Three-year-olds understand communicative intentions without language, gestures, or gaze

Richard Moore[1], Kristin Liebal[1,2] & Michael Tomasello[1]
[1]Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany / [2]Park-Krankenhaus Leipzig, Klinik für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie

The communicative interactions of very young children almost always involve language (based on conventions), gesture (based on bodily deixis or iconicity) and directed gaze. In this study, ninety-six children (3;0 years) were asked to determine the location of a hidden toy by understanding a communicative act that contained none of these familiar means. A light-and-sound mechanism placed behind the hiding place and illuminated by a centrally placed switch was used to indicate the location of the toy. After a communicative training session, an experimenter pressed the switch either deliberately or accidentally, and with or without ostension (in the form of eye contact and child-directed speech). In no condition did she orient towards the hiding place. When the switch was pressed intentionally, children used the light-and-sound cue to find the toy – and tended to do so even in the absence of ostensive eye contact. When the experimenter pressed the switch accidentally, children searched randomly – demonstrating that they were tracking her communicative intent, and not merely choosing on the basis of salience. The absence of an effect of ostension contradicts research that ostension helps children to interpret the communicative intentions underlying unfamiliar signs. We explain this by concluding that while it may play a role in establishing a communicative interaction, it is not necessary for sustaining one; and that even with a highly novel communicative act – involving none of the means of communication on which children typically rely – three-year-olds can comprehend the communicative intentions behind an intentionally produced act.

Keywords: communicative intentions; deixis; ostensive cues; three-year-olds

Communication is a pervasive feature of human social life. Understanding that some actions are produced with communicative intent, and grasping the intentions with which they are produced, therefore plays a foundational role in many interactions hypothesised as central to cognitive development. These include

Interaction Studies 14:1 (2013), **62–80**. DOI 10.1075/is.14.1.05moo
ISSN 1572–0373/E-ISSN 1572–0381

language acquisition (Bloom 2000; Sabbagh & Baldwin 2005; Tomasello 1999, 2008; Csibra & Gergely 2009), imitative learning (Gergely & Csibra 2005), pretend-play (Rakoczy, Tomasello & Striano 2005), and perhaps even joint attention (Carpenter & Liebal 2012).

Because words are used primarily for communicative purposes, in many cases the utterance of them will suffice to indicate when a speaker is communicating. However, in non-verbal interactions, interlocutors must rely on non-verbal cues to infer a speaker's communicative intent. These may be the same cues that pre-verbal children use in their language acquisition. A number of studies (Tomasello, Call & Gluckman 1997; Behne, Carpenter & Tomasello 2005; Senju & Csibra 2008; Leekam, Solomon & Teoh 2010) have used simple tasks to illuminate our knowledge of children's understanding of when – and what – others are communicating non-verbally. Two versions of these studies have often been pursued simultaneously: those that investigate (1) the sorts of signs (construed as intentionally produced vehicles of a speaker's message) from which children can infer an experimenter's communicative goals; and (2) those that investigate the conditions under which those signs will be interpreted as communicative.

Tomasello, Call and Gluckman (1997) investigated (1) by testing children's ability to find a hidden sticker when an experimenter indicated its location using three different signs: a point, a marker placed on the hiding container, and a replica of the hiding container held up in front of the experimenter and shown to the participant. They found that children of 30 and 36 months were successful using all these cues. However, they found the marker condition more difficult than the pointing one, and the replica condition more difficult still. This suggests that children start off able to infer a speaker's communicative goals from a limited repertoire of non-verbal actions – understanding simple deictic gestures like pointing (typically accompanied by gaze alternation between the interlocutor and the target of the point) before they understand more abstract non-deictic signs like replicas. By 'deictic' cues, what we mean is cues that involved the communicator either looking, moving or extending her body towards the hiding place, in such a way that features like proximity, physical contact and interest might give clues as to the location of the toy.

Taking young children's proficiency with pointing and gaze-following as a starting point, Behne, Carpenter and Tomasello (2005) investigated (2) – the circumstances in which children interpret familiar actions as communicative. Within a hiding game, an adult indicated the location of a hidden sticker by pointing or gazing towards the container in which the toy had been hidden. She did this in either a communicative or a non-communicative condition; where in the former but not the latter, the pointing and gazing actions were produced purposively and in conjunction with cues (including sustained ostensive eye contact)

associated with communication. Children from 14 months used both the gaze and the point to find the toy – but only in the communicative condition, suggesting that it is not simply the presence of a sign that children use, but also additional markers of communicative intent. These results are consistent with those of Senju and Csibra (2008), who found that six-month-olds followed an adult's 'referential' gaze toward an object only when it was preceded by ostensive eye contact or infant-directed speech. However, the cues used by Behne et al. differed along a number of dimensions between conditions, making the precise contributions of each difficult to establish. While in their communicative pointing condition, the experimenter extended her index finger in the direction of the hiding place, made ostensive eye contact and alternated gaze between the child and location, in the non-communicative control condition she held her hand in a point-like but limp gesture and inspected her wrist in a distracted manner, without gaze alternation or sustained ostensive eye contact. Consequently, this study did not differentiate between a number of features each of which might be necessary or sufficient for an action to be interpreted as communicative: the intentionality of the action, the presence of ostensive eye contact, and the role of gaze-alternation.

