions between different social cues (such as deictic or representational gestures) and their possible functions in particular learning contexts, and children's abilities to infer social information becomes more sophisticated with age.

How do these conclusions relate to those of the most closely relevant word learning theory, the Emergentist-Coalition Model (Hollich et al., 2000)? ECM is currently the model of word learning that can best accommodate some of the conclusions from our literature review. To recap, ECM consists of three central tenets: (1) children are sensitive to perceptual, social, and linguistic cues, (2) the weights of these cues change over time, and (3) principles of word learning also change over time. Our review speaks to the first two tenets. Indeed, consistent with the first tenet, prior findings point to children's robust ability to use social cues from a speaker. And like the second tenet, we can conclude that children increasingly infer social information over development. Overall, our findings are broadly consistent with ECM. However, in an effort to improve upon how the field conceives of social cues during word learning, we outline seven factors that we believe should be explicitly incorporated in theory-building.

5.1 | Age

First, the age of the child influences how social cues are used when learning new words—an idea consistent with ECM. Our review of the literature is consistent with ideas articulated previously that very young infants use social cues in a perceptually-based way (e.g., orienting attention) and that older children are more successful in using social cues for socially-based functions (e.g., indicating referential intent). These age-related changes suggest that researchers' predictions and conclusions are generally consistent with the predictions of ECM. However, ECM does not adequately contend with the dynamic functions of social cues during word learning across development. ECM primarily suggests a sequential change from perceptual functions to socially-based ones, but based upon our findings, these functions cannot be easily categorized. What is more likely, we argue, is that the functions of eye-gaze and gesture are intertwined from the beginning, spanning a range of functions that depend on a number of factors, including age. We suggest that theories of word learning should explicitly address the diversity of social cue functions both within and across ages.

5.2 | Types of social cues

One key dimension needed to explain how social cues are integrated during word learning is the type of social cue used. We find overlapping functions between eye-gaze and gesture cues, but some types of these cues are used differently than others. For example, we determine that the type of gesture (representational or deictic) may determine its function during learning. Representational gestures may be more often used to influence the information extracted from a learning event, while deictic gestures may be used to orient attention or indicate referential intent. Eye-gaze and gestures are not the only social cues that should be considered. Other cues include affect, prosody, and speaker reliability, among others (Berman et al., 2013; Doan, 2010; Herold et al., 2011; Jaswal & Neely, 2006; Kelly et al., 1996; Krogh-Jespersen & Echols, 2012; Nicely et al., 1999; Nygaard & Queen, 2008; Sobel et al., 2020). It is important to consider how different types of social cues may improve word learning both separately and simultaneously in order to obtain a comprehensive understanding of how social cues impact real-world word learning.

5.3 | Functions of social cues

In our literature review, we also find that there are four main functions of social cues: orienting attention, extracting information, disambiguating referents, and indicating referential intent. As we stated previously, these functions may or may not be separable, as it is difficult to determine (for example) how indicating referential intent can occur without also orienting attention. That is, these functions do not cleanly fit into perceptual vs. social categories; a single function, especially in natural interactions, is likely to be supported by related functions along a perceptual-to-social continuum. Researchers should aim to understand how certain functions manifest in diverse communicative contexts, and how multiple functions occur together.

5.4 | Word types

Our review primarily focused on how social cues intersect with children's learning of concrete nouns. This limited perspective precludes the potential for further variability as we consider various types of word learning goals. For example, there remains the possibility that some social cues are better suited for noun learning compared to verb learning. Research regarding the distribution of social cues and lexical types finds that caregivers

often use gaze and touch regularly when labelling objects in the environment (Custode & Tamis-LeMonda, 2020). By comparison, they do not point to and label ongoing actions with the same frequency as referring to objects (Gleitman, 1990; Tomasello & Kruger, 1992). To our knowledge, there has not yet been a study examining the different functions of social cue for learning different word classes (e.g., noun, verb, and adjective). It is quite likely that social cues serve differentiable roles as children begin homing in on the meanings of different word types.

