Cognitive Development 28 (2013) 312–329

Contents lists available at SciVerse ScienceDirect

Cognitive Development

Understanding communicative intentions and semiotic vehicles by children and chimpanzees

Jordan Zlatev[a,d,*], Elainie Alenkær Madsen[c,d], Sara Lenninger[b,d], Tomas Persson[c], Susan Sayehli[d,e], Göran Sonesson[b,d], Joost van de Weijer[a,e]

[a] Department of Linguistics, Centre for Languages and Literature, Lund University, Sweden
[b] Department of Semiotics, Centre for Languages and Literature, Lund University, Sweden
[c] Department of Cognitive Science, Lund University, Sweden
[d] Centre for Cognitive Semiotics, Lund University, Sweden
[e] Humanities Lab, Lund University, Sweden

ARTICLE INFO

Keywords:
Conventionality
Familiarity
Iconicity
Indexicality
Object-choice
Pointing

ABSTRACT

Developmental and comparative studies of the ability to understand communicative intentions using object-choice tasks raise questions concerning the semiotic properties of the communicative signals, and the roles of rearing histories, language and familiarity. We adapted a study by Tomasello, Call, and Gluckman (1997), in which a "helper" indicated the location of a hidden reward to children of three ages (18, 24, and 30 months) and to four chimpanzees, by means of one of four cues: Pointing, Marker, Picture and Replica. For the chimpanzees, we controlled for familiarity by using two helpers, one unfamiliar and one highly familiar. Even 18-months performed well on Pointing and Marker, while only the oldest group clearly succeeded with Picture and Replica. Performance did not correlate with scores for the *Swedish Early Communicative Development Inventory* (SECDI). While there were no positive results for the chimpanzees on the group level, and no effect of familiarity, two chimpanzees succeeded on Pointing and Marker. Results support proposals of a species difference in understanding communicative intentions, but also highlight the need to distinguish these from the complexity of semiotic vehicles and to consider both factors.

* Corresponding author at: Department of Linguistics, Centre for Languages and Literature, Lund University, Sweden. Tel.: +46 762981471.

E-mail address: jordan.zlatev@ling.lu.se (J. Zlatev).

0885-2014/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cogdev.2013.05.001

1. Introduction

When do children begin to understand communicative intentions? To what extent are other species capable of doing so? How do different means of communicating such as gaze, pointing gestures, replicas, or pictures, affect understanding? Tomasello, Call, and Gluckman (1997) used an object-choice task to address these questions. Subsequent studies have applied this method to address both developmental (Aureli, Perucchini, & Genco, 2009; Behne, Carpenter, & Tomasello, 2005; Behne, Liszkowski, Carpenter, & Tomasello, 2012) and comparative questions (Call & Tomasello, 2005; Hare & Tomasello, 2004; Hermann, Melis, & Tomasello, 2006; Leekam, Solomon, & Teoh, 2010), without reaching definitive conclusions. This may be because the two questions have not been differentiated from a third, that of semiotic vehicles.

We address all three questions theoretically and empirically by adapting Tomasello et al.'s (1997) study in which a reward was hidden in one of three boxes differing in shape and color. A second researcher communicated the reward location by (a) *pointing* to the correct box, (b) placing a wooden block *marker* on top of it, or (c) holding up a *replica* of the box. Thirty-month-old children found the reward in all cases. In contrast, apes (six chimpanzees and three orangutans) performed above chance only in the marker condition previously trained on. Tomasello (1999a) interpreted this as showing that "the apes were not able to understand that the human beings had intentions toward their own attentional states" (p. 102). Initially, this interpretation was thought to support the claim that apes do not understand intentions in general: ". . . the understanding of conspecifics as intentional beings like the self is a uniquely human cognitive competency" (p. 56). This conclusion, however, was shown to be problematic since a number of studies demonstrated that (at least) chimpanzees do understand others' intentions. Specifically, chimpanzees pursue contested food only if a (dominant) conspecific competitor cannot see it (Hare, Call, Agnetta, & Tomasello, 2000; Hare, Call, & Tomasello, 2001); they conceal their approach when competing against a human experimenter (Hare, Call, & Tomasello, 2006; Melis, Call, & Tomasello, 2006); they discriminate unwillingness from inability to deliver food (Call, Hare, Carpenter, & Tomasello, 2004); and they distinguish states of knowledge from states of ignorance (though not false beliefs) in a competitive game (Kaminski, Call, & Tomasello, 2008).