A recent series of studies (Leekam, Solomon & Teoh 2010) sought to disambiguate the contributions of some of these cues and shed further light on the varieties of signs that children comprehend by using a series of increasingly abstract signs (a point, an arrow, and a shown replica of the hiding place) to indicate the location of a hidden toy. They also varied whether or not the experimenter's face was covered, and her facial expression; the latter could be either neutral, or warm and engaging. They found that children of two and three years found the point and the arrow easier to interpret than the replica, and additionally that an adult's positive and engaging facial expression helped both younger and older age groups to interpret the experimenter's sign. That was particularly true when this sign was unfamiliar (for example, a replica as opposed to a point). Since in this study the experimenter's engaging facial expression included ostensive eye contact, in addition to a warm and friendly expression, this data can be interpreted as supporting evidence that ostensive eye contact facilitates the interpretation of communicative intent.

In our study we aimed to differentiate between the features of an action that would lead children to interpret it as communicative. Adapting the method of Behne, Carpenter and Tomasello (2005), an experimenter indicated to the child the location of a hidden toy. The intentions with which the experimenter did this were varied along two lines: whether her performance of the action was intentional, or accidental; and whether or not it was performed ostensively. An intentional action was operationalised as one performed deliberately. Following Senju and Csibra (2008), we used two markers of ostension: ostensive gaze, and directed speech.

We used a novel cue – a light-and-sound mechanism – to indicate the hiding location of the toy. This allowed us to (1) test children's understanding of others' actions as communicative independent from familiarity with the cue (they could not assume that the cue was communicative on the basis of prior experience with it), and (2) to use an identical cue across all four conditions. Additionally, the experimenter never looked towards the location of the toy, since the tendency to follow automatically the gaze of others to a referent (Hood, Willen & Driver 1998) could lead to children identifying its hiding place independently of any grasp of communicative intent (Freire, Eskritt & Lee 2004). We predicted that children would interpret the adult's behaviour as communicative when she illuminated the box intentionally and ostensively, but not when the box was illuminated accidentally. Although it is not currently known whether young children understand that non-deictic acts like the pressing of light switches can be used to perform referential acts, because the cue was unfamiliar, not iconic, and did not involve the experimenter extending her body in the direction of the illuminated box, we expected it to be difficult for them to interpret.[1]

1. Method

1.1 Participants

Ninety-six children of 36 months (48 girls, 48 boys; mean age = 35;29, range = 34;6–38;1) were recruited from a database of parents who had volunteered their children to take part in child development studies, in a predominantly Caucasian medium-sized German city. Nine additional children were tested but excluded from the final sample because of experimenter error (n = 1), parental interference (n = 1), inattentiveness (n = 1), and failure to complete the test (n = 6).

1.2 Materials and design

A light-and-sound mechanism consisting of two wooden boxes (20 × 26 × 12 cm, each covered with an orange, translucent Perspex sheet and containing an alarm and a blue flashing light) attached to a pair of large light switches mounted vertically on a central control panel (11 × 6 × 20 cm; with a sheet of clear Perspex to prevent child interference) was placed upon a testing table (160 × 38 × 42 cm). The light boxes were placed 120 cm apart, with the switches half way between (see Figure 1). The toys were hidden inside a pair of yellow cardboard containers (15 × 12 × 12 cm) with top-mounted lids. Various toys were used for hiding, including small wooden cubes and plastic farmyard animals. After retrieval, children were encouraged to slide the toys into a 'music box' (a box with a receiving

tube and xylophone keys fitted to its base), to produce a rewarding sound. In the training phase a yellow arrow (24 cm × 10 cm) and a wooden pointing stick (35 cm long) were also used.

Figure 1. The experimental set-up (ostensive/intentional condition). The toy is hidden in the illuminated box to the experimenter's left. Children could not see where the experimenter had hidden the toy

In a between-subjects design children received four trials in one of four conditions, varying the ostension and intentionality of the action: ostensive/intentional (boys n = 6; girls n = 6), non-ostensive/intentional (boys n = 6; girls n = 6); ostensive/accidental (boys n = 6; girls n = 6); and non-ostensive/accidental (boys n = 6; girls n = 6). In half the trials the toy was hidden in the left and in half in the right container, in fully counterbalanced order. Because of availability issues, two different female experimenters tested children. Each tested six boys and six girls in each condition, except intentional/non-ostensive (E1 7 girls, 6 boys; E2 5 girls, 6 boys).