5.5 | Time-scales of learning

Prior research has not focused on investigating another important aspect of word learning: how children learn from social cues across shorter and longer time-scales, from seconds to years. Our examination of the literature has only reviewed language learning in the context of naming events, and considered the information provided by social cues in reference to these smaller time-scales. These time-scales are typically referred to as 'fast-mapping' scenarios, where labels and referents are initially associated in memory in a first encounter (Carey, 1978). However, researchers have also considered the importance of viewing word learning as a culmination of learning events rather than as a singular exposure (Wojcik et al., 2022). This extended period of learning enriches and strengthens the associations created, known as slow-mapping (Carey, 1978; Swingley, 2010). Studies examining slow-mapping find that different underlying mechanisms influence this learning, compared to fast-mapping (Horst et al., 2011; McMurray et al., 2012; Munro et al., 2012; Twomey, 2013). Thus, it is imperative to consider that if social cues have been shown to improve word learning in a fast-mapping context, do they also facilitate slow-mapping between labels and objects, attributes, and events? Some studies have begun to examine this. For example, Capone and McGregor (2005) found that representational gestures improved word retrieval for novel nouns at both fast-mapping and slow-mapping intervals (i.e., on average 12 days between testing). Deictic gestures, on the other-hand, may not: in a control condition, Axelsson et al. (2012) found that pointing gestures facilitated fast-mapping, but not slow-mapping, novel labels. A deeper examination of the effect of social cues across larger time-scales of learning, using methods that include diverse retention intervals during exposure to novel labels, would help create a more complete view of the functions of such cues.

5.6 | Multiple levels of linguistic input in naturalistic settings

Another consideration is that the facilitation of social cues has largely been investigated in conjunction with labelling events in controlled laboratory experiments (e.g., *It's a dax*). This generally diverges from natural language input to children from their caregivers, which encompasses a variety of utterance types beyond just labelling events, including descriptive comments, imperatives, questions, and routinized interactions (Chang & Deák, 2019; Peters & Yu, 2020; Tamis-LeMonda et al., 2017, 2019). In real, complex input, children need to sift through relevant and irrelevant linguistic input and use social cues in a way that facilitates comprehension. Thus, it is important to consider language input in naturalistic vs. experimental settings, as caregiver and child behaviours may differ between the two (see Tamis-LeMonda et al., 2017 for an example). Frank et al. (2013) examined whether social cues assist in disambiguating visual referents in a child's real environment. In contrast to experimental research, they found that these cues did not cleanly delineate a given referent. Perhaps, then, social cues are not actually very helpful in real learning environments. In contrast, Cartmill et al. (2013) found individual differences in the informativity of social cues across play sessions of caregiver-child dyads, where clearer instantiations of social cues were related to more accurate identifications of referential intent. Thus, to truly understand whether and how children capitalize on social cues in their environment, we need more research in naturalistic settings to capture how social cues occur within and across different words, sentence types, and communicative acts.

5.7 | Individual differences

Finally, theories will need to grapple with individual differences in how children use social cues in service of word learning. Individual differences in how reliably infants follow the gaze of another have been positively related to their vocabulary growth at later ages (Adamson et al., 2004; Delgado et al., 2002; Morales et al., 1998; Morales, Mundy, Delgado, Yale, Messinger, et al., 2000; Mundy et al., 2007; Mundy & Gomes, 1998; Smith & Ulvund, 2003; Tomasello & Todd, 1983). Likewise, observing gestures during learning has been shown to positively enhance word learning (Capone & McGregor, 2005, Kapalková et al., 2016; Rowe et al., 2013). For example, Rowe et al. (2013) found that gestures were most beneficial among learners with less advanced language skills. Thus, investigations into variation in the use of social cues during word learning are necessary to create a clearer picture of early language development. Beyond just how well a given child exploits a social cue, there may be important variation in how different children interpret social cues in the first place. That is, they may differ in how much they interpret a social cue as perceptually- or socially-based, and there may be an advantage to learning sooner in development that social cues provide rich social information, such that children would better encode and retain word-object mappings. This form of variation could explain why children demonstrate varying degrees of learning in experimental tasks (e.g., Axelsson et al., 2012).

Furthermore, it will be important to consider variation in the use of social cues across different populations of children. For example, some studies have demonstrated that bilingual children seem to have a heightened sensitivity to social cues when learning new words, relative to monolinguals (Brojde et al., 2012; Groba et al., 2018; Yow et al., 2017; Yow & Markman, 2011). One idea behind this is that bilingual children may interpret these cues more readily as conveying social information, yielding an advantage in learning novel words in experiments. That is, given the complexities of dual-language use and/or the prevalence of communication breakdowns, they may need to learn from a young age that social cues can be used to navigate their social environment. This may lead bilingual children to capitalize on learning opportunities afforded by social cues more often than monolingual children. Thus, a complete theory of word learning will need to consider the range of factors that contribute to individual differences in children's interpretation and use of social cues. Going forward, it is necessary to acknowledge that social cues are not monolithic; they offer different information to different children in different circumstances.

6 | CONCLUSION

Social cues have been shown to support children's word learning, and the current review aimed to understand why this is the case. We propose that social cues such as eye-gaze and gestures serve at least four different functions: orienting attention, extracting information, disambiguating referents, and indicating referential intent. Similarly, social cues can serve different functional roles: while there may be broad age-related changes in how children use social cues for perceptual and social purposes, there is also evidence that infants younger than 12 months can use social cues both to guide attention *and* to read relevant social information. Contrary to implicit assumptions in prior studies, we argue that eye-gaze and gesture can, depending on the context, provide mostly perceptual information, mostly social information, or a combination of the two. However, there is considerable complexity left to be unexplained to gain a comprehensive view of children's word learning in a social world. Experimental and observational studies will need to scrutinize how diverse social cues may serve different purposes for different children's word learning across learning contexts, sentence contexts, and developmental time.