These capabilities were observed in competitive contexts, i.e. when the chimpanzees were motivated to consider the perspectives, goals and knowledge states of another (conspecific or human) individual, so as to maximize their own profits. Furthermore, even the brief description above makes it clear that it is specifically communicative intentions that are required for task success, a possibility Tomasello (1999b) admitted: ". . . in the current context I will simply maintain that non-human primates, whatever they may or may not understand about the simple intentions of others, clearly do not understand the communicative intentions of others" (p. 72). But what exactly are communicative intentions? The notion stems from Grice (1957), according to whom, to mean something by uttering/performing x is approximately equivalent to (a) intending to produce some effect on another individual and (b) intending this individual to recognize (a). Theorists who have employed the notion (Csibra, 2010; Moore, submitted for publication; Sperber & Wilson, 1995; Strawson, 1964; Zlatev, 2008a) differ in their interpretations, but there is general agreement that communicative intentions imply at least a second-order intention (b) to recognize the primary intention (a).

Tomasello, Carpenter, Call, Behne, and Moll (2005) and Tomasello (2008) offered this explanation of what distinguishes human from ape (social) cognition: (a) a motivation to share (information) and (b) the cognitive capacity for shared intentionality, i.e., engaging in and understanding joint intentions, both simple and communicative. Zlatev, Persson, and Gärdenfors (2005) and Zlatev (2008a) complemented this explanation by noting first that apes have restricted imitation capacity, in particular of bodily actions (Call, 2001; Custance, Whiten, & Bard, 1995; Hribar, Sonesson, & Call, in press; Myowa-Yamakoshi & Matsuzawa, 2000; although see Carrasco, Posada, & Colell, 2009). Second, imitation has been closely linked to empathy, both theoretically and empirically (Hurely & Chater, 2005; Tomasello & Carpenter, 2005), ever since the classical proposal of Lipps (1903). Third, imitation and empathy have been argued to serve as springboards for intentional communication in both child development (Piaget, 1962), and hominid evolution (Donald, 1991). Thus, an adaptation for bodily mimesis, implying improved volitional control of the body, could possibly explain why human being are particularly skillful (compared to non-human primates) in all three domains – imitation, empathy

Table 1
Classifying semiotic vehicles used in object-choice studies according to the factors bodily means, semiotic ground, directionality and representational relationship (see text).

Vehicle	Bodily	Ground	Directionality	Representation
Ostensive gaze	Yes	–	Yes	No
Proximal point	Yes	Indexical (+ symbolic)	Yes	No
Marker	No/Yes	Indexical	No/Yes	No/Yes
Picture	No	Iconic (+ symbolic)	No	Yes
Replica	No	Iconic	No	Yes

and (gestural) intentional communication. Since these are arguably prerequisites for language, no additional adaptations for the evolution of the latter (apart from vocal control) need be assumed (Zlatev, 2008a, 2008b). However, there are important unresolved questions and gaps in the evidence from object-choice studies to be addressed before any of these theoretical proposals can be considered well-supported.

1.1. Semiotic vehicles and communicative intentions

Apart from pointing, markers and replicas, object-choice studies with children and apes have employed a few other non-verbal means to communicate the location of a reward, such as ostensive gaze (Behne et al., 2005) and pictures (Hermann et al., 2006). We refer to these as semiotic vehicles. Various other terms are used in the psychological and ethological literature such as ‘cue’, ‘signal’, ‘symbol’ and ‘sign’, but they have the disadvantage of being used with specific meanings in different disciplines and theories, often in conflicting ways (cf. Heine & Kuteva, 2007, p. 132, on ‘symbol’). Semiotic vehicles differ at least with respect to four factors: (a) means, bodily or not; (b) semiotic ground – the motivating relation between vehicle and referent, which can be indexical (based on contiguity), iconic (similarity-based), or symbolic (convention-based) (Peirce, 1958; Sonesson, 1995); (c) directionality – whether the vehicle constitutes a vector, directing attention toward the referent (Sonesson, 1998), and (d) sign relation – whether the vehicle is understood by a conscious subject as representing a given object, action or event (Sonesson, 1989, 2009; Zlatev, 2009). The five most common vehicles used in object-choice studies may be characterized according to these four factors as shown in Table 1.

Only the first two vehicles (ostensive gaze and point) are entirely bodily, and only the last two (picture and replica) are clearly signs/representations. Marker relies on contiguity as semiotic ground, while picture and replica rely on iconicity. Directionality is involved in pointing, ostensive gaze, as well as in the act of placing the marker. Finally, understanding “symbolic” as conventional, point and picture may be said to share this property since they are both in part conventionalized and highly familiar to children in Western societies. In contrast, marker and replica can be considered as novel, unconventional vehicles, while ostensive gaze, in the case of communicative gaze alternation between addressee and intended target, is a paradigmatic example of a “natural”, non-conventional semiotic vehicle (Grice, 1957).