1.3 Procedure

Children were tested in a quiet room. After a familiarisation phase during which the experimenter and the child played for five minutes with unrelated toys, the introduction phase started when the experimenter said "I know a game we can play. Let me show you something!". She led the child to a chair placed centrally and 110 cm in front of the testing table and removed a linen sheet covering the light-and-sound mechanism. The two light-boxes faced the child. Sitting behind the switches, the experimenter explained how the mechanism worked. In turn she pressed each switch for three seconds and said: "When I press this switch,

this happens," while alternating her gaze between the child and the illuminated/beeping box. The child was then invited to try the mechanism herself, before the experimenter temporarily stored the light switches under the table. Finally, the experimenter demonstrated how the music box worked and encouraged the child to use it herself. This phase served to demonstrate the equipment to the child, prior to and independently from the test-phase.

As an entirely novel cue was used in the main procedure, next a training phase was administered at the same table. The goals of this training phase were two-fold: to introduce the child to the hiding game, and to encourage her to think about the possibility of using a variety of different signs for the same communicative goal. While the novel test cues did not involve the experimenter extending her body towards or looking at the hiding place to indicate a referent (we call this 'bodily deixis'), during the training we used somewhat familiar cues that required the experimenter to extend her arm towards one box (although without looking at it), therefore providing more spatial information about the toy's location. (We anticipated that this would make the training phase easier for young children, and so ease them into the more difficult test phase.)

In the first part of the training phase, the experimenter called the child's attention, showed her the toy, and placed it in one of the containers in view of the child – before asking the child to retrieve the toy and place it into the music box. She did this once per side, in random order. This phase was used to familiarise the child with the process of retrieving the toy and placing it into the music box.

In subsequent trials the experimenter called the child's attention and then hid the toy, as the child watched, by concealing it in her hands and placing it discreetly into one box (random order, at most two consecutive trials per side). In conjunction with ostensive cues (eye contact and a smiling, engaged facial expression), the experimenter then used one of three different signs to indicate its location (in fixed order): (1) an ipsilateral finger point, (2) an ipsilateral stick point, and (3) an ipsilateral arrow held vertically above the hiding place. Each sign was performed for three seconds, after which the experimenter paused and asked "Do you know where [it] is?" and encouraged the child to retrieve the toy and place it into the music box. If the child searched for the toy in the wrong location, the experimenter retrieved the toy herself, held it up, and stated in a reassuring tone "Oh, it was here! Let's try again!", before hiding the toy again.

Each child received each sign until she performed correctly in two consecutive trials, or until the maximum of six trials per sign was administered. Then the experimenter moved to the next sign, or the main procedure, respectively. She signalled the end of the training phase by clapping her hands together and announcing "That was a fun game!". A second experimenter then entered the room, left a book and exited[2], before the first experimenter said in a neutral way "I have an idea".

The experimenter replaced the light switches on the table and – using the same concealment method employed in the training phase – hid the toy in one of the yellow containers, which were placed in front of the light-boxes. She then pressed the switch that would illuminate the toy's hiding place in one of the four test conditions, leaving the mechanism turned on for three seconds.

In the ostensive conditions, E engaged the child's eye contact, smiled, and held her finger in the air Eureka-style, then said "And now …" in an engaging way before pressing the switch. While the flashing alarm was on, E maintained eye contact and positive affect with the child. In the non-ostensive conditions, E did not engage the child's eye contact after hiding the toy, but looked down at the switches as if inspecting them and said "And now …" as if speaking to herself, before pressing the switch. While the light and sound were on, E did not make eye contact with the child but kept looking at the switches. In all conditions, E left her left hand resting on the switches throughout the trial. In the intentional conditions, she raised her right hand and then pressed down on the switch with her index finger in a deliberate and confident manner. In the accidental conditions she shifted her weight, as if to make herself comfortable, and then 'accidentally' pressed the appropriate switch with her left hand. She then sat back, startled, threw her hands in the air in mock surprise and said "Oops!".

In all conditions, after switching the mechanism off, E paused, made eye contact with the child and asked playfully "Do you know where [it] is?".

1.4 Scoring and data analyses

Children's responses were coded from videotape. When a child chose the box containing the toy, this was counted as correct. Children were counted as choosing a box if they opened it (and so could see the object), picked it up (and so could feel the weight of the object), or if they indicated to E that they would like her to open the box – for example, by pointing to it. They did not count as choosing a box just by touching it – so if they touched one box and then opened the other, the latter action counted as their choice. Children could also make 'no choice' about the location of the hidden toy (2 trials in total –0.48%) by failing to choose within 30 seconds. This time limit was used in order to exclude trials where children became distracted and subsequently chose their search location on some basis other than the experimenter's action.

Additionally, children's attentiveness during the trial was coded. To be considered attentive, a child needed to be quiet and attending to the experimenter while the switch was pressed. Since trials were not initiated until children were attentive, scoring here was very high: over 384 trials, children were scored as attentive in 380 (99%), and only inattentive in 1. Three cases were marked as unclear.

