Article type : Paper

Pupillometry reveals communication-induced object expectations in 12- but not 8-months-old infants

Wiebke Pätzold & Ulf Liszkowski

Universität Hamburg, Fakultät für Psychologie und Bewegungswissenschaft

Von-Melle-Park 5, Hamburg, 20146, Germany

wiebke.paetzold@uni-hamburg.de

+49-40-42838-5409

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements: We thank all families who participated in these studies. We thank Jenny Yang and Katharina Orloff for their help in creating the stimuli and running pilot studies. We are grateful to Nicola Ballhausen, Janine Baumann, Jula Brüning-Wessels, Maria Häberlein, Saskia Kiehn, Jana Klose, Tine Maikowski and Lydia Rihm for the assistance with data collection. We thank Susanna Jeschonek-Seidel for valuable feedback in the early stages of the project.

doi: 10.1111/desc.12832

Research highlights

- We used a novel pupillometric paradigm to test for violations of expectations.
- Non-linguistic referential communication induced expectations about occluded objects at 12 months of age.
- These communication-induced object expectations were absent at 8 months of age.
- An attention-directing non-social cue to occluded sites did not induce object-expectations.

Abstract

Several interaction-based and looking-time studies suggest that 1-year-old infants understand the referential nature of deictic gestures. However, these studies have not unequivocally established that referential gestures induce object-expectations in infants prior to encountering a referent object, and have thus remained amenable to simpler attentional highlighting interpretations. The current study tested whether non-linguistic referential communication induces object expectations in infants by using a novel pupil dilation paradigm. In Experiment 1, 12-months-olds watched videos of a protagonist who either pointed communicatively towards an occluder in front of her or remained still. At test, the occluder opened to reveal one of two outcomes: an empty surface or a toy. Results showed that infants' pupils were larger for the unexpected outcome of an empty surface following a point compared to the control condition (an empty surface following no point). These differences were not caused by differences in looking times or directions. In Experiment 2, an attention-directing nonsocial control cue replaced the referential communication. The cue did direct 12-months-olds' attention to the occluder, but it did not induce an object expectation. In Experiment 3 we tested 8-months-olds in the setting of Experiment 1. In contrast to 12-

month-olds, 8-month-olds did not reveal object-expectations following communication. Findings demonstrate that communicative pointing acts induce object expectations at 12 months of age, but not at 8 months of age, and that these expectations are specific to a referential-communicative as opposed to an attention-directing nonsocial cue.

Introduction

Infants begin to follow adults' gestural attention-directing reference to entities in the environment in the first year of life (Leung & Rheingold, 1981; Mundy & Newell, 2007; Murphy & Messer, 1977). However, the underlying cognitive complexities of infants' point-following, and its ontogenetic emergence, are still contested. On the one hand, predictive relations to language acquisition and theory-of-mind suggest that point-following is causal in the emergence of higher social-cognitive skills and involves a mental understanding of others' reference (Brooks & Meltzoff, 2015) enabling a 'meeting of minds' (Bruner, 1995). On the other hand, several other species can follow the gaze of others, and sometimes also the canonical human pointing gesture (e.g. Call & Tomasello, 1994; Itakura, 2004; Range & Virányi, 2011; Tomasello & Call, 2008), suggesting that point-following involves simpler cognitive processes which are not directly related to higher social-cognitive skills and referential expectations.

Classic behavioral studies established that around the age of 10-12 months, depending on the distance and position of target stimuli, infants will shift their head and gaze towards a lateral target to which an interactant points, more often than to the opposite side (Butterworth & Cochran, 1980; Butterworth & Grover, 1988; Moore & Corkum, 1998; Scaife & Bruner, 1975). One interpretation is that infants follow the vectorial direction of a point because the cue orients their visual attention in that direction, which then results in the encounter of objects pointed at (Butterworth, 2003; Moore & Corkum, 1998). In support, more recent studies have used visual cueing paradigms and measures with a high temporal resolution, including EEG and eye tracking (Bertenthal, Boyer, & Harding, 2014; Gredebäck, Melinder, & Daum, 2010; Rohlfing, Longo, & Bertenthal, 2012). These studies establish that younger infants around 4-8 months of age orient attention covertly to a proximal stimulus that has been centrally cued. At 6 months, this cueing effect becomes specific to a hand with an extended index-finger - it does not work for a foil with a similar shape - and it appears within 100ms, but not 500ms after cue onset, which reveals that point following is initially an automatic, cue driven response (Bertenthal et al., 2014).

Looking-time studies have further established that from around their first birthdays infants relate the object to the person who pointed at it. As with reaching actions, infants interpret the point as object-directed (Woodward & Guajardo, 2002). This is also the case when infants observe a pointing interaction without being directly addressed, and, again, only when it is a point, not a fist control stimulus, presumably because only the point does direct attention to the object in the first place (Krehm, Onishi, & Vouloumanos, 2014).

Thus, while pointing directs attention to an object, and that object is then related to the person who directed attention (Paulus, 2011), communicative accounts of point comprehension posit that infants understand the meaning of the point by inferring the pointer's communicative intentions (Tomasello, Carpenter, & Liszkowski, 2007). With regard to the attention-directing component of point-following, on the referential level, the proposal is that infants expect a referent prior to encountering it (Csibra & Gergely, 2009; Gliga & Csibra, 2009). That is, the communicative pointing act should instigate an expectation about the presence of a referent before seeing the referent, which then leads infants to follow the point in search of the appropriate referent. Note that this is different from the interpretation that the cue orients attention to an object, and the object is then understood as target of the point. The key to distinguish between object-directed and referential interpretations are occlusion paradigms in which the infant sees something different than the pointer when following the point, and must draw an inference about the referent before seeing it.

Several interaction-based studies have tested the referential interpretation of point-following at 12 months of age using occlusion paradigms (for gaze-following, see Butler, Caron, & Brooks, 2009; Moll & Tomasello, 2004). Behne, Liszkowski, Carpenter, and Tomasello (2012) and Liszkowski and Tomasello (2011) have argued that in a hiding game 12-months-olds will not only follow a point to an indicated site at which an object is hidden, but expect to find the object at that site, and so uncover and retrieve it. Further, in both these studies point comprehension to occluded referents correlated with infants' own production of pointing, suggesting a bidirectional understanding of gestural reference. In support, in a looking-time study on gaze-following (Csibra & Volein, 2008) infants watched videos in which an actor turned to look to the right or left side of a surface in front of her, occluded to the infant's view. At test, the actor disappeared and the occluders were removed to reveal an object either on the indicated or the opposite side of the surface. Twelve-, and even 8-months-old infants looked longer at the empty side when it had been cued by the head turn than when it had not been cued, suggesting that referential expectations may emerge already before infants begin to point themselves.

Despite this solid body of research, several questions have remained. First, one problem is that the latter studies have remained amenable to an alternative, intermediate-level interpretation on which point-following does not involve a clear expectation about the referent object prior to finding it. Instead, the search may only be a consequence of some form of attentional highlighting. For example, the point to an occluded site, as in the Behne et al. (2012) and Liszkowski and Tomasello (2011) studies, may have served as a directive to go, or just look, to that location, but the uncovering of the object then happened independently of any a priori object expectations. Similarly, in the looking-time study by Csibra and Volein (2008), infants may have looked longer to the cued but empty location simply because the location had been cued, but not necessarily in expectation of an object. Second, it has remained unclear whether in similar contexts cues that direct attention to an occluder but are entirely non-social would also instigate similar object expectations. If so, this would rather indicate general attentional processes than a specific referential understanding. Third, it is possible that infants' point-following starts out simple and that a more complex understanding still develops across the first year of life. However, direct developmental comparisons probing differences in the cognitive complexities of point-following are still sparse.

One way to distinguish empirically between the different interpretations would be to test for object expectations while controlling for the allocation of visual attention. In that way one could exclude the leaner interpretation that the amount of allocated attention leads to searching for objects in that location. A promising method in this respect is pupillometry. When infants look at the same scene for the same amount of time, differences in pupil size can still reveal different attentional-cognitive processes underlying the processing of the scene. Pupil dilation has been a measure of cognitive effort for more than 50 years (Bradshaw, 1968; Hess & Polt, 1964; Ullwer, Ries, Foth, & Meer, 2010) and recent advances in eye tracking technology make it possible to effortlessly track infants' pupil sizes (Hepach & Westermann, 2016; Hochmann & Papeo, 2014; Sebastián-Gallés, 2013). For example, 8-months-old infants' pupils dilate to violations of object identity in impossible visual scenes (Jackson & Sirois, 2009; Sirois & Jackson, 2012), to violations of action goals at 6 months (Gredebäck & Melinder, 2010); and at 24 months to observed failures of assisting a person (Hepach, Vaish, & Tomasello, 2012).

In the current study, we designed a novel occlusion paradigm. Across three experiments, infants watched videos in which a small door occluded their line of sight, and then opened in a draw-bridge-like fashion to reveal either a toy or an empty surface. In Experiment 1, 12-month-olds watched as an actress sat behind the small door, visible to the infant, and either referred non-linguistically but communicatively (by pointing with an accompanying vocalization) or did not refer to the occluded site. If nonlinguistic referential communication indeed induces expectations about a referent object, the pointing cue should instigate a referential expectation about the presence of an object. On pointing trials, the toy outcome should thus be expected, and the empty outcome should result in a violation of that expectation. On non-pointing trials, there should be no specific expectation as to whether there is or is not an object behind the occluder.