AUTHOR CONTRIBUTIONS

Crystal Lee: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; validation; visualization; writing – original draft; writing – review and editing. **Casey Lew-Williams:** Funding acquisition; investigation; project administration; resources; supervision; validation; writing – review and editing.

CONFLICT OF INTEREST

We have no conflicts of interest to disclose.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/icd.2372.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Crystal Lee https://orcid.org/0000-0001-5452-6997

Casey Lew-Williams https://orcid.org/0000-0002-8781-4458

REFERENCES

Adamson, L. B., Bakeman, R., & Deckner, D. F. (2004). The development of symbol-infused joint engagement. *Child Development*, *75*(4), 1171–1187. https://doi.org/10.1111/j.1467-8624.2004.00732.x

Akechi, H., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2013). Brief report: Pointing cues facilitate word learning in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*(1), 230–235. https://doi.org/10.1007/s10803-012-1555-3

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, *67*(2), 635–645. https://doi.org/10.1111/j.1467-8624.1996.tb01756.x

Akhtar, N., & Gernsbacher, M. A. (2007). Joint attention and vocabulary development: A critical look. *Language and Linguistics Compass*, *1*(3), 195–207. https://doi.org/10.1111/j.1749-818X.2007.00014.x

Akhtar, N., & Tomasello, M. (1998). Intersubjectivity in early language learning and use. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 316–335). Cambridge University Press.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford University Press.

Aussems, S., & Kita, S. (2019). Seeing iconic gestures while encoding events facilitates children's memory of these events. *Child Development*, *90*(4), 1123–1137. https://doi.org/10.1111/cdev.12988

Axelsson, E., Churchley, K., & Horst, J. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. *Frontiers in Psychology*, *3*, 88. https://doi.org/10.3389/fpsyg.2012.00088

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*(5), 875–890. https://doi.org/10.1111/j.1467-8624.1991.tb01577.x

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*(5), 832–843. https://doi.org/10.1037/0012-1649.29.5.832

Baldwin, D. A. (1996). Understanding the link between join attention and language. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development*. Lawrence Erlbaum Associates, Inc.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, *67*(6), 3135–3153. https://doi.org/10.1111/j.1467-8624.1996.tb01906.x

Bannard, C., & Tomasello, M. (2012). Can we dissociate contingency learning from social learning in word acquisition by 24-month-olds? *PLoS One*, *7*(11), e49881. https://doi.org/10.1371/journal.pone.0049881

Bates, J. E. (1976). Effects of children's nonverbal behavior upon adults. *Child Development*, *47*(4), 1079–1088. https://doi.org/10.2307/1128445

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, *30*(3), 359–375. https://doi.org/10.1111/j.2044-835X.2011.02043.x

Beier, J. S., & Carey, S. (2014). Contingency is not enough: Social context guides third-party attributions of intentional agency. *Developmental Psychology*, *50*(3), 889–902. https://doi.org/10.1037/a0034171

Berman, J. M. J., Graham, S. A., Callaway, D., & Chambers, C. G. (2013). Preschoolers use emotion in speech to learn new words. *Child Development*, *84*(5), 1791–1805. https://doi.org/10.1111/cdev.12074

Booth, A. E., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-pragmatics and attention: Contributions to gesturally guided word learning in toddlers. *Language Learning and Development*, *4*(3), 179–202. https://doi.org/10.1080/15475440802143091

Boyce, K. C. M. (2021). *Can you see what I mean? The effects of iconic gestures on early word learning* [Master's thesis, Dalhousie University]. DalSpace.

Brojde, C. L., Ahmed, S., & Colunga, E. (2012). Bilingual and monolingual children attend to different cues when learning new words. *Frontiers in Psychology*, *3*, 155.

Brooker, I., & Poulin-Dubois, D. (2013). Is a bird an apple? The effect of speaker labeling accuracy on infants' word learning, imitation, and helping behaviors. *Infancy*, *18*(s1), E46–E68. https://doi.org/10.1111/infa.12027

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, *8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, *35*(1), 207–220. https://doi.org/10.1017/S030500090700829X

Bruner, J. (1983). Play, thought, and language. *Peabody Journal of Education*, *60*(3), 60–69.