To assess reliability, 25% of all trials (96 in total, from 3 children per condition per experimenter) were coded from tape by a second observer unaware of the experimental hypotheses and ignorant of the particular conditions. There was 100% agreement between coders about which containers children chose (κ = 1), and 99% agreement on children's attentiveness (κ = .99).

For each child, the sum successful retrievals after four trials was calculated.

2. Results

2.1 Comparisons to chance

In each condition, children's ability to locate the toy was compared to chance. Since children were tested in four trials each, with two possible hiding locations, finding the toy twice would constitute chance performance.

Children retrieved the toy well above chance in the ostensive/intentional condition (one sample t-test: mean retrieval = 2.75, t(23) = 3.30, p = .003), and with marginal significance in the non-ostensive/intentional condition (mean = 2.5, t(23) = 2.02, p = .056). In both accidental conditions children found the toy only at chance (ostensive/accidental: mean = 2.04, t(23) = 0.20, p = .85; non-ostensive/accidental: mean = 2.12, t(23) = 0.59, p = .56; see Figure (2).

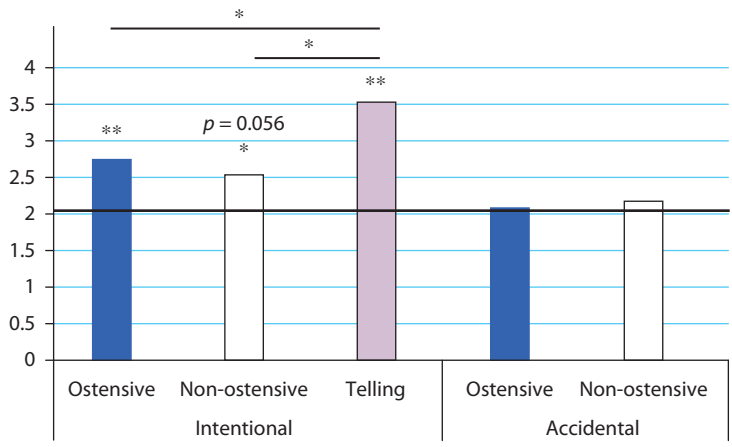

Figure 2. Children's mean success at finding the hidden toy (over four trials), including the follow-up 'telling' condition. The horizontal bar indicates chance performance. In intentional conditions E pressed the switch that illuminated the light and sound mechanism deliberately; and in accidental conditions she pressed the switch by falling on it 'accidentally'. In ostensive conditions, E pressed the switch in conjunction with child-directed speech and ostensive eye contact. In non-ostensive conditions E inspected the switch before pressing it, did not make ostensive eye contact with the child, and spoke only to hersel

2.2 Preliminary analyses

We also used an ANOVA to measure the proportion of variance explained by each variable in the experimental setup.

There was no main effect of experimenter ($F(1, 96) = 2.33$, MSE = 1.07, $p = .17$), sex (F(1, 96) = 0.04, MSE1.07, $p = .85$) or hiding order (F(1, 96) = 3.14, MSE = 1.07, $p = .084$) on children's ability to locate the hidden toy. Consequently the data was collapsed across these variables.

2.3 Main analyses

A two-way ANOVA testing the effects of intentionality and ostension on children's retrieval of the toy found a main effect of intentionality ($F(1, 92) = 5.32$, MSE = 1.22, $p = .023$), but not ostension (F(1, 92) = 0.077, MSE = 1.22, $p = .78$) and no interaction effect (F(1, 92) = 0.42, MSE = 1.22, $p = .52$). The effect size of intentionality was small (Cohen's $d = 0.47$, $r = 0.23$).

2.4 Additional analyses

Additionally, in order to explore the performance of children at the individual level, we plotted the number of trials on which individuals found the toy (their 'sum success') across each of the four conditions. In both accidental conditions there was a unimodal distribution of data, with the mode performance at chance (ostensive/accidental, n = 11; non-ostensive/accidental, n = 11). In contrast, in the intentional conditions there was a bimodal distribution of data. Although children found the toy on average above chance, in fact only around 1/3 of children (ostensive/intentional, n = 9; non-ostensive/intentional, n = 7) performed at ceiling (retrievals = 4/4 trials) – and similar numbers of children performed at chance (retrievals = 2: ostensive, n = 10; non-ostensive, n = 8). This supports the prediction that children found interpreting this cue difficult.

Further evidence that children found the test cue comparatively difficult to interpret comes from the training phase, which 67 children (70%) completed in only six trials (ceiling performance). Only ten children (10%) made 2 or more mistakes in the finger-point training, compared to six (6%) in the stick-point training, and four (4%) in the arrow training.