Because the two perceptually different outcomes differed in their luminance, one cannot compare across outcome events, because pupil sizes would differ naturally. Our comparisons of interest instead focused on the referential communication manipulation. Our main comparison concerned the case of the perceptually identical empty outcomes: On the referential expectation hypothesis, infants' pupils should be more dilated in the point manipulation, when the empty outcome was a violation of a referential expectation, than in the no point manipulation, when the empty outcome was not a violation, because no expectation had been formed. The null-hypothesis was thus that infants had no referential expectations, and pupil sizes would not differ (or dilate even less). For our main analysis we thus conducted a planned comparison between the two empty outcomes when they were cued or not cued. A secondary possibility for comparison concerned the two toy outcomes. This comparison, however, does not test for a violation of an expectation. One possibility was that the appearance of an expected object in the point manipulation would result in a relative decrease in pupil diameter compared to the no point manipulation, because the communicative cue would lead to expect the state of affairs, thus requiring less cognitive effort, while in the absence of a communicative cue infants would be in a state of uncertainty and still needed to process the new state of affairs. Alternatively, the toy outcome in the point manipulation could yield no further decrease in pupil diameter compared to the no point manipulation, because infants in the no point manipulation were equally ready to expect a toy or no toy.

In Experiment 2, we addressed the question whether in a similar context a cue which directs attention to the occluder but is entirely non-social (a transient color change), would also induce object expectations. If this was the case, it would severely limit the conclusion about the comprehension of referential communication in Experiment 1. We used the same basic paradigm and measures as in Experiment 1 but used a color cue instead of a communicative pointing cue. We did not use cues that resembled pointing cues shape-wise (like a "fist-point", Krehm et al., 2014; "distracted point", Behne, Carpenter, & Tomasello, 2005; or a "foil", Bertenthal et al., 2014) because previous research has shown that these cues do not direct attention to begin with. If the alternative hypothesis that directing attention to the occluder was enough to induce an object expectation in the current context, infants should dilate their pupils in the cued empty outcome condition significantly more than in the no cued empty outcome condition, just like on the referential expectation hypothesis of Experiment 1. Further, one would expect no significant difference between the two attention-directing cues (point or color change) for the empty outcome.

In Experiment 3, we addressed the ontogenetic question about the origins of communication-induced object expectations. We re-ran Experiment 1 with 8-month-olds, who orient attention following a point, to test whether they also expect an object when following others' points.

Experiment 1

Method

Participants. Seventeen 12-months-old infants (7 males, 10 females) were included in the final sample. Mean age was 12 months, 24 days (range: 12 months, 17 days – 13 months, 1 day). Four additional infants participated (3 males, 1 female) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (1), missing data in one or more conditions during baseline or test (1) or fussiness (2). Infants provided data for a median of 10 (range 4 - 12) out of 12 trials.

All infants were recruited via birth records and had a middle to high socioeconomic, western cultural background. Infants were included in the final analysis when they provided data for at least one trial of each condition. Trials were included when looking times indicated that the child had watched at least 50% of the manipulation.

Apparatus. We measured eye gaze and pupil dilation with a Tobii x120 eye tracker (Tobii Technology, Stockholm, Sweden) which was attached to a standard 51.50 x 32.00 cm computer screen. The presentation screen and the eye tracker were placed in a testing booth built for this purpose, with black canvases behind and on both sides of the screen. A video camera above the screen allowed the researchers to monitor the infant's behavior during the session. The size of the stimuli presentation was 1280 x 1024 px on a 1920 x 1200 px screen, with the rest of the screen appearing black throughout the experiment. The size of the display was 34.50 x 27.50 cm on the screen, which corresponds to a visual angle of 32.08° horizontally and 25.81° vertically.

Stimuli. The videos consisted of recorded clips of a female actress sitting behind an empty table. During stimulus recording, props served to mark a location in the middle of the table to which the actress referred with her non-linguistic gestural communication. Additional objects were superimposed onto the recording using Adobe After Effects®. In all videos, the same clip of the pointing gesture was used. As in many other studies, the actress wore a visor to conceal her eyes. Piloting had shown that infants had difficulties disengaging from looking at the actress' eyes and face. In addition to this methodological advantage it also eliminated the possibility of following only the gaze. The edited clips showed the actress sitting behind the table. In front of her, on top of the table, there was a green door-like occluder held in a door frame superimposed onto the video. The door was animated to open and close in a draw-bridge-like fashion such that the frame remained and the outcome was revealed within the frame. When the door opened, it revealed a view through the door frame onto the table surface that was either empty or held a toy. There were 3 toys (car, helicopter and truck) in total, one for each trial. They were constructed from Lego Duplo® blocks and were similar in size. It was possible in all constellations to construe the actress as seeing the target object, even when it was occluded to the child. The vector line from the extended index-finger extended directly to the area on the screen covered by the door. In order to hide the agent from view, a white curtain was superimposed onto the video and lowered between the agent and table after the manipulation.

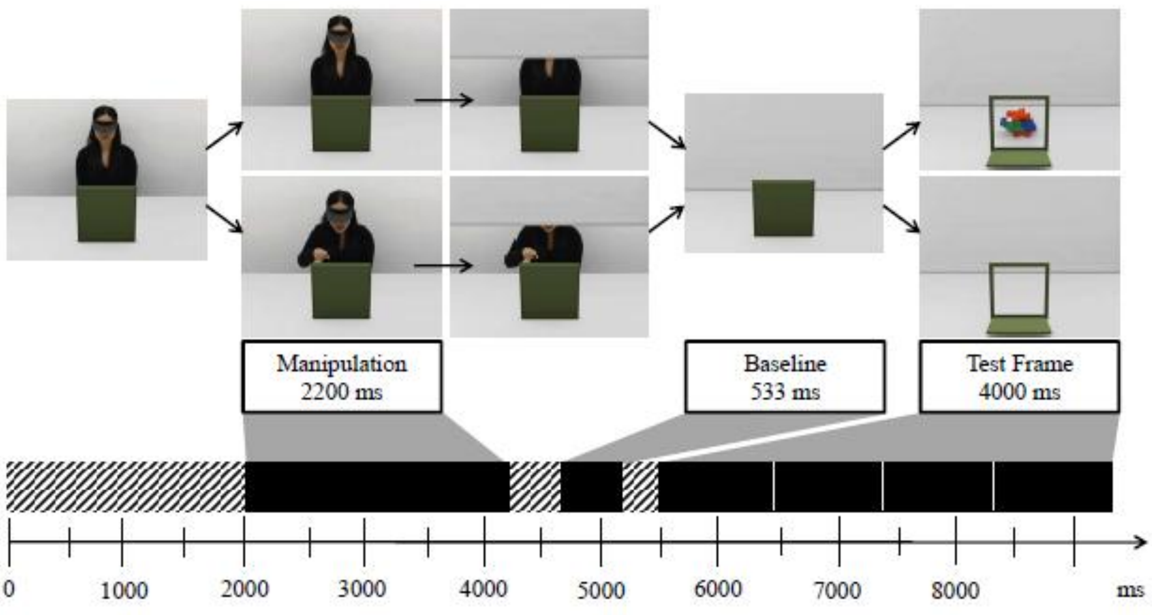


Figure 1. Schematic depiction of the Experimental Design for Experiment 1.

All trials started with a 4 second sequence of animated soap bubbles on an evenly illuminated blue background in order to create an even level of luminance for each trial (Hepach et al., 2012). The bubbles then cut to the video showing the actress sitting at the table, with the door closed in front of her from the child's view (2000 ms, see Fig.1).

There were four conditions: two types of manipulation each paired with one of two outcomes. During the manipulation phase (2200 ms), the protagonist either pointed to the area on the screen covered by the door with her right hand, saying "ah" with an excited expression ("point" manipulation) or remained still ("no point" manipulation). The intonation was child-directed and ostensive[1]. Following this manipulation, the curtain hid the agent and then the door remained closed for 533 ms, displaying a perceptually identical scene in all conditions (baseline). After that, the door opened and revealed one of two outcomes: a toy ("toy" outcome) or the empty table ("empty" outcome). After 4000 ms of showing a still frame of the outcome, the animated bubbles were presented again for 4000 ms. Each child was presented with 3 blocks each containing all 4 conditions. The sequence of presentation was counterbalanced across blocks and subjects.

Procedure. The experimenter explained the study to the caregiver and obtained their informed consent. The caregiver was seated on a swivel chair in the testing booth 64 cm away from the presentation screen with the child on his or her lap. A 9-point-calibration ensured that the child's eyes were properly captured by the eye tracker. The experimental stimuli were then presented on the screen as long as the child was willing to watch or until 12 trials were completed. The caregivers were instructed to neither talk to nor point for the child and to keep their eyes closed for the duration of the experiment.

[1] Infants are more likely to follow referential gaze when it is preceded by ostensive signals like mutual gaze or child-directed prosody, see Senju and Csibra (2008).

Data Processing. Pupil size and gaze location of both eyes were recorded at a 120 Hz rate. Pupil size from the left and right eye was averaged to mean pupil size. If data from one eye was missing, the data from the other eye was used. If data from both eyes were missing, no substitution was made. Careful visual inspection of the time line of the pupil data showed that changes in luminance in the video stimuli (e.g. bubbles cutting to the start of the scene) caused pupillary restriction (PC) with a delay of 500 ms across all conditions. This delay in response corresponds to previous reports showing that infants' pupils take about 500ms to adapt to luminance changes in stimuli (Verschoor, Paulus, Spapé, Bíró, & Hommel, 2015). We therefore placed all time windows of analyses concerning pupil data 500 ms later than the corresponding events in the video. Also, by not including the phases of PC following a stimulus onset in the average, a more accurate representation of the tonic pupil dilation is achieved (see Hepach & Westermann, 2016). Analyses focused on the respective time windows starting 500 ms after the manipulation phase, baseline phase; and test phase. For the looking time data, we preserved the temporal correspondence between events in the video and the analyzed time windows.