Butler, S. C., Caron, A. J., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development*, *1*(4), 359–377. https://doi.org/10.1207/S15327647JCD0104_01

Capone, N. C., & McGregor, K. K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research*, *48*(6), 1468–1480. https://doi.org/10.1044/1092-4388(2005/102)

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). MIT Press.

Caron, A. J., Caron, R. F., Caldwell, R. C., & Weiss, S. J. (1973). Infant perception of the structural properties of the face. *Developmental Psychology*, *9*(3), 385–399. https://doi.org/10.1037/h0034910

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–174. https://doi.org/10.2307/1166214

Cartmill, E. A., Armstrong, B. F., III, Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences*, *110*(28), 11278–11283.

Çetinçelik, M., Rowland, C. F., & Snijders, T. M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology*, *11*, 3627. https://doi.org/10.3389/fpsyg.2020.589096

Chang, L. M., & Deák, G. O. (2019). Maternal discourse continuity and infants' actions organize 12-month-olds' language exposure during object play. *Developmental Science*, *22*(3), e12770. https://doi.org/10.1111/desc.12770

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, *112*(2), 347–382. https://doi.org/10.1037/0033-295X.112.2.347

Cook, S. W. (2018). Chapter four - enhancing learning with hand gestures: Potential mechanisms. In K. D. Federmeier (Ed.), *Psychology of learning and motivation, Vol. 69* (pp. 107–133). Academic Press. https://doi.org/10.1016/bs.plm.2018.10.001

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, *34*, 28–38. https://doi.org/10.1037//0012-1649.34.1.28

Custode, S. A., & Tamis-LeMonda, C. (2020). Cracking the code: Social and contextual cues to language input in the home environment. *Infancy*, *25*(6), 809–826.

Csibra, G., & Volein, Á. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology*, *26*(1), 1–11. https://doi.org/10.1348/026151007X185987

De Ruiter, J. P. (2000). 14 The production of gesture and speech. *Language and Gesture*, *2*, 284.

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18- month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology*, *36*(4), 511–523. https://doi.org/10.1037/0012-1649.36.4.511

Deák, G. O., Krasno, A. M., Jasso, H., & Triesch, J. (2018). What leads to shared attention? Maternal cues and infant responses during object play. *Infancy*, *23*(1), 4–28. https://doi.org/10.1111/infa.12204

Deák, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science*, *17*(2), 270–281. https://doi.org/10.1111/desc.12122

Deák, G. O., Walden, T. A., Yale Kaiser, M., & Lewis, A. (2008). Driven from distraction: How infants respond to parents' attempts to elicit and re-direct their attention. *Infant Behavior and Development*, *31*(1), 34–50. https://doi.org/10.1016/j.infbeh.2007.06.004

Delgado, C. E. F., Mundy, P., Crowson, M., Markus, J., Yale, M., & Schwartz, H. (2002). Responding to joint attention and language development. *Journal of Speech, Language, and Hearing Research*, *45*(4), 715–719. https://doi.org/10.1044/1092-4388(2002/057

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-months-old infants. *Developmental Psychology*, *47*(6), 1499–1503. https://doi.org/10.1037/a0025659

D'Entremont, B. (2000). A perceptual–attentional explanation of gaze following in 3- and 6-month-olds. *Developmental Science*, *3*(3), 302–311. https://doi.org/10.1111/1467-7687.00124

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569–572. https://doi.org/10.1016/S0163-6383(97)90048-5

Desrochers, S., Morissette, P., & Ricard, M. (1995). Two perspectives on pointing in infancy. In *Joint attention: Its origins and role in development* (pp. 85–101). Lawrence Erlbaum Associates, Inc.

Doan, S. N. (2010). The role of emotion in word learning. *Early Child Development and Care*, *180*(8), 1065–1078. https://doi.org/10.1080/03004430902726479

Farrant, B. M., & Zubrick, S. R. (2012). Early vocabulary development: The importance of joint attention and parent-child book reading. *First Language*, *32*(3), 343–364. https://doi.org/10.1177/0142723711422626

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, *16*(8), 1320–1326. https://doi.org/10.1162/0898929042304787

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), i–185. https://doi.org/10.2307/1166093

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, *9*(1), 1–24. https://doi.org/10.1080/15475441.2012.707101

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*(3), 347–353. https://doi.org/10.1111/j.1467-9280.2009.02295.x

Gliga, T., Elsabbagh, M., Hudry, K., Charman, T., & Johnson, M. H. (2012). Gaze following, gaze reading, and word learning in children at risk for autism. *Child Development*, *83*(3), 926–938. https://doi.org/10.1111/j.1467-8624.2012.01750.x

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878–894.

Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to maternal multimodal naming by 6- to 8-month-old infants and learning of word–object relations. *Infancy*, *9*(3), 259–288. https://doi.org/10.1207/s15327078in0903_1

Goldin-Meadow, S., & Singer, M. (2003). From children's hands to adults' ears: Gesture's role in the learning process. *Developmental Psychology*, *39*, 509–520. https://doi.org/10.1037/0012-1649.39.3.509

Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Baby wordsmith: From associationist to social sophisticate. *Current Directions in Psychological Science*, *15*(1), 30–33. https://doi.org/10.1111/j.0963-7214.2006.00401.x

Goodrich, W., & Kam, C. L. H. (2009). Co-speech gesture as input in verb learning. *Developmental Science*, *12*(1), 81–87. https://doi.org/10.1111/j.1467-7687.2008.00735.x

Goodwyn, S. W., Acredolo, L. P., & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior*, *24*, 81–103.

Graham, S. A., Nilsen, E. S., Collins, S., & Olineck, K. (2010). The role of gaze direction and mutual exclusivity in guiding 24-month-olds' word mappings. *British Journal of Developmental Psychology*, *28*(2), 449–465. https://doi.org/10.1348/026151009X424565

Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, *13*(1), 252–263. https://doi.org/10.1111/j.1467-7687.2009.00871.x

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839–848. https://doi.org/10.1111/j.1467-7687.2009.00945.x

Groba, A., De Houwer, A., Mehnert, J., Rossi, S., & Obrig, H. (2018). Bilingual and monolingual children process pragmatic cues differently when learning novel adjectives. *Bilingualism: Language and Cognition*, *21*(2), 384–402.

Harris, M., Barlow-Brown, F., & Chasin, J. (1995). The emergence of referential understanding: Pointing and the comprehension of object names. *First Language*, *15*(1), 19–34. https://doi.org/10.1177/014272379501500101

Hembacher, E., deMayo, B., & Frank, M. C. (2020). Children's social information seeking is sensitive to referential ambiguity. *Child Development*, *91*(6), e1178–e1193. https://doi.org/10.1111/cdev.13427

Herold, D. S., Nygaard, L. C., Chicos, K. A., & Namy, L. L. (2011). The developing role of prosody in novel word interpretation. *Journal of Experimental Child Psychology*, *108*(2), 229–241. https://doi.org/10.1016/j.jecp.2010.09.005

Hirotani, M., Stets, M., Striano, T., & Friederici, A. D. (2009). Joint attention helps infants learn new words: Event-related potential evidence. *Neuroreport*, *20*(6), 600–605. https://doi.org/10.1097/WNR.0b013e32832a0a7c

Hoehl, S., Wahl, S., Michel, C., & Striano, T. (2012). Effects of eye gaze cues provided by the caregiver compared to a stranger on infants' object processing. *Developmental Cognitive Neuroscience*, *2*(1), 81–89. https://doi.org/10.1016/j.dcn.2011.07.015

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). I. What does it take to learn a word? *Monographs of the Society for Research in Child Development*, *65*(3), 1–16. https://doi.org/10.1111/1540-5834.00091

Horst, J., Parsons, K., & Bryan, N. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology*, *2*, 17. https://doi.org/10.3389/fpsyg.2011.00017

Houston-Price, C., Plunkett, K., & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology*, *95*(1), 27–55. https://doi.org/10.1016/j.jecp.2006.03.006

Hutman, T., Harrop, C., Baker, E., Elder, L., Abood, K., Soares, A., & Jeste, S. S. (2016). Joint engagement modulates object discrimination in toddlers: A pilot electrophysiological investigation. *Social Neuroscience*, *11*(5), 525–530. https://doi.org/10.1080/17470919.2015.1114966

Iverson, J. M., Capirci, O., Longobardi, E., & Cristina Caselli, M. (1999). Gesturing in mother-child interactions. *Cognitive Development*, *14*(1), 57–75. https://doi.org/10.1016/S0885-2014(99)80018-5

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, *17*(9), 757–758. https://doi.org/10.1111/j.1467-9280.2006.01778.x

Kalagher, H., & Yu, C. (2006). The effects of deictic pointing in word learning. In *Proceedings of the 5th International conference of development and learning*.

Kapalková, S., Polišenská, K., & Süssová, M. (2016). The role of pictures and gestures as a support mechanism for novel word learning: A training study with 2-year-old children. *Child Language Teaching and Therapy*, *32*(1), 53–64. https://doi.org/10.1177/0265659015585373

Kelly, J. F., Morisset, C. E., Barnard, K. E., Hammond, M. A., & Booth, C. L. (1996). The influence of early mother—Child interaction on preschool cognitive/linguistic outcomes in a high-social-risk group. *Infant Mental Health Journal*, *17*(4), 310–321. https://doi.org/10.1002/(SICI)1097-0355(199624)17:4<310::AID-IMHJ3>3.0.CO;2-Q

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge University Press.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245–266. https://doi.org/10.1037/rev0000059