2.5 Discussion

Children in the ostensive/intentional condition, at least considered as a group, clearly understood the experimenter's use of the novel sign to indicate the location of the toy. Additionally, children in the non-ostensive/intentional condition did so marginally. In both accidental conditions children found the toy only at

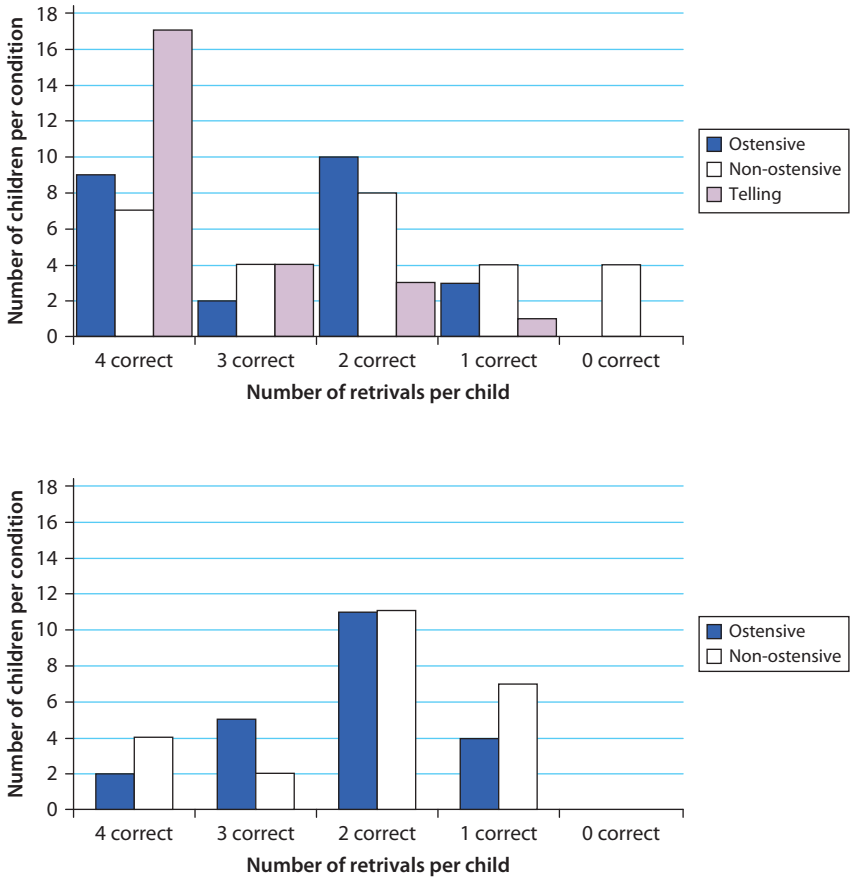

Figure 3. Children's summed success at finding the hidden toy (over four trials). The top figure shows performance in intentional conditions, and includes the follow-up 'telling' condition; the bottom figure shows accidental conditions

chance – indicating that they did not use the light mechanism alone as a cue to the toy's location. Successful performance cannot, then, be explained solely by appeal to the greater salience of one box relative to the other, and supports the hypothesis that in the intentional conditions, children took the experimenter's actions to be revelatory of communicative intent. It might be objected here that children's performance in the accidental conditions could be explained by their greater attention to the experimenter, because of her surprised response in the accidental action. This distraction would make the illuminated box comparatively less salient in accidental conditions. While this is possible, we think it highly unlikely that it contributed to children's different performance across conditions. The light and sound mechanism was both visually salient and very noisy. In all conditions, the

illumination continued for three seconds – considerably longer than it took for the experimenter to express her surprise. It's therefore unlikely that children simply failed to notice which box was illuminated, even in accidental conditions.

The lack of a main effect of ostension or of an interaction effect, and the marginal success of children in the non-ostensive/intentional condition, suggest that at least in this task the experimenter's production of ostensive cues played no significant role in children's recognition of her communicative intent. Despite the appearance of a difference between the ostensive/intentional and non-ostensive/intentional conditions (children in the former performed significantly above chance; in the latter, marginally below significance) the absence of an effect of ostension seems unlikely to be attributable to a lack of power, given then very low p-value shown by the ANOVA ($p = .78$). This contradicts the finding of Leekam et al. (2010), who found that a positive, engaging face facilitated the comprehension of unfamiliar signs.

A likely explanation of this difference can be found in the warm-ups used in the respective studies. In their 'pre-test' phase, the participants of Leekam et al. (2010) played a game in which they had to find hidden stickers, but where the experimenter did not seek to provide any sign that would indicate their location. By contrast, in our training phase the experimenter used a variety of signs to show the toy's location, so that the test phase took place against a background of cooperative and communicative interactions between child and caregiver. Consequently, even in the absence of ostensive cues, it may be that children took the experimenter's actions to be a continuation of the hiding game they had previously played; and so interpreted her intentional actions as communicative. This suggests that the contribution of ostensive cues to interpretation may be contextual: where the communicative nature of an interaction is already common ground (Clark 1996; Liebal et al. 2009) between participants, this may be sufficient to indicate to an audience that an individual's actions are communicative, making ostensive cues – at least in the form of directed speech and ostensive eye contact – inessential to at least some communication. For example, while such cues may be necessary for establishing a communicative interaction, they are not necessary for sustaining it.