Looking time data was processed in relation to the door area (380 x 355 px) and the entire display (1280 x 1024 px). Time looked away from the screen was defined as the difference between the duration of the analysis window and the time looked to the display. It thus included untracked gaze due to blinks or technical failure of the eyetracker, which represents minimal noise that must be assumed to be equally distributed. For our time window of the main analyses we chose the first 1000ms of the test sequence because we reasoned that infants would pay equal amount of attention to the outcomes during the first second.[2]

[2] Because infants may tend to look away from the screen over time, we chose to report the pupil dilation of the first 1000 ms for our main analyses. However, to provide a fuller picture of the dataset, we also report secondary analyses on the entire test period of 4000 ms.

Because pupil data as collected by remote eye trackers such as the Tobii X120 may be susceptible to error introduced by gaze direction (Brisson et al., 2013), we report looking time patterns for looking at the video and at AOIs during the baseline and the test phase to control for potential influences of looking behavior on the pupil diameter. All significance tests are reported two-tailed.

Results

Preliminary analyses

Figure 2 displays the mean pupil diameters for the four conditions across a trial. Visual inspection revealed an unexpected difference between point and no point trials after the manipulation, before the outcomes were shown.

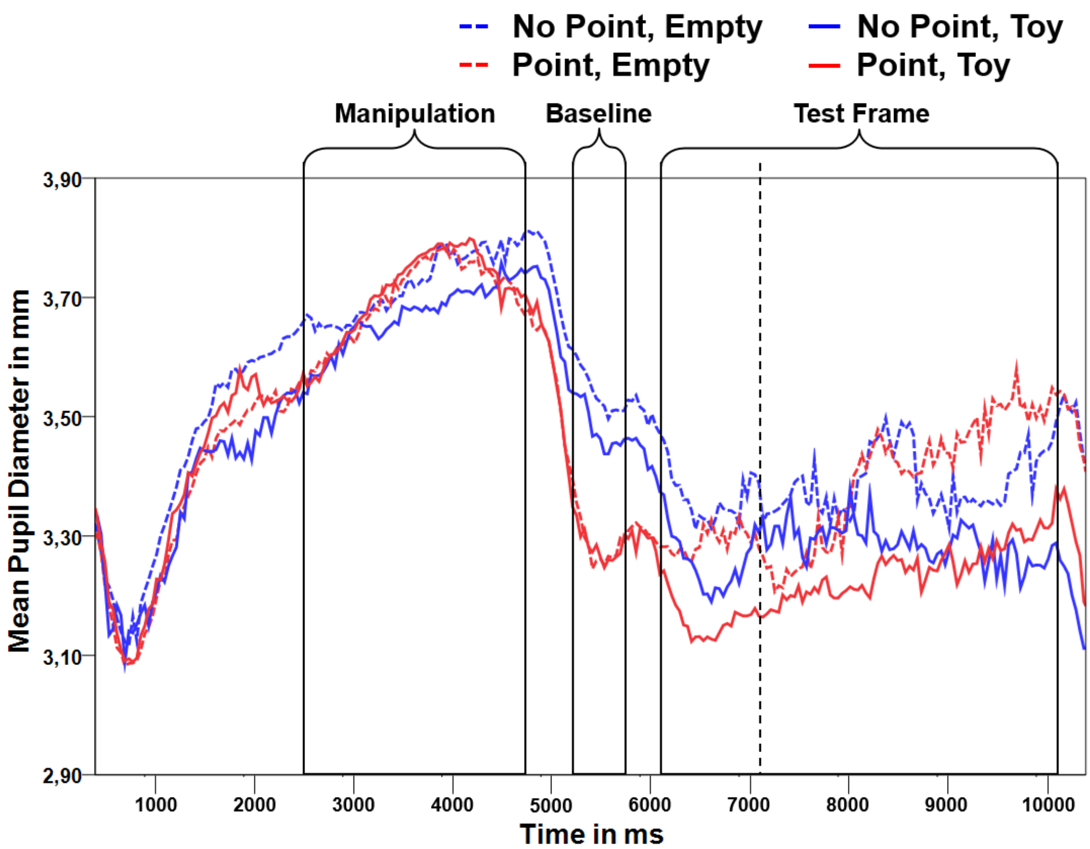
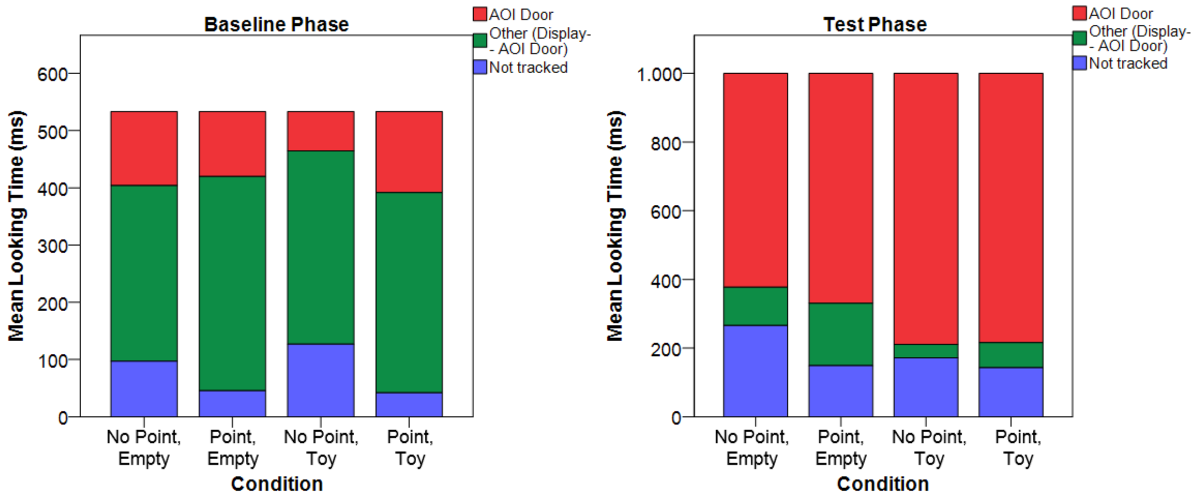


Figure 2. Mean pupil diameter in in 12-month-olds Experiment 1. Windows of analyses begin 500 ms later compared to the timeline in Fig. 1.

During the 533 ms long baseline period, the curtain has come down and the door frame is still closed, rendering videos of all conditions perceptually identical. On no point trials, infants had larger pupils than on point trials (M_{point}= 3.29 mm, SD_{point}= .38 mm, $M_{no\ point}$= 3.39 mm, $SD_{no\ point}$= .37 mm, $t_{(16)} = 6.54$, $p < .001$). Control analyses of infants' looking pattern suggest that in point trials, infants looked longer at the screen in this time period than in the no point trials, see Figure 3a, although this differences did not reach statistical significance (M_{point}= 519.37 ms, SD_{point}= 26.97 ms, $M_{no\ point}$= 489.10 ms, $SD_{no\ point}$= 81.23 ms, $t_{(16)} = 1.47$, $p = .160$).[3]



(a) Baseline phase. *(b)* 1000 ms of test phase.

Figure 3. Distribution of looking time in Experiment 1.

Main analyses

For our main analyses, in order to adjust for the different levels of pupil dilation before the outcomes, we calculated a relative change score (Hepach et al., 2012): *(Pupil size (test) – Pupil size (baseline)) / Pupil size (baseline)*. Figure 4 shows the relative change scores across conditions, averaged over the duration of the first 1000 ms. Direct comparison following our main hypothesis confirmed that in the surprising 'point, empty' condition, the pupil was relatively larger compared to the 'no point, empty' condition ($M_{point,\ empty}$= - .005,

[3] Similar patterns emerged in Experiments 2 and 3, where the differences did reach significance.

$SD_{point, empty}$= .069, $M_{no point, empty}$= -.042, $SD_{no point, empty}$= .042, $t_{(16)}$ = 5.18, p < .001). Regarding our secondary comparison, pupils in the 'point, toy' condition did not differ in their relative change compared to the 'no point, toy' condition ($M_{point, toy}$ = - .046, $SD_{point, toy}$ = .063 $M_{no point, toy}$= -.057, $SD_{no point, toy}$= .069, $t_{(16)}$ = .56, p = .582).

The difference of the 'point – empty' condition relative to the other three conditions is also apparent when comparing the relative change in pupil size to baseline (Figure 2). Pupils decreased significantly in all conditions except for the 'point, empty' condition (difference from baseline: no point, empty: $t_{(16)}$ = -4.12, p = .001, no point, toy: $t_{(16)}$ = -3.39, p = .004, point, toy: $t_{(16)}$ = -3.04, p = .008; point, empty: $t_{(16)}$ = -.56, p = .584).

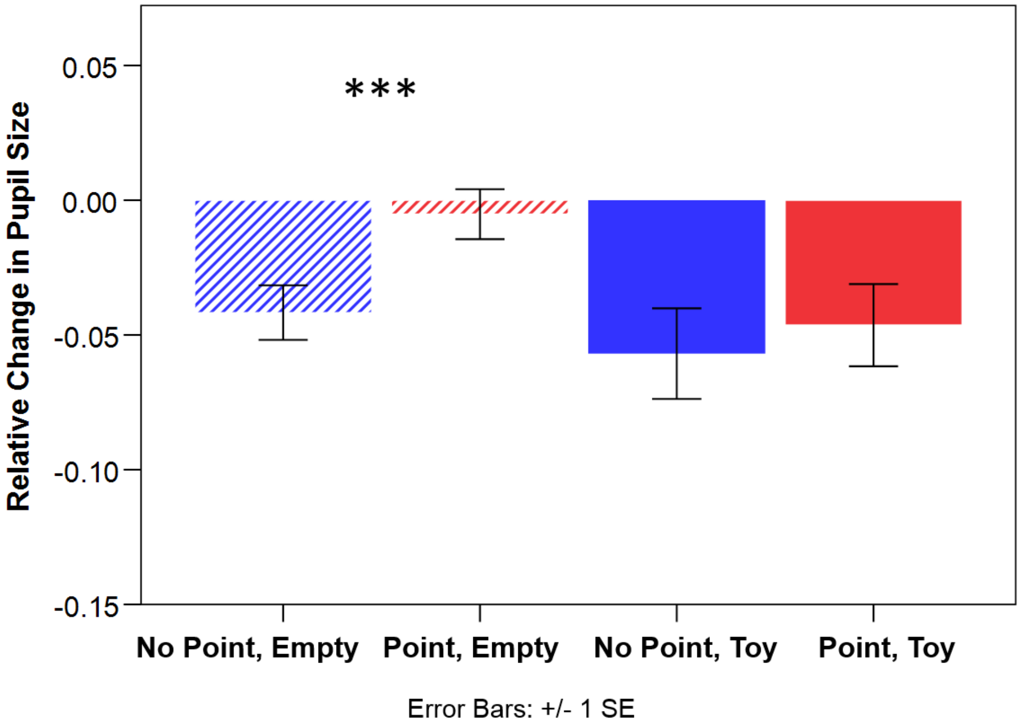


Figure 4. Mean Change in Pupil Size in Experiment 1. ***indicates highly significant difference between the two empty outcomes.