Kopp, F., & Lindenberger, U. (2011). Effects of joint attention on long-term memory in 9-month-old infants: An event-related potentials study. *Developmental Science*, *14*(4), 660–672. https://doi.org/10.1111/j.1467-7687.2010.01010.x

Kopp, F., & Lindenberger, U. (2012). Social cues at encoding affect memory in 4-month-old infants. *Social Neuroscience*, *7*(5), 458–472. https://doi.org/10.1080/17470919.2011.631289

Krogh-Jespersen, S., & Echols, C. H. (2012). The influence of speaker reliability on first versus second label learning. *Child Development*, *83*(2), 581–590. https://doi.org/10.1111/j.1467-8624.2011.01713.x

Law, B., Houston-Price, C., & Loucas, T. (2012). Using gaze direction to learn words at 18 months: Relationships with later vocabulary. *Language Studies Working Papers*, *4*, 3–14.

Lüke, C., & Ritterfeld, U. (2014). The influence of iconic and arbitrary gestures on novel word learning in children with and without SLI. *Gesture*, *14*, 204–225. https://doi.org/10.1075/gest.14.2.04luk

Mangardich, H., & Sabbagh, M. A. (2018). Children remember words from ignorant speakers but do not attach meaning: Evidence from event-related potentials. *Developmental Science*, *21*(2), e12544. https://doi.org/10.1111/desc.12544

Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. In M. R. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition* (pp. 59–101). Lawrence Erlbaum Associates, Inc.

Markus, J., Mundy, P., Morales, M., Delgado, C. E., & Yale, M. (2000). Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Social Development*, *9*(3), 302–315.

Maurer, D., & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development*, *47*(2), 523–527. https://doi.org/10.2307/1128813

McGregor, K. K. (2008). Gesture supports children's word learning. *International Journal of Speech-Language Pathology*, *10*(3), 112–117. https://doi.org/10.1080/17549500801905622

McGregor, K. K., Rohlfing, K. J., Bean, A., & Marschner, E. (2009). Gesture as a support for word learning: The case of under. *Journal of Child Language*, *36*(4), 807–828.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831–877. https://doi.org/10.1037/a0029872

Merriman, W. E., & Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development*, *62*(6), 1288–1301. https://doi.org/10.2307/1130807

Michel, C., Wronski, C., Pauen, S., Daum, M. M., & Hoehl, S. (2019). Infants' object processing is guided specifically by social cues. *Neuropsychologia*, *126*, 54–61. https://doi.org/10.1016/j.neuropsychologia.2017.05.022

Moore, C. (2008). The development of gaze following. *Child Development Perspectives*, *2*(2), 66–70. https://doi.org/10.1111/j.1750-8606.2008.00052.x

Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology*, *35*(1), 60–68. https://doi.org/10.1037/0012-1649.35.1.60

Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology*, *16*(4), 495–503. https://doi.org/10.1111/j.2044-835X.1998.tb00767.x

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, *21*(3), 283–298. https://doi.org/10.1016/S0193-3973(99)00040-4

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Neal, R., & Schwartz, H. K. (2000). Gaze following, temperament, and language development in 6-month-olds: A replication and extension. *Infant Behavior and Development*, *23*(2), 231–236. https://doi.org/10.1016/S0163-6383(01)00038-8

Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, *21*(2), 373–377. https://doi.org/10.1016/S0163-6383(98)90014-5

Mumford, K. H., & Kita, S. (2014). Children use gesture to interpret novel verb meanings. *Child Development*, *85*(3), 1181–1189. https://doi.org/10.1111/cdev.12188

Mundy, P., Block, J., Delgado, C., Pomares, Y., Hecke, A. V. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*(3), 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, *21*(3), 469–482. https://doi.org/10.1016/S0163-6383(98)90020-0

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*(5), 269–274. https://doi.org/10.1111/j.1467-8721.2007.00518.x

Munro, N., Baker, E., Mcgregor, K., Docking, K., & Arciuli, J. (2012). Why word learning is not fast. *Frontiers in Psychology*, *3*, 41. https://doi.org/10.3389/fpsyg.2012.00041

Nelson, K. (1988). Constraints on word learning? *Cognitive Development*, *3*(3), 221–246. https://doi.org/10.1016/0885-2014(88)90010-X

Nicely, P., Tamis-LeMonda, C. S., & Bornstein, M. H. (1999). Mothers' attuned responses to infant affect expressivity promote earlier achievement of language milestones. *Infant Behavior and Development*, *22*(4), 557–568. https://doi.org/10.1016/S0163-6383(00)00023-0

Nygaard, L. C., & Queen, J. S. (2008). Communicating emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 1017–1030. https://doi.org/10.1037/0096-1523.34.4.1017