The bimodal distribution of children's performance in the intentional conditions suggests that even in the ostensive/intentional condition, many of the children tested were guessing the location of the toy. Consistent with our predictions, this shows that three-year-olds found this task difficult. An interesting question is why children struggled with the task. One low-level possibility – that children didn't understand the light-and-sound mechanism – is unlikely because they had both been taught about and played with it previously.

A more likely explanation is that three year olds found this task difficult because it required them to infer not only the experimenter's communicative goal (her message), but to do so from her performance of an unfamiliar

action (the vehicle of her message); and because they failed to anticipate that this particular action could be used to perform a referential communicative act. Previous studies (including Tomasello, Call & Gluckman 1997; Leekam, Solomon & Teoh 2010) have shown that while children readily understand the use of deictic, directional markers like pointing and referential gaze to refer to a location, the use of non-deictic markers in this way is not intuitive for them. This effect may be amplified where unfamiliar signs are used.

An explanation of why a minority of children nonetheless performed well, at least in the intentional conditions, may be that successful children had somehow acquired the insight that the non-deictic cues could be used referentially. With this insight in hand, interpreting the experimenter's action as indicating the hiding place of the toy becomes relatively straightforward. However, to those who lack this insight, the task would remain difficult.

We ran a follow-up condition designed to test the hypothesis that children performed poorly because they failed to understand that our light mechanism could, even in the absence of deictic cues, be used to refer to the location of the hidden toy.

3. Follow-up 'Telling' condition

3.1 Method

3.1.1 *Participants*

Twenty-five children of 36 months (12 girls, 13 boys; mean age = 35;21, range = 34;1–37;26) were recruited from a database of children, in the same German city as the previous conditions. The same experimenters as were used previously each tested half of the children (E1 n = 6 girls, 6 boys; E2 n = 6 girls, 7 boys).

3.1.2 *Materials*

The materials were the same as those used in the previous conditions.

3.1.3 *Procedure*

The procedure in the 'telling' condition was, with one difference, the same as in the ostensive/intentional condition previously tested. In this condition, after hiding the toy, E made eye contact with the participant, smiled, lifted her finger and said "And now I'm going to tell you where it is," in an engaging tone of voice. She then pressed the light switch in a deliberate manner, maintaining eye contact and positive affect with the child for the three seconds during which the light and sound mechanism was activated.

3.1.4 *Scoring and data analysis*

Scoring and data were carried out as in the previous conditions. Over 100 trials, children were judged to be attentive in 99 (99%), and in all trials children made a choice about the location of the toy (100%).

Additionally, reliability coding was carried out on 24% of cases (three children per experimenter), by a coder unaware of the experimental hypotheses and ignorant of the condition being tested. There was 100% agreement between coders about which containers children chose ($\kappa = 1$), and 98% agreement on children's attentiveness ($\kappa = 0.96$).

4. Results

4.1 Comparison to chance

Children retrieved the toy well above chance in the telling condition (one sample t-test: mean retrieval = 3.48, $t(24) = 8.49$, $p < .001$).

4.2 Preliminary and main ANOVA analyses

Additionally, data from the 'telling' condition were compared with those from the intentional conditions previously tested. A preliminary analysis was conducted using a four-way ANOVA, including ostension, sex, experimenter, and hiding order as factors. This analysis showed no main effect of experimenter or sex. There was a main effect of hiding order ($F(5, 73) = 8.62$, MSE = 0.50, $p = .029$) – showing that children in some counterbalancing permutations found the toy more often than in others. However, since there was no interaction between ostension and counterbalancing ($F(10, 73) = 2.17$, MSE = 1.09, $p = .24$), and since all participant groups were fully counterbalanced, we collapsed across this variable in further analyses.

Subsequently, treating 'telling' as a variety of ostension, a one-way ANOVA was used to test the effects of ostension on children's retrieval of the toy. A main effect was found ($F(2, 73) = 5.66$, MSE = 6.57, $p = .005$). In addition, post-hoc pairwise tests performed using a Bonferroni correction showed that children were significantly more likely to retrieve the toy in the telling condition than in the ostensive condition (MSE = 1.16, $p = .043$; effect size: Cohen's $d = 0.73$, $r = 0.34$). They were also significantly more likely to find the toy in the telling condition than the non-ostensive condition (MSE = 1.16, $p = .007$; Cohen's $d = 0.93$, $r = 0.42$). However, they were not more likely to retrieve the toy in the ostensive condition than in the non-ostensive one (MSE = 1.16, $p > .999$).

4.3 Additional analyses

As before, we also plotted children's 'sum success' in the new condition. In contrast to the ostensive and non-ostensive conditions, in the telling condition there was a unimodal distribution of data, with 17 children performing at ceiling (retrievals = 4/4), and only two children performing at chance (retrievals = 2/4) – suggesting that children found this condition easy. This interpretation is supported by the fact that in the training phase of the study children performed similarly to those previously tested. 15/25 (60%) completed the warm-up in only six trials (ceiling performance). Only one (4%) made 2 or more mistakes in the finger-point training, compared to three (12%) in the stick-point training, and none (0%) in the arrow training.