Control analyses

To control for the possibility that differences in looking patterns led to differences in pupil sizes, we analyzed looking times to the three different AOIs during the 1000 ms of test time: 1) Looks within the frame of the opened door, 2) looks at the video in general and 3) untracked looks (e.g., looks away from the screen). The latter include track loss due to blinks and technical failure which represents noise that must be assumed to be equally distributed across time and conditions. Figure 3b shows that infants looked longer to the door when there was a toy than when it was empty (main effect for presence of toy: $F_{(1, 16)} = 5.89$, $p = .027$, $\eta_p^2 = .269$). However, regarding our planned crucial comparisons within the same outcomes, the analyses revealed no significant differences in terms of looking time to the door frame, to the video, or not registered gaze points (e.g. off screen). This means that influences from adaptation to the dark testing booth, or systematic measurement deviations from diverted pupils cannot explain our pupillary finding, and that the differences in pupil size reflect differences in the cognitive processing induced by the manipulation.

Additional analyses

Since the experiment was designed with a test phase of 4000 ms, we also analyzed average relative change in pupil size for the entire duration of the test phase. Regarding our main hypothesis, the direct comparison between conditions confirmed that in the surprising 'point, empty' condition, the pupil was relatively larger compared to the 'no point, empty' condition ($M_{\text{point, empty}} = .037$, $SD_{\text{point, empty}} = .048$, $M_{\text{no point, empty}} = -.014$, $SD_{\text{no point, empty}} = .040$, $t_{(16)} = 5.02$, $p < .001$). Pupils in the 'point, toy' condition were relatively larger compared to the 'no point, toy' condition ($M_{\text{point, toy}} =- .013$, $SD_{\text{point, toy}} = .070$, $M_{\text{no point, toy}} = -.0401$, $SD_{\text{no point, toy}} = .059$, $t_{(16)} = 2.18$, $p = .045$). Also the looking time pattern remained similar: infants looked longer to the door when there was a toy than when it was empty (main effect for presence of toy: $F_{(1, 16)} = 49{,}14$, $p < .001$, ηp2 = .754). However, regarding our crucial comparisons within the same outcomes, the analyses revealed no other significant differences between any conditions of the same outcome in terms of looking time to the door frame; or to the video in total; or regarding gaze points not registered on the screen.

Although our main hypothesis focused on the comparison between the violation of expectation ('point, empty') and the perceptual control ('no point, empty'), we also report repeated measures ANOVAs on the relative change score for both the 1000 ms and the 4000 ms time windows, because a post-hoc analysis of all four conditions may be informative for a fuller picture. For both time windows there were main effects for Cue (respectively, $F_{(16)} = 4.64$, $p = .047$; $F_{(16)} = 17.66$, $p = .001$) and Outcome (respectively, $F_{(16)} = .7.55$, $p = .014$; $F_{(16)} = 13.62$, $p = .002$). The interaction between Cue and Outcome failed to reach a two-sided significance in the 4000 ms time window ($F_{(16)} = 3.11$, $p = .096$) and in the 1000 ms time window ($F_{(16)} = 2.05$, $_{\text{p}} = .172$). For both time windows, Bonferroni-corrected post hoc paired comparisons (for 6 comparisons) between all conditions confirmed the referential expectation hypothesis. For the 1000ms time window, the 'point, empty' condition was significantly different from all other three conditions, respectively 'point, toy' ($t_{(16)} = 3.86$, $p = .001$), 'no point, toy' ($t_{(16)} = 3.16$, $p = .007$) and 'no point, empty'($t_{(16)} = 5.18$, $p < .001$), and all other three conditions were not different from each other. Similarly, for the 4000 ms time window, the 'point, empty' condition was significantly different from all other three conditions, respectively 'point, toy' ($t_{(16)} = 4.33$, $p = .001$), 'no point, toy' ($t_{(16)} = 5.81$, $p < .001$) and 'no point, empty'($t_{(16)} = 5.02$, $p < .001$) and all other three conditions were not different from each other.

Discussion

Infants' pupils were relatively larger in the 'point – empty' condition, when the actress pointed behind the door but the opening of the door revealed no object, compared to the control condition, when the opening of the door also revealed no object but there had been no pointing cue. This finding is consistent with the referential expectation hypothesis that infants expect a referent object following nonlinguistic referential communication before they encounter the referent object.

Although infants' pupils were generally larger on pointing than non-pointing trials before any outcome was presented, our additional analyses excluded the possibility that our main finding of relatively larger pupils in the 'point – empty' condition was solely driven by this initial difference between pointing and non-pointing trials. Firstly, at test infants' pupils for the toy outcome were not significantly different between pointing and non-pointing trials, thus rejecting the possibility that pupil size at test was solely driven by the differences after the manipulation. Secondly, by calculating the relative change from baseline to test and comparing it across conditions, we accounted for the differences at baseline. Further, our control analyses on the looking times at test confirmed that infants focused equally long to the same areas in the stimulus material across conditions, thus excluding that differences in pupil size were spurious and just driven by differences in luminance or pupil orientation during test (Brisson et al., 2013).

Although not target of our investigation, we note that non-pointing trials led to larger pupils compared to pointing trials right after the manipulation, before any outcome was shown. One interpretation in line with the referential expectation hypothesis is that pointing reduces uncertainty about the next sequence, given a general expectation of referent objects following pointing, hence resulting in a relatively smaller pupil size. Alternatively, it may be that the cue oriented attention to the screen more so than no cue, resulting in pupil change when looking less onto the screen. While not significant, our control analyses on looking times may suggest descriptively that infants' gaze on non-pointing trials was less often registered on the screen.

Pupil size in the 'point, empty' condition did not increase relative to the preceding still phase sequence after the manipulation. However, this lack of relative increase must be interpreted in the perceptual context of the stimulus material and its change in luminance which induced a significant pupil constriction in all other conditions. The crucial analysis pertained to the relative change compared to the control manipulation, in which the pupil size decreased to a significantly larger extent than in the 'point, empty' condition, as was the case for the other two conditions.

The current findings exclude leaner alternative interpretations of infant point-following as an orienting response (Bertenthal et al., 2014), or that infants search for referent objects at indicated locations due to differences in attentional allocation (Behne et al., 2012). Instead, by 12 months of age, infants have developed referential expectations which enable them to infer the referent of a non-linguistic referential communicative gesture. It remains unknown from the current experiment whether this induced object expectation effect is specific to the communicative-referential pointing act, and perhaps primarily driven by one or several of its distinct facial-vocal-affective components that make pointing a referential communicative act, or whether a non-social cue that directs attention successfully to the occluder but lacks any of these components would equally induce object expectations. A future task could be to unravel which of the communicative referential cues that we employed in our manipulation may be sufficient or necessary to instigate referential object expectations. However, to first establish that it is indeed at all the communicative referential nature of the cue which induced object expectations, it is necessary to investigate the effect with an attention-directing cue that is non-social and non-communicative, i.e. a cue that does direct attention to the occluder, but not for referential communicative reasons. We investigated this question in Experiment 2. A further question is whether the induced object expectations at 12 months of age appear already earlier in development, perhaps even before infants begin to point (Gredebäck & Melinder, 2010). We addressed that second question in Experiment 3.

Experiment 2

Method

Participants. Twenty-three 12-months-old infants (13 males, 10 females) were included in the final sample. Mean age was 12 months, 12 days (range: 12 months, 3 days – 12 months, 28 day). Twenty-seven additional infants participated (16 males, 11 females) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (12), missing data in one or more conditions during baseline or test (4), technical error (1) or fussiness (10). On average, each infant provided data for a median of 9 (range 4 - 12) out of 12 trials. Drop-out rates were higher in this experiment compared to experiment 1, most likely due to the nonsocial, non-communicative nature of the stimulus material.

Stimuli. The videos consisted of a tabletop scene identical to the one in Experiment 1, but without a human agent. Instead of the pointing gesture, in the experimental manipulation a red circle appeared and disappeared on the closed door, accompanied by a bell chiming, directing attention to the area on the screen covered by the door. The timing of this nonsocial audiovisual cue was identical to the pointing gesture and vocalization in Experiment 1. When the door opened, it revealed a view on the table surface that was either empty or held a toy, identical to the ones in Experiment 1. In the control manipulation, no cue appeared. The videos in Experiment 2 were presented in exactly the same way as in Experiment 1, including the presentation of bubbles in the beginning and the end, the order, and the number of trials. Apparatus and Procedure were identical to Experiment 1.

Data Processing. Data processing was the same as in Experiment 1. As in Experiment 1, we analyzed looking behavior to control for possible influences on pupil diameter.