What is the relation between type of semiotic vehicle and communicative intentions? The two can be considered independent dimensions (Andrén, 2010), in a single communicative act. Any act performed with deliberate expressiveness for the sake of an addressee can be interpreted as intentionally communicative, whether or not it stands for something (Sperber & Wilson, 1995). Yet the sign relation can function without communicative intention, as when a hunter tracks an animal by “reading the signs”, or when a child engages in “symbolic play” with no one else present. Semiotic theories tend to privilege the role of vehicles, while Gricean (and psychological) approaches tend to focus on intentions. The cognitive-semiotic approach that we adopt suggests that both are non-reducible to one another though closely interacting, aspects of meaning. Csibra (2010) has similarly argued that the Gricean communicative intention (my intention for the other to grasp that I want to communicate) and the informative intention (the content) may exist independently of each other and can be expressed by independent signals. For example, ostensive mutual gaze with an addressee can be regarded as “enacting” the second-order communicative intention (Gergely & Csibra, 2006; Gómez, 1994; Moore, submitted for publication).

Thus there is no *prima face* reason to consider use of any of the semiotic vehicles in Table 1 in object-choice tasks, nor in other experimental paradigms, as "more simple" than others. Indeed, studies involving various subsets of these vehicles yield a rather ambiguous picture. Behne et al. (2005) showed that children as young as 14 months could perform the object-choice task when the experimenter pointed to the correct box, gaze-shifting between the box and the addressee, but not when pointing to the box while looking elsewhere. Ostensive gazing alone often led to finding the reward, although 24-month-olds performed better than 14- and 18-month-olds. Behne et al. (2012) showed that even 12-month-olds could solve the task (under a simplified design) when pointing and ostensive cues were combined. As for other semiotic vehicles, while Tomasello et al. (1997) showed similar results for point, marker and replica for 30- and 36-month-olds, children below age two were reported to succeed only given the cues ostenstive gaze and point (Leekam et al., 2010).

Hermann et al. (2006) applied the object-choice task with photographs and replicas to 12 chimpanzees, six orangutans, six gorillas and four bonobos and found that when photos and replicas of the hidden food were shown on top of the baited box, participants performed well. However, when a color photograph depicting the baited box was shown "between the two boxes and so had no *indexical* relationship to the baited container" (p. 126, our emphasis), no participant reached above-chance performance. Persson (2008) reported similar results with four gorillas, one chimpanzee, one orangutan and one bonobo. A possible conclusion is that non-human primates simply do not understand (the referential function of) pictures and replicas, and perhaps that the understanding of fully-fledged iconic signs (Table 1) is a uniquely human capacity.

Other studies speak against this conclusion, however. First, chimpanzees have been suggested to use scale models to search for hidden items after watching an experimenter hide a miniature item in a scale model (Kuhlmeier & Boysen, 2002). The experimental conditions resembled the "hiding game" paradigm developed by DeLoache (2004) and colleagues in which children see pictures and scale models of objects and are then asked to find the actual objects in another room. These studies have suggested that children's understanding of pictures is not stable until 30 months, and that of scale models even later, although these results may also be due to the inherent greater cognitive complexity of DeLoache's hiding-game task. Suddendorf (2003) showed that some understanding of pictures as representations of real-world objects emerges by 24 months.

In sum, theoretical generalizations are hampered by the fact that different developmental and comparative studies have used various semiotic vehicles with different experimental paradigms, with children of different ages. A further complication, to which we now turn, is that age also interacts with language proficiency.

1.2. Language proficiency and "enculturation"

Explanations of success in object-choice tasks based on social-cognitive capacities such as shared intentionality (Tomasello, 2008) and bodily mimesis (Zlatev, 2008a) assume that these capacities precede language, in both hominid evolution and child development. While relevant studies of children at 12 and 14 months have been reported (Behne et al., 2005, 2012), their results cannot be directly compared with those of Tomasello et al.'s (1997) original study, since only pointing overlaps among the semiotic vehicles used.[1] Thus, it is possibly older children's language proficiency that allowed them to comprehend the sign relation of pictures and replicas.

An analogous problem in ape studies concerns not participants' age, but rather their level of "enculturation". Call and Tomasello (1996) distinguished four increasing levels of "the effect of humans on the cognitive development of apes": (a) mother-raised in captivity, (b) nursery-raised, (c) laboratory-raised (with language-training), and (d) raised in a human culture, or perhaps, a shared "pan-human culture" (Segerdahl, Fields, & Savage-Rumbaugh, 2005). Several studies have reported that human

[1] A secondary reason is that the three-box design from the original study was changed for *two* boxes (also in subsequent studies with apes) – without any explicit discussion, apart from motivations such as "this should make choice easier" (Hermann et al., 2006, p. 126). However, this also increases the potential influence of implicit, non-intentional cues (such as posture), and requires an increased number of trials for conclusive statistics. On the other hand, the 2-box design allows the cue/vehicle to be equidistant to the boxes, and thus to control for proximity more easily.

enculturaltion, or socialization (Tomasello & Call, 2004), affects apes' performance on tasks such as object-choice. For example, while poor comprehension of pointing is commonly reported (Call & Tomasello, 2005), "environmental factors, particularly access to a sociolinguistically rich environment, directly influence great apes ability to comprehend declarative signals" (Lyn, Russell, & Hopkins, 2010, p. 360), as it was shown that chimpanzees and bonobos raised in such an environment succeed in an object-choice task. Persson (2008) reported positive results on picture comprehension with the bonobos Kanzi and Panbanisha, who should be at the highest of the Call–Tomasello enculturation levels, strongly suggesting that at least language-trained bonobos understand pictures as fully-fledged signs. Nonetheless, a lower level of enculturation may be sufficient for the latter: Hribar et al. (in press) showed that a nursery-raised chimpanzee trained in do-as-I-do imitation could recognize actions shown in still photos.