Okumura, Y., Kanakogi, Y., Kanda, T., Ishiguro, H., & Itakura, S. (2013). The power of human gaze on infant learning. *Cognition*, *128*(2), 127–133. https://doi.org/10.1016/j.cognition.2013.03.011

O'Neill, M., Bard, K. A., Linnell, M., & Fluck, M. (2005). Maternal gestures with 20-month-old infants in two contexts. *Developmental Science*, *8*(4), 352–359. https://doi.org/10.1111/j.1467-7687.2005.00423.x

Özçalışkan, Ş., & Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition*, *96*(3), B101–B113. https://doi.org/10.1016/j.cognition.2005.01.001

Paulus, M., & Fikkert, P. (2014). Conflicting social cues: Fourteen- and 24-month-old infants' reliance on gaze and pointing cues in word learning. *Journal of Cognition and Development*, *15*(1), 43–59. https://doi.org/10.1080/15248372.2012.698435

Peters, R. E., & Yu, C. (2020). Examining a developmental pathway of early word learning: From qualitative characteristics of parent speech, to sustained attention, to vocabulary size. In *42nd annual meeting of the cognitive science society*. Cognitive Science Society.

Pomiechowska, B., & Csibra, G. (2020). Nonverbal action interpretation guides novel word disambiguation in 12-month-olds. *Open Mind*, *6*, 51–76. https://doi.org/10.31234/osf.io/tkhsp

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon, E. A. (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development*, *77*(2), 266–280. https://doi.org/10.1111/j.1467-8624.2006.00869.x

Rader, N., & Zukow-Goldring, P. (2012). Caregivers' gestures direct infant attention during early word learning: The importance of dynamic synchrony. *Language Sciences*, *34*, 559–568. https://doi.org/10.1016/j.langsci.2012.03.011

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*(6), 1763–1766. https://doi.org/10.1111/j.1460-9568.2005.03986.x

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *Neuroreport*, *15*(16), 2553–2555.

Rowe, M. L., Silverman, R. D., & Mullan, B. E. (2013). The role of pictures and gestures as nonverbal aids in preschoolers' word learning in a novel language. *Contemporary Educational Psychology*, *38*(2), 109–117.

Rowe, M. L., Özçalışkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, *28*(2), 182–199. https://doi.org/10.1177/0142723707088310

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*(4), 1054–1070. https://doi.org/10.1111/1467-8624.00334

Sabbagh, M. A., & Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition*, *112*(3), 415–422. https://doi.org/10.1016/j.cognition.2009.06.005

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, *69*(1), 94–104. https://doi.org/10.1111/j.1467-8624.1998.tb06136.x

Saxon, T. F. (1997). A longitudinal study of early mother-infant interaction and later language competence. *First Language*, *17*(51), 271–281. https://doi.org/10.1177/014272379701705112

Scofield, J., & Behrend, D. A. (2008). Learning words from reliable and unreliable speakers. *Cognitive Development*, *23*(2), 278–290. https://doi.org/10.1016/j.cogdev.2008.01.003

Scofield, J., & Behrend, D. A. (2011). Clarifying the role of joint attention in early word learning. *First Language*, *31*(3), 326–341. https://doi.org/10.1177/0142723710395423

Scott, K., Sakkalou, E., Ellis-Davies, K., Hilbrink, E., Hahn, U., & Gattis, M. (2013). Infant contributions to joint attention predict vocabulary development. In *35th annual conference of the cognitive science society*. Cognitive Science Society.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671. https://doi.org/10.1016/j.cub.2008.03.059

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition*, *108*(2), 303–319. https://doi.org/10.1016/j.cognition.2008.02.009

Smith, C. B., Adamson, L. B., & Bakeman, R. (1988). Interactional predictors of early language. *First Language*, *8*(23), 143–156. https://doi.org/10.1177/014272378800802304

Smith, L., & Ulvund, S. E. (2003). The role of joint attention in later development among preterm children: Linkages between early and middle childhood. *Social Development*, *12*(2), 222–234. https://doi.org/10.1111/1467-9507.00230

Smith, L. B. (2000). An associative crane. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 51–80). Oxford University Press.

Snyder, K. A. (2010). Neural correlates of encoding predict Infants' memory in the paired-comparison procedure. *Infancy*, *15*(3), 270–299. https://doi.org/10.1111/j.1532-7078.2009.00015.x

Sobel, D., Sedivy, J., Buchanan, D., & Hennessy, R. (2011). Speaker reliability in preschoolers' inferences about the meanings of novel words. *Journal of Child Language*, *39*, 90–104. https://doi.org/10.1017/S0305000911000018

Sobel, D. M., Luchkina, E., & Tummeltshammer, K. (2020). Mechanisms for evaluating others' reliability when learning novel words. In J. B. Childers (Ed.), *Language and concept acquisition from infancy through childhood: Learning from multiple exemplars* (pp. 179–196). Springer International Publishing. https://doi.org/10.1007/978-3-030-35594-4_9

Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to goal-directed action. *Journal of Cognition and Development*, *5*(3), 289–316.