4.4 Discussion

Children tested in the telling condition were significantly more likely to find the hidden toy than in both ostensive/intentional and non-ostensive/intentional conditions. A likely explanation of this finding is that young children do not find it intuitive to interpret unfamiliar non-deictic markers referentially – although a minority are capable of doing so. When the experimenter told the children that her action – pressing the light switch – would indicate the location of the toy, they found it easy to infer where the toy was hidden. This confirms the hypothesis that children's difficulty in the previous conditions lay not in inferring communicative intentions in general, but in understanding that the light mechanism could be used to refer to the location of the hidden toy.

4.5 General discussion

The fact that children used the light mechanism to retrieve the hidden toy only in intentional conditions shows that they tracked speakers' communicative intentions, and did not simply respond to salience when looking for the toy. This shows that children know that communicative acts must be intentionally produced. However, the difficulty that children had in the originally tested intentional conditions suggests that three-year-olds nonetheless exploit a number of everyday bodily cues in interpreting communicative intent; and that in the absence of these cues, inferring even simple messages can become difficult for them. For example, children's ability to interpret a speaker's action as a referential communicative act may depend not only upon its being performed intentionally and serving to make a location salient, but also upon the speaker's pattern of gaze, the orientation of her body towards the referent, and the hearer's familiarity with the vehicle of meaning by which the referent is picked out. As part of their facility for inferring referential

intent, it may be that young children initially expect that only a limited variety of actions – for example, acts of speech, gaze, and deictic gesture – can be used to refer. This would not be surprising: after all, these actions are overwhelmingly the ones by which speakers do refer in the course of their every day interactions. The insight that any action could, in principle, be used to refer might then be acquired only later – perhaps as children become more familiar with the wide range of actions that speakers perform with referential intent, where these are typically produced in conjunction with acts of ostension and bodily deixis that scaffold their comprehension.

Although children clearly performed better when told of the experimenter's intention to use the 'light mechanism' referentially, as yet it remains unclear where breakdown in comprehension occurred for those children who struggled to understand the speaker's message in the originally tested intentional conditions. Two possibilities are worth considering here, particularly because they bear directly on recent debates on the role of ostension in children's cognition of communicative intent.

First, it may be that in the originally tested intentional conditions children recognised *that* the speaker was acting with communicative intent, but could not infer *what* she was telling them. This might happen if children expected that, were the experimenter trying to tell them the location of the toy, then she would have pointed to its hiding place – since this would be the most natural action to perform in this context. Since the experimenter did not point, children may have assumed that she wanted to tell them something other than the location of the toy – such that her action was not a reliable indicator of the toy's location. For example, they may have taken her to be explaining how the light mechanism worked. Their assumption that E was not telling them the location of the toy would gain additional support if they also failed to recognise her pressing the light switches as a potentially referring act.

While this explanation is possible and consistent with the above data, its plausibility is somewhat undermined by children's experience of the experimenter playfully indicating the toy's location in a variety of ways; and by the experimenter's having previously demonstrated the mechanism.

An alternative explanation – the 'that' explanation (to be contrasted with the 'what' explanation above) – locates the source of children's breakdown at an even earlier stage of comprehension. On this story, it may be that some of the tested children failed to realise that, when the experimenter pressed the light switch, she was doing so with the intention to communicate. This would be the case if children failed to recognise that light switches could be used communicatively – let alone to communicate referentially. It may be that children initially expect that only certain kinds of action – for example, words and gestures – are

communicative; and that they learn only later that any action could, in principle, be used as a vehicle of meaning.

The latter possibility would present a puzzle for some accounts of the role of ostension in the recognition of communicative intent. In a series of recent papers, Gergely and Csibra have developed the hypothesis that humans possess an adaptive 'Pedagogy' mechanism that evolved to facilitate social learning (Gergely & Csibra 2005; Gergely, Egyed & Király 2007; Csibra & Gergely 2009; Csibra, 2010). On this view, children are hardwired to interpret a variety of ostensive cues – including ostensive eye contact, infant directed speech ('motherese'), and contingent turn-taking behaviour – as indicating an individual's acting with communicative intent. Against the Pedagogy view, on the second interpretation children in our study would have failed to grasp that the experimenter acted with communicative intent, in spite of her production of ostensive cues.

In order to deepen our understanding of the cognitive processes by which children infer communicative intent, future research should try to shed led on precisely where children's performance in our study broke down. Current evidence is consistent with both of the explanations proffered above. Not least, this is because the telling condition that we tested added two pieces of information not found in the ostensive/intentional conditions. First it informed the child *that* the experimenter was acting with communicative intent; and second, it specified some of *what* that content of the communicative act would be. Either piece of information could suffice to explain children's improved performance in the telling condition.