Thus, in addition to age, level of language proficiency should be considered for children, as well as level of enculturation/socialization for apes. The latter can be seen as a special case of "rearing history".

1.3. Rearing history and familiarity

Leavens, Hopkins, and Bard (2008) have argued that it is problematic to compare middle-class children and encaged (and often orphaned) chimpanzees and draw conclusions about species differences. Close emotional contact with a caregiver may be essential for developing the intersubjective skills necessary for understanding communicative and cooperative intentions, and pointing in particular. The positive results from (high-level) enculturants on pointing, including one orangutan in the Tomasello et al. (1997) study, can be interpreted as supportive of this possibility. In a very different experimental paradigm for the study of intersubjectivity/empathy, namely contagious yawning, a familiarity-effect has recently been suggested: Both chimpanzees (Campbell & de Waal, 2011) and dogs (Silva, Bessa, & de Soussa, 2012) appear to be more susceptible to yawning contagion when the model is someone with whom they are familiar. There are reasons to expect such an effect in the object-choice task as well. Grasping communicative intention of a "helper" is not a purely cognitive matter. It requires a presumption that he or she is indeed trying to help one find the desired reward. In other words, performance should be enhanced by trust, which is greater for in-group than out-group members (Over & Carpenter, 2012).

1.4. Goals and research questions

We sought to address these issues by adapting the original Tomasello et al. (1997) study with respect to: (i) *Age*: Children of three different age-groups were to be included: 18, 24 and 30 months. Thus, the age of the youngest group was to coincide with the average onset of the "vocabulary spurt" and the oldest group overlapped in age with the younger of the two groups tested in the original study, thus allowing comparison; (ii) *Role of language*: Measures of the children's receptive and productive vocabulary were to be collected using the *Swedish Early Communicative Development Inventory* (SECDI) (Berglund and Eriksson, 2000), allowing us to examine the relationship between these measures and the children's performance on the object-choice task; (iii) *Familiarity*: One of two helpers in the experiment with chimpanzees was an animal caretaker who had had a close relationship with the chimpanzees since infancy or juvenility. Both children and chimpanzees were exposed to four semiotic vehicles – Point, Marker, Replica and Picture – while controlling for gaze alternation between addressee and target as carefully as possible. These modifications allowed us to address the following questions:

1. Does children's and chimpanzees' understanding of Point differ from that of Marker? Both can be considered indexical semiotic vehicles, but there are a number of important differences (Table 1): Point is a familiar vehicle (at least for children) and thus involves a degree of conventionality, while Marker does not. When accompanied by ostensive gaze, Point is clearly an act of intentional communication, while Marker is more ambiguous. While both (proximal) Point and Marker involve spatial contiguity, Point functions by specifying a direction toward the intended referent. Marker may, though need not, be understood in the manner of Picture and Replica as a fully-fledged sign for the location of the reward.

2. Is there evidence of a developmental progression, as suggested by previous research, such that understanding of indexical vehicles precedes that of iconic ones, and, within the iconic category, comprehension of Picture precedes that of Replica? Evidence of developmental precedence for Point would support a positive role for at least one of the factors that distinguish Point from Marker, or a negative (delaying) role of the final factor indicated above (status as sign).
3. Is there a correlation between linguistic skills and comprehension of communicative intentions and semiotic vehicles? The theories of shared intentionality (Tomasello, 2008) and bodily mimesis (Zlatev, 2008b) predict no positive correlations.
4. Do chimpanzees perform better with familiar than unfamiliar helper, and if so, for which vehicles? If Leavens et al. (2008) are correct that a close emotional bond between helper and addressee is essential for intentional (and cooperative) communication, such a difference is expected.
5. Is there a difference between children and chimpanzees in understanding communicative intentions, and if so, for which vehicles and ages? If only older children performed better than chimpanzees and their performance correlated with language proficiency, the supposition would be supported that it is language, rather than either shared intentionality or bodily mimesis, that accounts for the species difference. However, if even younger children perform much better than the chimpanzees, and there is no correlation with language, it is likely the task is tapping cross-species differences of a non-linguistic nature.

While we made efforts to keep the experimental conditions as similar as possible, we agree with Tomasello et al. (1997) that "making cross-species comparisons in experimental settings is always difficult because a key condition is comparability of the setting *from the subject's perspective*" (p. 1069, our emphasis). Hence, we adapted the experimental design for each species.