Results

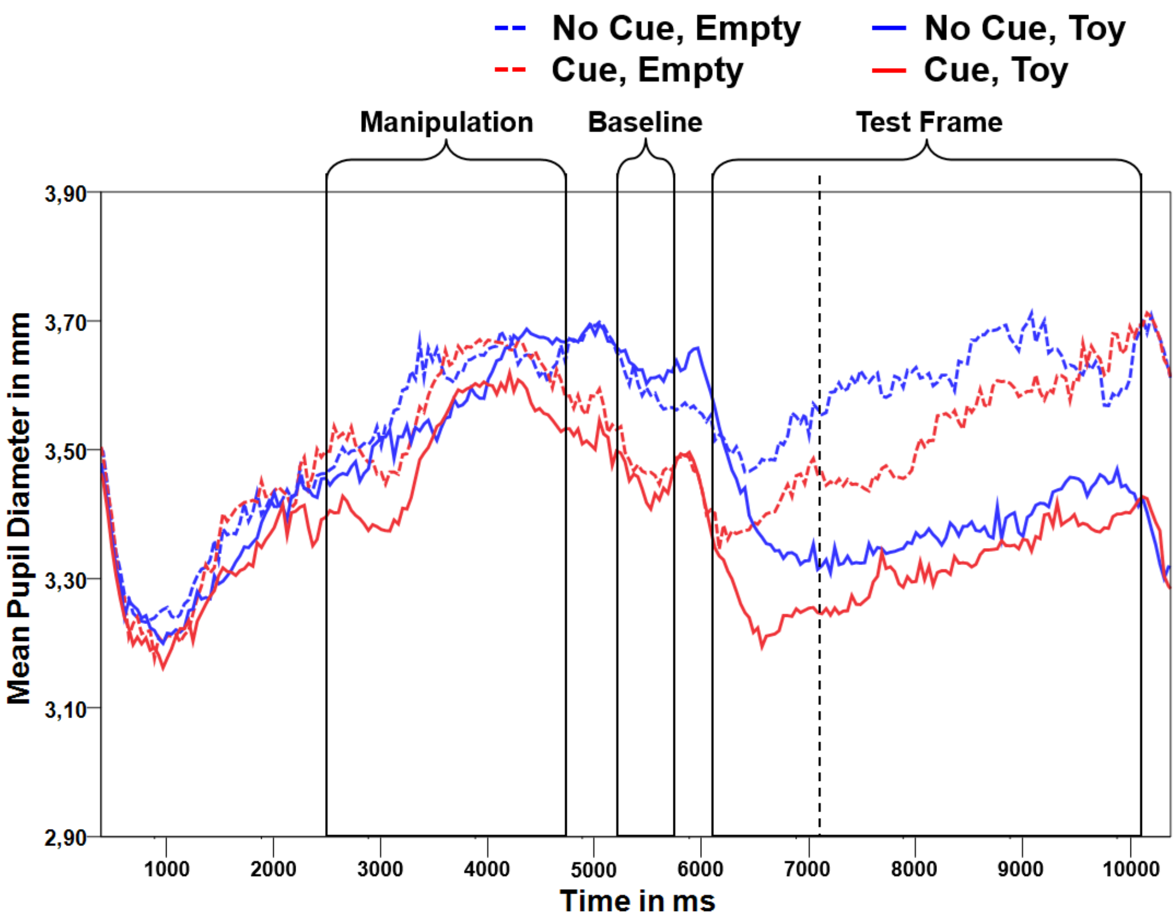
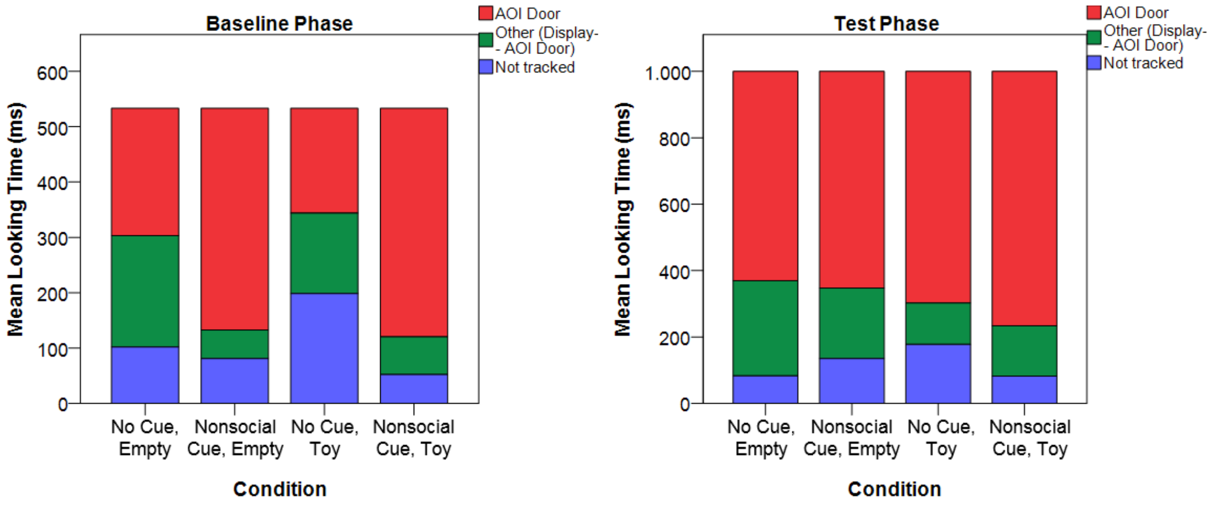


Figure 5. Mean pupil diameter of 12-month-olds in Experiment 2. Windows of analyses begin 500 ms later compared to the timeline in Fig. 1.

Preliminary analyses

Figure 5 displays the mean pupil diameters for the four conditions across a trial. As in Experiment 1, visual inspection revealed a difference in pupil size right after the manipulation, already before the outcome. The nonsocial cue manipulation led to smaller pupils during baseline compared to the no cue control manipulation (M_{cue} = 3.43 mm, SD_{cue} = .37 mm, $M_{no\ cue}$ = 3.58 mm, $SD_{no\ cue}$= .37 mm, $t_{(22)}$ = 2.99, p = .007). A control analysis of looking time during baseline, revealed that infants spent significantly more time looking at the screen when there was a cue than when there was no cue (M_{cue} = 466.85 ms, SD_{cue} = 126.92 ms, $M_{no\ cue}$ = 379.15 ms, $SD_{no\ cue}$ = 90.53 ms, $t_{(22)}$ = 2.77 p = .011), see Figure 6a. Assuming that failed tracking and blinks were equally distributed, this suggests that infants in the no cue manipulation looked more often away from the screen, which may have resulted in larger pupils as their eyes adapted to the darker surrounding of the testing booth.



(a) Baseline phase. *(b)* 1000 ms of test phase.

Figure 6. Distribution of looking time in Experiment 2.

Main analyses

Given our preliminary analyses, we applied the same baseline correction as in Experiment 1. Contrary to the alternative hypothesis, a direct comparison showed no significant differences in relative changes between the two manipulation conditions for the empty outcome, and the means were in the opposite direction ($M_{cue, empty}$= -.026, $SD_{cue, empty}$ = .051, $M_{no\ cue, empty}$ = -.012, $SD_{no\ cue, empty}$ = .042, $t_{(22)}$ = 1.43, p = .166). Regarding our secondary comparison, there was also no difference between manipulations for the toy outcome ($M_{cue, toy}$ = -.059, $SD_{cue, toy}$ = .062, $M_{no\ cue, toy}$ = -.060, $SD_{no\ cue, toy}$= .050, $t_{(22)}$ = .092, p = .927), see Figure 7.

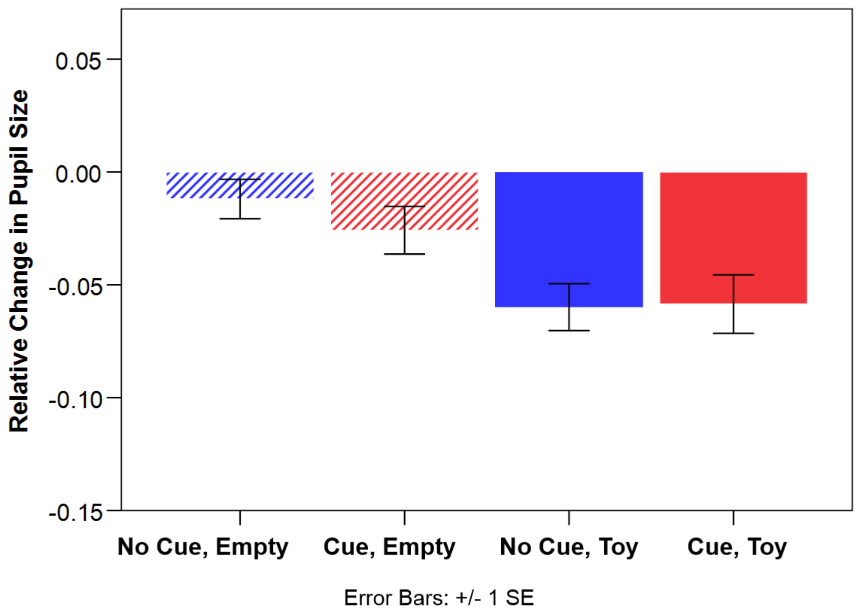


Figure 7. Mean Change in Pupil Size in Experiment 2.

Control analyses

Again, we checked for looking times to different AOIs during the test time in order to control for the influence of gaze direction on the pupil (Fig. 6b). An analysis of looking patterns revealed no significant differences between any conditions in terms of looking time to or away from the screen. As in Experiment 1, infants looked longer at the door AOI when there was a toy than when it was empty ($F_{(1, 22)}$ = 10.76, p = .003, η_p^2 = .328). However, there was no interaction between the two factors, ruling out that the toy effect varied systematically with manipulation, and no significant differences concerning our crucial main comparisons between the two similar outcomes.