One possibility is that the primary function of ostensive cues may not be to indicate communicative intent, but simply be to highlight an individual's performance of an action, and to indicate to an interlocutor that an utterance has been performed *for her* (Tomasello 2008; Moore, submitted; in press). Ostensive cues may be unnecessary to indicate this, since this property of an action can be inferred from the context of its production. Additionally, ostensive cues may be insufficient to indicate communicative intent, because this additionally requires that the speaker is performing an action intentionally, to solicit some response from the hearer, and its being common ground that the action that a speaker performs could be used as the vehicle for a communicative act.

One valuable line of future research would be to investigate the ways in which the contributions to the understanding of communicative intent of the various cues used in this study vary with contextual features of the interaction between speaker and hearer. Follow-up studies should also try to disambiguate the competing explanations of the role of ostension in communication. In the meantime, both explanations are consistent with the possibility that children

start with a relatively small repertoire of actions from which they can infer a speaker's message.

Acknowledgements

This research was made possible by a DAAD Early Career Research Fellowship to the first author. We would like to thank our experimenters, Susanna Jorek and Maria Ploetner, and our research assistants Manja Teich and Gesa Volland. We would also like to thank Marike Schreiber for the illustration in Figure 1, Georg Keller and Kerstin Esau for reliability coding, and Roger Mundry and Harriet Over for discussion of the manuscript. Finally, we acknowledge with gratitude the parents and children whose participation made this study possible.

Notes

1. This prediction was supported by extensive piloting, which indicated that children of 30 months failed to find the toy, even in ostensive conditions.

2. A second experimenter was used to control for a planned follow-up study in which the experimenter was changed between the training and test-phases, in order to evaluate the role of common ground in the comprehension of unfamiliar signs. This follow-up study has not yet been run.

References

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492–499.

Bloom, P. (2000). *How children learn the meanings of words*. London: MIT.

Carpenter, M., & Liebal, K. (2012). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint attention: New developments* (pp. 159–181). Cambridge, MA: MIT Press.

Clark, H. (1996). *Using language*. Cambridge: Cambridge University Press.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141–168.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Science*, *13*(4), 148–153.

Freire, A., Eskritt, M., & Lee, K. (2004). Are eyes windows to a deceiver's soul? Children's use of another's eye gaze cues in a deceptive situation. *Developmental psychology*, 40(6),1093–1104.

Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, *6*(3), 463–481.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science*, *10*(1), 139–146.

Hood, B.M., Willen, J.D., & Driver, J. (1998). Adults' eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*, 131–134.

Leekam, S.R., Solomon, T.L., & Teoh, Y-S. (2010). Adults' social cues facilitate young children's use of signs and symbols. *Developmental Science*, *13*(1), 108–119.

Liebal, K. (2009). *Infants' and children's understanding of common ground and markedness in communication*. Leipzig: Universtität Leipzig.

Moore, R. Enacting and understanding communicative intent. Manuscript submitted for publication.

Moore, R. (In press). Ontogenetic constraints on Paul Grice's theory of communication. In D. Matthews (Ed.), *Pragmatic development*. Amsterdam: John Benjamins.

Rakoczy, H., Tomasello, M., & Striano, T. (2005). How children turn objects into symbols: A cultural learning account. In L. Namy (Ed.), *Symbol use and symbol representation* (pp. 67–97). New York, NY: Erlbaum.

Sabbagh, M., & Baldwin, D. (2005). Understanding the role of communicative intentions in word learning. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention, communication and other minds* (pp. 165–184). Oxford: Oxford University Press.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668–671.

Tomasello, M. (1999). *The cultural origins of human cognition*. London: Harvard University Press.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, *68*(6), 1067–1080.

Addresses for correspondence

Richard Moore
Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Deutscher Platz 6
D-04103 Leipzig, Germany

r.t.moore@gmail.com

Kristin Liebal
Park-Krankenhaus Leipzig GmbH
Klinik für Kinder- und Jugendpsychiatrie
Psychosomatik und Psychotherapie
Morawitzstraße 2, D-04289
Leipzig, Germany

kristin.liebal@parkkrankenhaus-leipzig.de

Michael Tomasello
MPI-EVA (as above)

tomas@eva.mpg.de

Author's biography

Richard Moore is a post-doctoral researcher in the Department of Developmental and Comparative Psychology at the Max Planck Institute for Evolutionary Anthropology in Leipzig. He holds a Ph.D. in Philosophy (Warwick 2010) and his research focuses on the nature of cognition required for intentional communication.

Kristin Liebal was a post-doctoral researcher in the Department of Developmental and Comparative Psychology at the Max Planck Institute for Evolutionary Anthropology in Leipzig, before retraining as a clinical psychologist. Her research investigates the socio-cognitive abilities of normally developing children and children with autism, particularly in communicative and cooperative interactions.

Michael Tomasello is Director of the Department of Developmental and Comparative Psychology at the Max Planck Institute for Evolutionary Anthropology in Leipzig.