2. Study of children's object choice

The experiment was designed as a hiding-finding game, following Tomasello et al. (1997). Like Tomasello et al., we considered it pragmatically odd (at least in the study with the chimpanzees) for the same individual to be first hiding a reward ("baiting a box"), and then helping the participant to find it, especially since our intention was to manipulate the degree of "trust" between chimpanzee participants and researcher. Hence, a 'Hider' performed the first role and a 'Helper' the latter. The warm-up and pretest phases were intended to ensure a degree of familiarity between child and Helper. In the key trials, the Helper used one of four semiotic vehicles to indicate where the reward was hidden. When the Helper was satisfied that the child had seen the semiotic vehicle(s), the Hider pushed the board with the boxes toward the child to make a choice.

2.1. Method

2.1.1. Participants

Seventy-two participants (35 girls) were recruited through letters to families living in Sweden. Each three age groups had 24 children: 18 months (9 girls, $M = 18.07$ months, SD = 8 days), 24 months (12 girls, $M = 24.00$ months, SD = 9.69 days), and 30 months (14 girls, $M = 30.01$ months, SD = 8.92 days). Children received a gift for their participation. Parents' completion of the Swedish Early Communicative Development Inventory (SECDI) (Berglund and Eriksson, 2000) indicated no obvious developmental delays. Ten additional participants were excluded – two (one in the youngest and one the middle age group) who failed the pretest, four who lost the motivation to continue during the session (three in the youngest and one in the middle group), and four because parents did not complete the SECDI (two in the youngest, one in the middle and one in the oldest group).

2.1.2. Materials

The apparatus (Fig. 1) was similar, but not identical to the one used by Tomasello et al. (1997). Three opaque wooden boxes, of different shape and color, were aligned on the longer side of a wooden board (52 cm × 36 cm). The top of each box was placed facing downward, and its edge attached to the board, so that it could be opened by tipping it backward. After piloting, it was decided to attach a

Fig. 1. The apparatus used in the object-choice task with children (but without the Plexiglas shield with chimpanzees).

transparent Plexiglas "shield" to the front of the board, with round openings in front of each box, with 8 cm diameter. This constrained the child's hand movements when choosing a box (and thus facilitated coding) and discouraged playing with the boxes.

The leftmost box (from the point of view of the participant) was red, 10 cm × 7 cm × 10 cm. The middle box was black, 7 cm × 14 cm × 7 cm. The rightmost box was white, 14 cm × 7 cm × 7 cm. The distance between the boxes was 4.5 cm. An opaque cardboard blind (41 cm × 70 cm, not shown in Fig. 1) was used to occlude the participant's view of the hiding process. On the side of the blind facing the child, a poster of a popular fictional character *Pippi Longstocking* was pasted, giving children something interesting to look at during the hiding process. Rewards were stickers of animals and colorful objects. When children found one, they could (with help from the parent) paste it on a sheet that could be taken home. With permission of the parents, raisins were used as rewards if this helped maintain the child's motivation.

Two versions of the SECDI (one for children below and one above 16 months; Eriksson & Berglund, 2002) were sent to parents one week before the study. Only the two most extensive measures from each were used: a checklist of 370 common Swedish words (common nouns, verbs, and certain closed-class words) from the first version and a checklist of 710 words from the second.

2.1.3. Design and procedure

Parent, child, and the two female researchers (Hider and Helper) first spent approximately 10 min in a child-friendly reception room. While the Hider explained the procedure to the parent and obtained informed consent, the Helper interacted with the child.

Once in the testing room, child and parent were seated on one side of the table with the apparatus, while the two experimenters were positioned on the other side. The Hider was seated, while the Helper stood to her left, from the child's point of view. Children sat in a child seat, with the parent sitting beside them, facing toward the child and away from the apparatus. The parent was asked to encourage the child and help fasten the reward on the sheet, but otherwise refrain from assisting. If the child was clearly not comfortable with this arrangement, we allowed the parent to sit facing the apparatus and hold the child. In that case, however, we asked parents to close their eyes when the Helper communicated the location of the reward. Location of the reward was randomized across trials, with the constraint that it did not appear in the same box more than two trials in a row.

Warm-up phase. The Hider pushed the apparatus toward the child, showing how the boxes opened and closed and how a sticker could be put under one of them. When the Hider opened a box and showed it to contain a sticker, the child was allowed to retrieve it and told that more could be found in the game that was to begin.

Pretest phase. The Hider said that she was about to "hide it", then placed the cardboard blind in front of the apparatus (from the child's view), and put a sticker in one of the boxes, according to a counterbalanced, pre-arranged schedule. When the blind was removed, the Hider turned to her notes, while the Helper, standing beside the Hider, bent down and opened the correct box, made eye contact with the child, and shifted gaze between reward and child, using "accomplice" facial expressions (raised eye-brows, smiles, nods, etc.). When content that the child had seen the reward, she closed the box and surreptitiously signaled the Helper, who then turned toward the child and pushed the board forward asking "Can you find it?". Verbal interaction was kept to a minimum. The child had to find the sticker in three trials in a row to proceed to the next phase.