Additional analyses

As in Experiment 1, we repeated the analyses with data averaged across the entire test phase of 4000 ms, to test for any prolonged effects. Repeated measures t-Test following our hypotheses revealed no differences between conditions for the empty outcomes ($M_{cue, empty}$=.005, $SD_{cue, empty}$=.061, $M_{no\ cue, empty}$= = .008, $SD_{no\ cue, empty}$= = .034, $t_{(22)}$ = .17, p = .870) and for the toy outcomes ($M_{cue, toy}$ = -.039, $SD_{cue, toy}$ = .055, $M_{no\ cue, toy}$ =-.057, $SD_{no\ cue, toy}$ =.047, $t_{(22)}$ = -.24, p = .228).

Infants looked longer to the open door frame over the duration of 4000 ms when there was a toy than when it was empty (main effect for presence of toy: $F_{(1, 22)}$ = 29.69, p < .001, η_p^2 = .574), however, regarding our crucial comparisons, the analysis revealed no significant differences between any conditions of the same outcome in terms of looking time in total or gaze points not registered on the screen.

Although our main comparison focused on testing for a violation of expectation effect between the 'cue, empty' condition and the perceptual control condition ('no cue, empty' condition), we also report repeated measures ANOVAs on the relative change score for both 4000 ms and 1000 ms time windows to provide a fuller picture, as in Experiment 1. For both time windows there were main effects for Outcome (respectively $F_{(16)}$ = 19.92, p < .001; $F_{(16)}$ = 35.58; p < .001), but no main effects for Cue (F's < .60) and no interactions (respectively $F_{(16)}$ = 1.21, p = .284; $F_{(16)}$ = 1.35, p = .259).

For the 1000 ms time window, Bonferroni-corrected posthoc paired comparisons (for 6 comparisons) between all conditions showed significant differences for two comparisons: ‘no cue, toy’ and ‘cue, toy’ were both significantly smaller than ‘no cue, empty’ ($t_{(22)} = 4.\ 18$, $p < .001$ and $t_{(22)} = 3.46$, $p = .002$; respectively), reflecting the main effect of outcome. No other comparisons were significant. For the 4000 ms time window, Bonferroni-corrected paired comparisons (for 6 comparisons) between all conditions showed a significant difference only for the comparison ‘no cue, toy’ and ‘no cue, empty’ ($\mathrm{t}_{(22)} = 4.95$, $p < .001$) and no other significant differences between conditions, again likely reflecting the main effect of outcome.

Comparison between Experiment 1 and Experiment 2

Experiments 1 and 2 aimed to establish with planned experimental comparisons the presence or absence of significant condition differences predicted by the referential expectation hypothesis, rather than directly compare the size of all possible effects (including those due to luminance differences, or presence/absence of actress) across Experiment 1 and Experiment 2. Of relevance, however, may be a comparison of the difference between the cued and the not cued empty outcomes across the two experiments. If the effect of Experiment 1 was selective to the communicative referential act, then it should be significantly greater compared to any difference induced by the attention-directing non-communicative cue of Experiment 2. Indeed, a direct t-test comparison between the Experiments on the mean difference between the cue-empty and no cue-empty outcomes based on relative difference scores revealed a significantly greater pupil increase in Experiment 1 than Experiment 2 ($t_{(38)} = 3.94$, $p < .001$).

Discussion

Experiment 2 showed that the cue effectively directed attention to an area, as in Experiment 1, but it did not induce an expectation about a referent object. Infants' reaction to the outcome scene did not vary as a function of the previously seen manipulation. Therefore, we conclude that it is neither the amount of allocated attention to the door, nor the fact that any cue allocates attention to the door, that causes infants to expect an object. Instead, it is the communicative referential aspect of the attention-directing cue: Infants at 12 months expect the communicator to refer to something, so when there is nothing, they are surprised.

Simply highlighting an area does not excite this same expectation. Other studies have compared index-finger pointing cues to non-pointing cues with a similar surface structure (Behne et al., 2005; Bertenthal et al., 2014; Krehm et al., 2014) and find that those control cues do not direct infants' attention to objects. The current experiment built on that work. Instead of using a control cue that does not direct attention to a target area, we employed a cue that did direct attention to the target area. Given our sensitive physiological measure of visual attention, it was important to test whether infants, when being successfully cued to attend to an area, would derive similar expectations about a referent.

Experiment 2, especially in light of results from Experiment 1, reveals that referential expectations are specific to a communicative referential cue, but not to other, non-social, non-communicative attention-directing cues. These findings thus exclude domain-general attention effects to account for the main finding of Experiment 1. It may be an important future asset to dissect the specific characteristics of non-linguistic referential communication, like the visual attention-directing component, vocalizations, prosody, or emotions, and to determine which ones are necessary or sufficient. First, however, in order to better understand the development of communication-induced object expectations we found in Experiment 1, we tested in Experiment 3 a younger age group of 8-month-olds with the same design of Experiment 1 and compared its results to Experiment 1.

Experiment 3

Experiment 3 was similar to Experiment 1, except that we tested 8-months-old infants instead of 12-month-olds. If results for the 8-month-olds were to mirror those of the 12-months-olds, it would provide strong evidence that an understanding of the referential nature of communicative pointing gestures appears already before infants can point themselves; and before they reliably show point comprehension in behavioral settings. Such evidence would make accounts less likely which assume a protracted development of referential understanding in the second year of life due to caregivers' shaping of infant behavior (Carpendale & Carpendale, 2010) and instead support views on early cognition as basis for the emergence of interactional behaviors (Csibra, 2003).

Method

Participants. Seventeen 8-months-old infants (8 males, 9 females) were included in the final sample. Mean age was 8 months, 18 days (range: 8 months, 4 days – 9 months, 5 days). Nine additional infants participated (6 males, 3 females) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (8) or missing data in one or more conditions during baseline or test (1). On average, each infant provided data for a median of 8 (range 5 - 11) out of 12 trials. Apparatus, Stimuli and Procedure were identical to Experiment 1.

Data processing. As in the previous experiments, we analyzed gaze patterns to control for the influence of looking behavior on the pupil diameter.

Results

Preliminary analyses

Figure 8 displays the mean pupil diameters for the four conditions across a trial. As in Experiment 1, visual inspection revealed a difference in pupil size right after the manipulation, already before the outcome. The pointing conditions led to smaller pupils during the baseline than the no point conditions (M_{point}= 3.52 mm, SD_{point}= .51 mm, $M_{\text{no poin''}}$= 3.69 mm, $SD_{\text{no poin''}}$= .52 mm, $t_{(16)}$ = 6.13, p < .001).

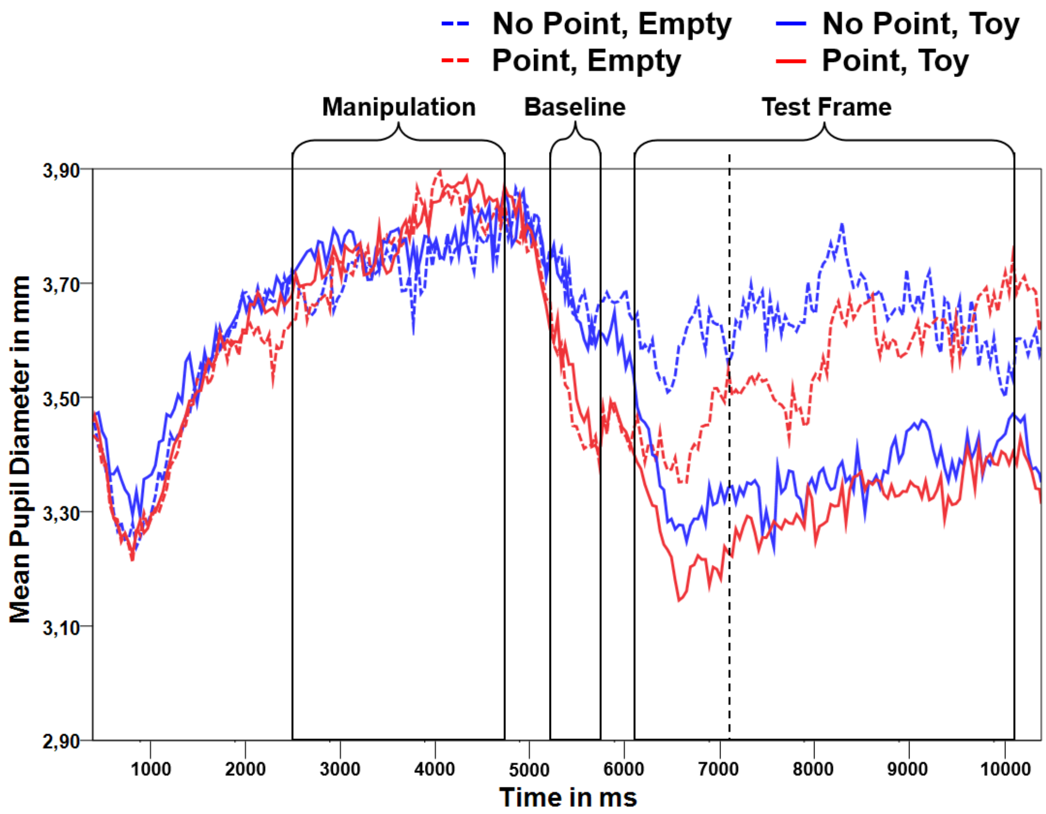
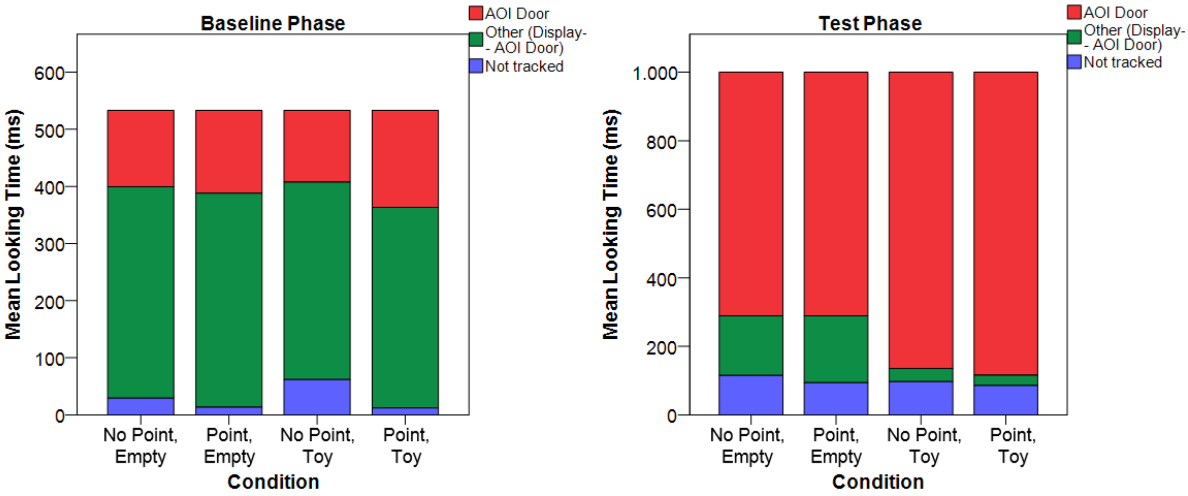


Figure 8. Mean pupil diameter of 8-month-olds in Experiment 3. Windows of analyses begin 500 ms later compared to the timeline in Fig. 1.