Striano, T., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology*, *3*(3), 289–299. https://doi.org/10.1080/17405620600879779

Striano, T., & Reid, V. M. (2006). Social cognition in the first year. *Trends in Cognitive Sciences*, *10*(10), 471–476. https://doi.org/10.1016/j.tics.2006.08.006

Swingley, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development*, *6*(3), 179–183. https://doi.org/10.1080/15475441.2010.484412

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech directed to infants during home activities. *Child Development*, *90*(6), 2135–2152. https://doi.org/10.1111/cdev.13089

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, *20*(6), e12456. https://doi.org/10.1111/desc.12456

Tenenbaum, E., Amso, D., Abar, B., & Sheinkopf, S. (2014). Attention and word learning in autistic, language delayed and typically developing children. *Frontiers in Psychology*, *5*, 490. https://doi.org/10.3389/fpsyg.2014.00490

Thoermer, C., & Sodian, B. (2001). Preverbal infants' understanding of referential gestures. *First Language*, *21*(63), 245–264. https://doi.org/10.1177/014272370102106303

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Erlbaum.

Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *10*(4), 401–413. https://doi.org/10.1075/prag.10.4.01tom

Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In *Language development: The essential readings* (pp. 111–128). Blackwell Publishing.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, *30*(5), 639–650. https://doi.org/10.1037/0012-1649.30.5.639

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–1463. https://doi.org/10.2307/1130423

Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, *19*(2), 311–333. https://doi.org/10.1017/S0305000900011430

Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, *23*(1), 157–176. https://doi.org/10.1017/S0305000900010138

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197–211. https://doi.org/10.1177/014272378300401202

Twomey, K. E. (2013). *An investigation of fast and slow mapping*. [Doctoral dissertation, University of Sussex]. Sussex Research Online.

Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition*, *2*(1), 59–87. https://doi.org/10.1080/13506289508401722

Vigliocco, G., Motamedi, Y., Murgiano, M., Wonnacott, E., Marshall, C., Milán-Maillo, I., & Perniss, P. (2019). *Onomatopoeia, gestures, actions and words: How do caregivers use multimodal cues in their communication to children?*. PsyArXiv. https://doi.org/10.31234/osf.io/v263k

Vogt, S., & Kusschke, C. (2017). Observing iconic gestures enhances word learning in typically developing children and children with specific language impairment. *Journal of Child Language*, *44*(6), 1458–1484. https://doi.org/10.1017/S0305000916000647

Wahl, S., Marinović, V., & Träuble, B. (2019). Gaze cues of isolated eyes facilitate the encoding and further processing of objects in 4-month-old infants. *Developmental Cognitive Neuroscience*, *36*, 100621. https://doi.org/10.1016/j.dcn.2019.100621

Wahl, S., Michel, C., Pauen, S., & Hoehl, S. (2013). Head and eye movements affect object processing in 4-month-old infants more than an artificial orientation cue. *British Journal of Developmental Psychology*, *31*(2), 212–230. https://doi.org/10.1111/bjdp.12001

Wojcik, E. H., Zettersten, M., & Benitez, V. L. (2022). The map trap: Why and how word learning research should move beyond mapping. WIREs. *Cognitive Science*, *13*(4) Portico. https://doi.org/10.1002/wcs.1596

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology*, *47*(5), 1220–1229.

Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology*, *107*(2), 118–136. https://doi.org/10.1016/j.jecp.2010.04.014

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, *105*(36), 13690–13695. https://doi.org/10.1073/pnas.0804388105

Yow, W. Q., Li, X., Lam, S., Gliga, T., Chong, Y. S., Kwek, K., & Broekman, B. F. P. (2017). A bilingual advantage in 54-month-olds' use of referential cues in fast mapping. *Developmental Science*, *20*(1), e12482. https://doi.org/10.1111/desc.12482

Yow, W. Q., & Markman, E. M. (2011). Young bilingual children's heightened sensitivity to referential cues. *Journal of Cognition and Development*, *12*(1), 12–31.

Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, *70*(13–15), 2149–2165. https://doi.org/10.1016/j.neucom.2006.01.034

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. https://doi.org/10.1016/j.cognition.2012.06.016

Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science*, *20*(2), e12349. https://doi.org/10.1111/desc.12349

How to cite this article: Lee, C., & Lew-Williams, C. (2023). The dynamic functions of social cues during children's word learning. *Infant and Child Development*, *32*(1), e2372. https://doi.org/10.1002/icd.2372