Test phase. The test was like the pretest, but instead of opening the box and showing the reward, the Helper used one of four semiotic vehicles: (a) *Point*: a proximal, dynamic, index finger pointing gesture toward the correct box, at an angle of 45°, with the tip of the finger at a distance of 5–10 cm to the box; (b) *Marker*: reaching out and placing a yellow "post-it" note on the top of the box; (c) *Picture*: holding up a color photograph of the correct box in mid-position, approximately 30 cm above the middle box (Fig. 1, with some displacement to the left, due to the position of Helper); (d) *Replica*: holding up an identical replica of the correct box in the same position as Picture. For Point and Marker, the Helper's gaze alternated between the box indicated and the child; for Picture and Replica, between the vehicle, the box and the child.

In all cases except Marker, the Helper continued showing the semiotic vehicle until the child had made a choice. Since the Hider was at the same time extending the tray with the apparatus toward the child, this implied that the Helper was in part leaning over the Hider. The order in which children viewed the four vehicles was counterbalanced across participants. Each child viewed the four vehicles six times, divided in two rounds of three. A novelty in our design was that, after each round of three, an "easy" trial (as in the pretest) was presented, so children would not become discouraged if they had consistently failed. After the first set of 16 trials (4 vehicles × 3 trials + 4 easy) a break was introduced and the child offered a snack. Each trial took about 20 s, and a total session, including break, lasted about 40 min. Children's box choice was operationalized as the gesture through which they either touched or indicated one of the three boxes, and this was immediately coded by the 'Hider'. In addition, all sessions were recorded by two cameras, and 20% of the data were coded by two different raters independently. Inter-rater was 96.67%, yielding a Kappa value of 0.94.

2.1.4. Predictions

It was predicted (a) that semiotic vehicles that do *not* require to be understood as (fully-fledged) signs (Point and Marker) would lead to better performance at all ages and would be comprehended earlier than the vehicles that presuppose understanding a representational relationship (Picture and Replica). Also predicted were (b) an advantage for Picture compared to Replica, to the extent that the results from other experimental paradigms can be extended to the object-choice task; (c) an advantage for Point over Marker, given Point was more familiar (and conventionalized) than Marker; (d) no correlations between SECDI scores and performance if the latter were tapping into essentially non-linguistic skills.

2.2. Results and discussion

Mean scores by age and vehicles type appear in Fig. 2. We tested whether these deviated significantly from chance level (=2.00, indicated by the dashed line) for the age groups and vehicles separately (one-sample t-test). Mean scores for the youngest group were above chance for Point = 3.88, $t(23) = 5.48$, $p < .001$, and for Marker = 3.63, $t(23) = 4.71$, $p < .001$, but not for Picture = 2.04, $t(23) = .21$, $p = .833$, nor Replica = 2.17, $t(23) = 1.28$, $p = .213$. Mean scores of the middle group were all above chance: Point = 5.29, $t(23) = 13.90$, $p < .001$, for Marker = 4.71, $t(23) = 8.74$, $p < .001$, for Picture = 2.46, $t(23) = 3.82$, $p = .001$, and for Replica = 2.75, $t(23) = 3.56$, $p = .002$. For the oldest group mean scores were all significantly above chance: Point = 5.50, $t(23) = 18.39$, $p < .001$, Marker = 4.96, $t(23) = 9.74$, $p < .001$, Picture = 4.04, $t(23) = 6.59$, $p < .001$, and Replica = 3.38, $t(23) = 3.67$, $p = .001$.

Numbers of correct responses were compared in a repeated measures analysis of variance (ANOVA) with four factors: age-group, sex, and presentation-order as between-subjects factors and vehicle-type as a within-subjects factor. Main effects of vehicle-type, $F(3,195) = 68.42$, $p < .001$, partial eta-squared = .513 and age, $F(2,65) = 16.95$, $p < .001$, partial eta-squared = .343, were significant. Unlike in the Tomasello et al. (1997) study, the interaction of vehicle type with age group was also significant,