Our looking time control analysis revealed that in point trials, infants looked significantly longer at the screen than in the no point trials (see Fig. 9a, M_{point} = 480.97 ms, $SD_{\text{point''}}$ = 68.76 ms, $M_{\text{no poin''}}$= 404.05 ms, $SD_{\text{no point}}$= 102.71 ms, $t_{(16)}$ = 3.28, p = .005). We suspect that when infants looked away they focused on the darker test booth, which may explain why the pupils in the no point trials were larger than in the point trials.



(a) Baseline phase. *(b)* 1000 ms of test phase.

Figure 9. Distribution of looking time in Experiment 3.

Main analyses

To test our main question, we applied the same baseline correction as in Experiment 1. A repeated measures t-test could not confirm the referential expectation hypothesis and revealed that the relative changes in pupil size in the empty conditions were not significantly different after a point compared to after no point ($M_{point, empty}$= -.037, $SD_{point, empty}$= .038, $M_{no\ point, empty}$= -.049, $SD_{no\ point, empty}$ = .037, $t_{(16)}$ = 1.31, p = .207, see Fig. 10). Pupil size in the toy conditions was significantly less decreased after a point compared to after no point ($M_{point, toy}$= -.071, $SD_{point, toy}$= .062, $M_{no\ point, toy}$=-.105, $SD_{no\ point, toy}$=.052, $t_{(16)}$ = 2.84, p = .012).

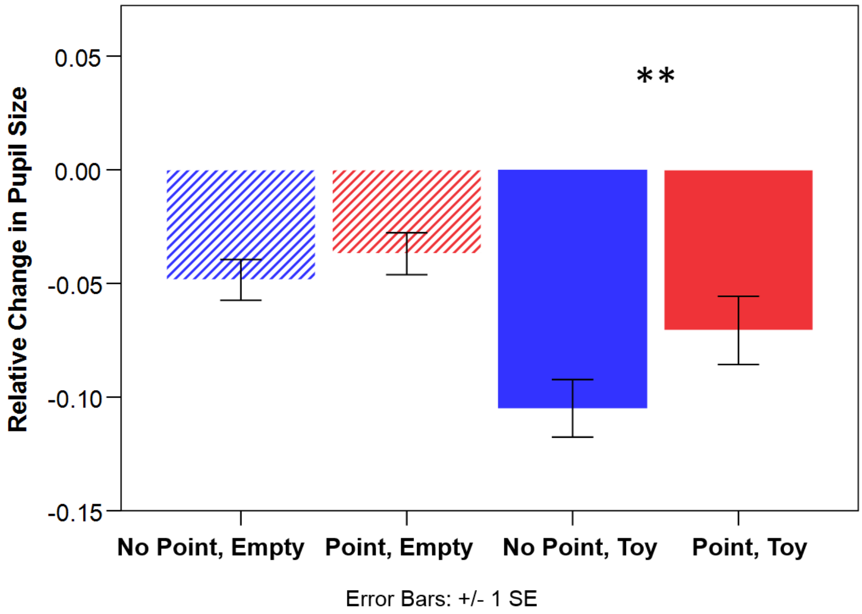


Figure 10. Mean Change in Pupil Size in Experiment 3.

Control analyses

To confirm the results from the pupil data, we again analyzed looking times to different AOIs during the test phase. Paired comparisons between conditions revealed no significant differences between any conditions in terms of looking time to the video as a whole or looks away from the screen. A 2x2 repeated measures ANOVA with manipulation and presence of toy as within-subject factors revealed that infants looked more towards the door when there was a toy than when it was empty ($F_{(1, 16)} = 5.98$, $p = .026$, $\eta_p^2 = .272$) but no main effect for manipulation and no interaction between manipulation and presence of toy, thus excluding that the pupil data were an artefact of difference in looking patterns (see Fig. 9b).

Additional analyses

We repeated the analysis for the test phase of 4000 ms to check for a later emergence of the effect. Repeated measures *t*-Test following our hypotheses revealed that relative change in pupil size in the empty conditions were not significantly different after a point compared to after no point ($M_{point,\ empty}$= -.009, $SD_{point,\ empty}$= .044, $M_{no\ point,\ empty}$= -.023, $SD_{no\ point,\ empty}$ = .042, $t_{(16)} = 1.06$, $p = .304$). Relative change size in the toy conditions was significantly less decreased after a point compared to after no point ($M_{point,\ toy}$= -.059, $SD_{point,\ toy}$= .045, $M_{no\ point,\ toy}$=-.096, $SD_{no\ point,\ toy}$=.052, $t_{(16)} = 3.44$, $p = .003$).

Infants looked longer to the door area over the duration of 4000 ms when there was a toy than when it was empty (main effect for presence of toy: $F_{(1,\ 16)} = 23.81$, $p < .001$, η_p^2 = .598), however, regarding our crucial comparisons, the analysis revealed no significant differences between any conditions of the same outcome in terms of looking time to the door frame; to the video in total; or looking away from the screen.

As in Experiment 1, we performed also the omnibus repeated measures ANOVAs in the 1000 ms and the 4000 ms time windows to provide a more general view on the data. For both time windows there were main effects for Cue (respectively, $F(_{16}) = 8.34$, $p = .011$; $F_{(16)} = 8.06$, $p = .012$) and Outcome (respectively, $F_{(16)} = 23.72$, $p < .001$; $F_{(16)} = 32.87$; p < .001), but no interactions between the two conditions (respectively, $F_{(16)} = 2.68$, p = .122; $F_{(16)} = 2.14$, $p = .168$).

Bonferroni-corrected post hoc paired comparisons (for 6 comparisons) between all conditions for the first 1000 ms of the test phase showed that two comparisons were significant: ‘no point, empty’ and ‘point, empty’ were both significantly larger than ‘no point, toy’ ($t_{(16)} = 4.89$, $p < .001$ and $t_{(16)} = 5.67$, $p < .001$; respectively), again likely reflecting the main effect of outcome. No other comparisons yielded significant differences. For the 4000 ms time window, three comparisons became significant: as for the 1000ms time window, both ‘no point, empty’ ($t_{(16)}$ = 4.58, p < .001) and ‘point, empty’ ($t_{(16)}$ = .7.07, p < .001) were significantly larger than ‘no point, toy’. Further, ‘point, empty’ was significantly larger than ‘point, toy’ ($t_{(16)}$ = .5.05, p < .001). Again, these results only reflected the main effect of outcome (larger pupils for ‘empty’ than for ‘toy’ outcomes).

Comparison between Experiment 1 and Experiment 3

The purpose of Experiment 1 and Experiment 3 was to test for the presence or absence of a referential expectation effect in each of the two different age groups, rather than to compare its size across ages. However, because the stimuli were the same we can also provide the fuller picture by reporting the results of the ANOVA with Cue and Outcome as within-subject factors and age as a between-subject factor. As expected we found a main effect for Cue ($F_{(1,32)}$ = 11.83, p = .002, η_p^2= .270) and Outcome ($F_{(1,32)}$ = 28.12, p < .001, η_p^2= .468), further a tentative main effect for age ($F_{(1,32)}$ = 4.090 p = .052, η_p^2= .113), and the expected three-way interaction between Cue, Outcome and age ($F_{(1,32)}$ = 4.55, p = .041, η_p^2= .124) confirming the findings of the individual Experiments, and no other interactions. Similarly, a direct t-test comparison between the Experiments on the mean difference between the ‘point, empty’ and ‘no point, empty’ outcomes based on relative difference scores revealed a significantly greater pupil increase in Experiment 1 than Experiment 3 ($t_{(32)}$ = 2.23, p = .033).

Discussion

Eight-month-olds showed no difference in relative change of pupil size from baseline between the expectation-violating conditions and the control conditions. In contrast to the 12-month-olds of Experiment 1, 8-month-olds did not differentiate between the ‘point, empty’ condition and the ‘no point, empty’ condition. Importantly, across our focal comparison conditions, the 8-month-olds watched the videos for the same amount of time, and paid attention to the same areas of interest, thus excluding perceptual differences as one explanation. By the same standards of Experiment 1, the finding thus suggests that pointing does not induce object expectations in 8-month-old infants. Note also that the difference in pupil sizes following the pointing cue (before the outcome) is more parsimoniously explained by our control analysis on looking times indicating perceptual differences.

The current finding of a developmental difference in the emergence of communication-induced object expectations contradicts previous findings which have suggested similar performance at 8 and 12 months of age (Csibra & Volein, 2008). One should note, however, that the latter study remained amenable to a leaner attentional highlighting interpretation (longer looking to the empty side because attention has been directed to that side). Further, it did not establish the effect for each age group separately. Of course it could be possible that other ostensive-referential cues would elicit referential expectations at 8 months, and the current study can only speak to the quotidian canonical pointing cue we employed, which is likely the most frequent gesture infants see at that age (Salomo & Liszkowski, 2013). Our current finding of age differences is less supportive of views on which reference comprehension emerges early as a unitary complex system (Csibra, 2003) and instead provides room for developmental accounts on which several factors may interact to yield an increasing refinement in the development of understanding gestural reference.

General Discussion

The current study investigated the complexity and development of infant nonlinguistic reference comprehension across three experiments. The main question was whether the observation of a nonlinguistic communicative referential pointing act led infants to expect a referent object before they could see it. The alternative was that infants simply follow the vectorial direction of a point without prior expectation of encountering an object. Findings from Experiment 1 rejected the lean alternative for 12-month-old infants and revealed that referential communication with a pointing gesture induces the expectation of an object, prior to encountering it. Findings from the control Experiment 2 revealed that the effect of induced object expectation was not due to being cued to attend to the location where an object would appear, thus excluding attention-highlighting interpretations and general attention-directing processes as one alternative explanation of the effect. Results from the younger age group of Experiment 3 revealed that the nonlinguistic communicative pointing cue did not induce object expectations at 8 months, an age at which infants themselves do not yet produce the canonical index-finger pointing gesture, suggesting a protracted development of reference comprehension across the first year of life. The current study thus provides new evidence for a referential interpretation of infant pointing at 12 months of age, while at the same time it suggests that leaner alternative interpretations apply earlier in the first year of life.

The importance of the finding of a developmental difference between 12 and 8 months of age cannot be overstated. Current findings strongly suggest that point comprehension does not emerge as one parcel early in ontogeny but instead likely undergoes developmental change: Very early point-following in the first half of the first year appears rather reflexive (Bertenthal et al., 2014). In the early part of the second half of the first year point-following responses are governed only by the directional nature of the gesture, but not yet by referential expectations. While points may then be understood as being connected to the visible environment to which they direct attention (Woodward & Guajardo, 2002), it is only toward the end of the first year of life, as the current findings attest, that gestural reference induces referential expectations about unseen objects. Given this developmental gradient, it is plausible that a cognitively even more advanced distinction between intended referent and indicated site, e.g. based on false belief attribution, still awaits further development within the second or third year of life (Liszkowski, 2018). Here, we can only speculate about the developmental mechanisms of change between 8 and 12 months of age. While infants certainly gain more experiences with processing pointing during this period (see Salomo & Liszkowski, 2013), another major developmental change pertains to the emergence of point production in that period (see Liszkowski & Tomasello, 2011, Behne et al., 2012). And so, while the current study demonstrates that by 12 months of age infants partake in a 'meeting of minds' which forms a basis for further cultural learning and understanding others, it remains to be tested to what extent this 'meeting of minds' is perhaps as much a product of socialization and earlier forms of social interactions in the first year of life.

References

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*, 492–499. https://doi.org/10.1111/j.1467-7687.2005.00440.x

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

Bertenthal, B. I., Boyer, T. W., & Harding, S. (2014). When do infants begin to follow a point? *Developmental Psychology*, *50*, 2036–2048. https://doi.org/10.1037/a0037152

Bradshaw, J. L. (1968). Pupillary changes and reaction time with varied stimulus uncertainty. *Psychonomic Science*, *13*, 69–70.

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, *45*, 1322–1331. https://doi.org/10.3758/s13428-013-0327-0

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67–78.

Bruner, J. S. (1995). From joint attention to the meeting of minds: An introduction. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 1–14). Hillsdale, NJ: Erlbaum.

Butler, S. C., Caron, A. J., & Brooks, R. (2009). Infant Understanding of the Referential Nature of Looking. *Journal of Cognition and Development*, *1*, 359–377. https://doi.org/10.1207/S15327647JCD0104_01

Butterworth, G. E. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9–33). Mahwah, N.J: Erlbaum. Retrieved from http://www.worldcat.org/isbn/9780805840148

Butterworth, G. E., & Cochran, E. (1980). Towards a Mechanism of Joint Visual Attention in Human Infancy. *International Journal of Behavioral Development*, *3*, 253–272. https://doi.org/10.1177/016502548000300303

Butterworth, G. E., & Grover, L. (1988). The origins of referential communication in human infancy. In L. Weiskrantz (Ed.), *Oxford science publications. Thought without language.* Oxford [England], New York: Clarendon Press; Oxford University Press.

Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *108*, 307–317. https://doi.org/10.1037/0735-7036.108.4.307

Carpendale, J. I. M., & Carpendale, A. B. (2010). The Development of Pointing: From Personal Directedness to Interpersonal Direction. *Human Development*, *53*, 110–126. https://doi.org/10.1159/000315168

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*, 447–458. https://doi.org/10.1098/rstb.2002.1235

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*, 148–153. https://doi.org/10.1016/j.tics.2009.01.005

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

Dörrenberg, S., Rakoczy, H., & Liszkowski, U. (2018). How (not) to measure infant Theory of Mind: Testing the replicability and validity of four non-verbal measures. *Cognitive Development.* Advance online publication. https://doi.org/10.1016/j.cogdev.2018.01.001

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

Gredebäck, G., & Melinder, A. M. D. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition*, *114*, 197–206. https://doi.org/10.1016/j.cognition.2009.09.004

Gredebäck, G., Melinder, A. M. D., & Daum, M. M. (2010). The development and neural basis of pointing comprehension. *Social Neuroscience*, *5*, 441–450. https://doi.org/10.1080/17470910903523327

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, *23*, 967–972. https://doi.org/10.1177/0956797612440571

Hepach, R., & Westermann, G. (2016). Pupillometry in Infancy Research. *Journal of Cognition and Development*, *17*, 359–377. https://doi.org/10.1080/15248372.2015.1135801

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*, 1190–1192.

Hochmann, J.-R., & Papeo, L. (2014). The invariance problem in infancy: a pupillometry study. *Psychological Science*, *25*, 2038–2046. https://doi.org/10.1177/0956797614547918

Itakura, S. (2004). Gaze-following and joint visual attention in nonhuman animals. *Japanese Psychological Research*, *46*, 216–226. https://doi.org/10.1111/j.1468-5584.2004.00253.x

Jackson, I. R., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental Science*, *12*, 670–679. https://doi.org/10.1111/j.1467-7687.2008.00805.x

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings. Biological Sciences / the Royal Society*, *282*. https://doi.org/10.1098/rspb.2015.1683

Krehm, M., Onishi, K. H., & Vouloumanos, A. (2014). I See Your Point: Infants Under 12 Months Understand That Pointing Is Communicative. *Journal of Cognition and Development*, *15*, 527–538. https://doi.org/10.1080/15248372.2012.736112

Leung, E. H., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology*, *17*, 215–220. https://doi.org/10.1037/0012-1649.17.2.215

Liszkowski, U. (2018). Emergence of shared reference and shared minds in infancy. *Current Opinion in Psychology*, *23*, 26–29. https://doi.org/10.1016/j.copsyc.2017.11.003

Liszkowski, U., & Tomasello, M. (2011). Individual differences in social, cognitive, and morphological aspects of infant pointing. *Cognitive Development*, *26*, 16–29. https://doi.org/10.1016/j.cogdev.2010.10.001

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*, F1-F9. https://doi.org/10.1111/j.1467-7687.2004.00315.x

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

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*, 269–274.

Murphy, M., & Messer, D. J. (1977). Mothers, infants and pointing: A study of a gesture. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 325–354). London: Academic Press.

Paulus, M. (2011). How infants relate looker and object: evidence for a perceptual learning account of gaze following in infancy. *Developmental Science*, *14*, 1301–1310. https://doi.org/10.1111/j.1467-7687.2011.01076.x

Range, F., & Virányi, Z. (2011). Development of Gaze Following Abilities in Wolves (*Canis Lupus*). *PLoS ONE*, *6*, e16888 EP -. https://doi.org/10.1371/journal.pone.0016888

Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, *15*, 426–435. https://doi.org/10.1111/j.1467-7687.2012.01139.x

Salomo, D., & Liszkowski, U. (2013). Sociocultural settings influence the emergence of prelinguistic deictic gestures. *Child Development*, *84*, 1296–1307. https://doi.org/10.1111/cdev.12026

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*, 265–266. https://doi.org/10.1038/253265a0

Sebastián-Gallés, N. (2013). Eyes wide shut: linking brain and pupil in bilingual and monolingual toddlers. *Trends in Cognitive Sciences*, *17*, 197–198. https://doi.org/10.1016/j.tics.2013.03.002

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

Sirois, S., & Jackson, I. R. (2012). Pupil Dilation and Object Permanence in Infants. *Infancy*, *17*, 61–78. https://doi.org/10.1111/j.1532-7078.2011.00096.x

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*, 907–912. https://doi.org/10.1111/j.1467-7687.2009.00946.x

Tomasello, M., & Call, J. (2008). Assessing the validity of ape-human comparisons: a reply to Boesch (2007). *Journal of Comparative Psychology*, *122*, 449–452. https://doi.org/10.1037/0735-7036.122.4.449

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*, 705–722. https://doi.org/10.1111/j.1467-8624.2007.01025.x

Ullwer, D., Ries, J., Foth, M., & Meer, E. van der. (2010). Resource allocation and learning in individuals differing in cognitive abilities: Insights from pupillometry and eye - movement analysis. *EyeTrackConf.*

Verschoor, S. A., Paulus, M., Spapé, M. M., Bíró, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64–78. https://doi.org/10.1016/j.cognition.2015.01.005

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*, 1061–1084. https://doi.org/10.1016/S0885-2014(02)00074-6