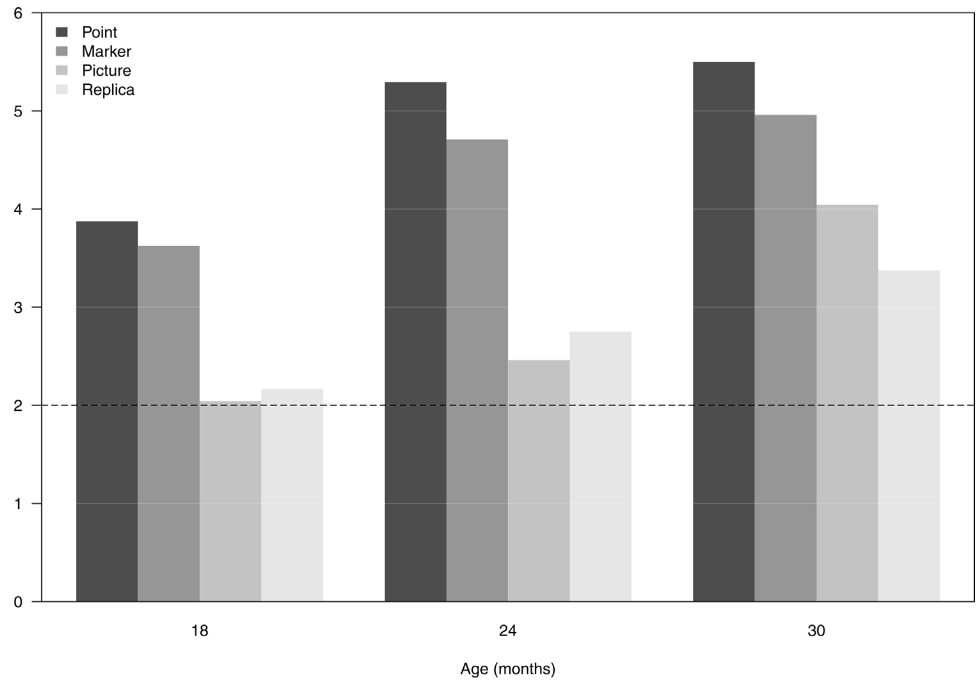

Fig. 2. Mean scores (0–6) by vehicles and age group.

$F(6,195) = 2.34$, $p = .034$, partial eta-squared = .067. The interactions of presentation-order by vehicle-type, and sex by vehicle-type were not significant.

The significant interaction between age and vehicle-type was followed by an analysis of simple main effects. For each age group, significant differences between vehicle types were found only between Picture and Replica and between Point and Marker, although for 30-month-olds the difference between Marker and Picture approached significance ($p = .056$), as did the difference between Picture and Replica ($p = .059$). Thus, Point and Marker were "easier" to understand than Picture and Replica for all age groups, with some indication that Picture was easier than Replica for the 30-month-olds.

We then compared the three age groups for each vehicle type. For Point and Marker, the difference between 18- and 24-month-olds was significant, but not the difference between 24- and 30-month-olds. For Picture and Replica, the difference between 18-months-olds and 24-month-olds was not significant. For Replica the difference between 24- and 30-month-olds was not significant, but for Picture it was. For all vehicle types differences between 18- and 30-month-old groups were significant. Thus, the understanding of Point and Marker appeared to develop earlier than Picture and Replica. Hence, in both respects (rate of correct performance, and order of development) the first prediction was confirmed, while there was only tentative support for the second prediction (Picture > Replica) at 30 months.

These conclusions are supported by using a second criterion, the percentage of children who performed correctly on at least five of the six trials for each vehicle (binominal test, $p < .05$).[2] As Fig. 3 shows, a large proportion of the 18-month-olds and a clear majority of 24- and 30-month-olds succeeded with the indexical non-representational vehicles. With the iconic representational vehicles,

[2] If a child was just guessing, the chance of selecting the correct alternative equals 1/3. Since there were six trials per child, the probability of guessing five or more correct of six equals 0.0178 (binomial distribution). A binomial distribution usually implies that the trials are independent, which is not the case here, given that the trials come from the same child. Still, this is a standard way of gauging individual children's performance in the literature, and following this practice allows straightforward comparison.

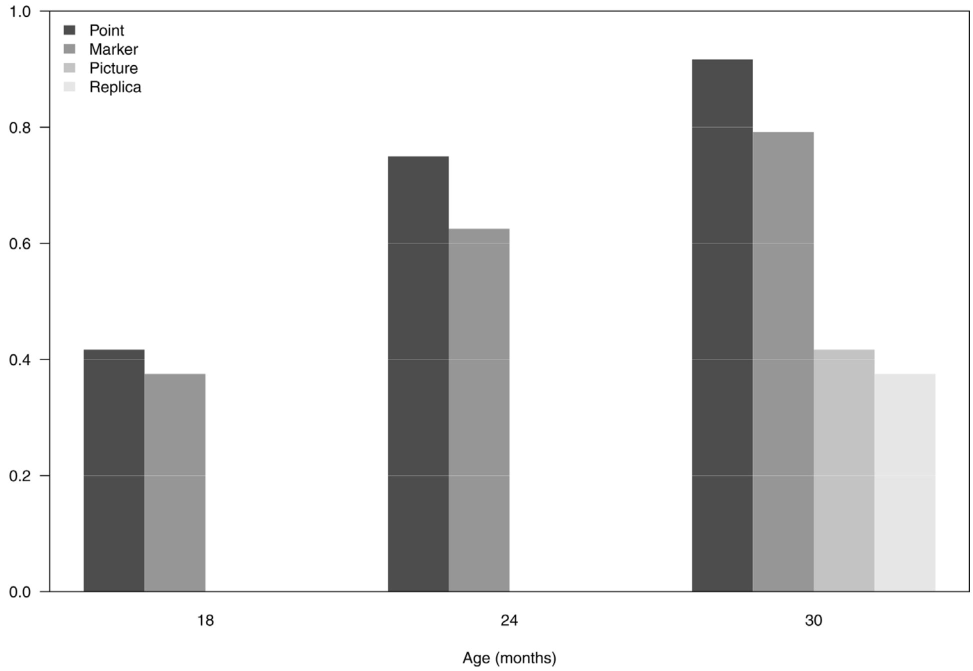

Fig. 3. Proportions of children who performed significantly above chance (at least five of six trials per vehicle type) by age group.

however, not a single child performed significantly above chance-level within the two younger age groups. Then, seemingly abruptly, about 40% of 30-month-olds (a nearly identical proportion as that for the indexical vehicles for the 18-month-olds) succeeded with the representational vehicles above chance level.[3]

Concerning the third prediction, that Point would be at an advantage to Marker, Figs. 2 and 3 both suggest such a tendency. However, as noted above, the differences between the two vehicle types were not significant, neither within nor between age groups. We considered the possibility that the success rates for Marker might have been inflated. In the pilot phase we initially used a wooden block as Marker, but switched to a "post-it" sticker, since children often tried to grab the block, even after we had introduced the Plexiglas shield. It remains a possibility that in some of the Marker trials the children were actually reaching for the "post-it" note, and when failing to do so (due to the shield), adapting their hand movements and touching the box. To control for this possibility, two independent coders assessed the marker condition trials from 20% of randomly selected children in each age group in terms of whether the child reached for the box or for something else (this was usually the Marker, but the coding categories were defined in this way to include a number of ambivalent reactions). Inter-rater agreement was 92% (Cohen's Kappa = 0.81). In 28% (Rater 1) or 26% (Rater 2) of the cases the child reached for something else than the box. Thus, in over one quarter of the trials, we cannot know for certain whether the child treated the Marker as a semiotic vehicle, as opposed to as an object of inherent interest. However, since a corresponding analysis for the other vehicle types has not been performed, we cannot draw the conclusion that the Marker condition was special in this respect. Even if it was, in a clear majority of the double-coded cases the children selected the box.

[3] Interestingly, this performance was considerably better than the Replica results for 30-month-olds in the original Tomasello et al. (1997) study, where the mean score was 2.29, and only a single child passed at least 5 trials. The only explanation for this difference we can suggest is that the Helper in our study may have used "triadic" gaze alternation (i.e., between vehicle, box and child) more systematically.

Finally, concerning the prediction of lack of correlation between the scores for (productive and receptive) vocabulary and object choice, multiple regression analyses were performed with the two SECDI measures, Word Comprehension (0–370) and Word Production (0–710), and age group as predictors. Age significantly predicted performance for Point, Picture and Replica (although not Marker, due to four extreme outliers among 30-month-olds). Most important, neither children's word comprehension nor production scores their understanding of the semiotic vehicle, confirming our prediction.

We should nonetheless exercise caution before concluding that there was no relation between language development and performance on the object-choice task, since we relied only on children's parents to assess vocabulary levels. However, boxplots of comprehension and production scores for each age group showed no severe departures from normality, apart from a few outlying observations. Moreover, mean Production scores (18-month = 35.50, 24-month = 305.50, 30-month = 505) fell within the normal range for Swedish children (Eriksson & Berglund, 2002).

In sum, the most robust finding is the clear difference between the indexical, non-representational vehicles Point and Marker, on the one hand, and the iconic, representational Picture and Replica on the other. We can view the results represented in Figs. 2 and 3 as evidence for a developmental "spurt" between 18 and 24 months for understanding Point and Marker in acts of intentional communication, and for an analogous spurt for Picture (and to some extent Replica) comprehension between 24 and 30 months. Still, even the 30-month-olds appeared to have difficulties with the latter.

3. Study of chimpanzees' object choice

The design and procedure in a study of four chimpanzees were nearly identical to those used in the previously described study with children. The major difference, apart from the much lower number of participants, is inclusion of the independent variable *familiar/unfamiliar Helper*.

3.1. Method

3.1.1. Participants

The participants were three adult and one juvenile chimpanzees (*Pan troglodytes*): Santino (male, zoo-born in Münich, Germany in 1978, hand-reared), Penny (female, zoo-born in Borås, Sweden in 1980, mother-reared), Linda (female, born ca. 1984, wild caught in Liberia, mixed rearing) and Manda (female, zoo-born in Kolmården, Sweden in 2004, mixed rearing). They were housed at Furuvik Zoo/Lund University Primate Research Station Furuvik (LUPRSF). At the time of testing (2010), the participants had experienced few previous tests, and none requiring the understanding of communicative intentions, although one participant (Linda) had been included in a test of understanding pictorial vehicles (Persson, 2008). The animals were not food deprived during the study, and water was available ad libitum.

A main reason for conducting the study with these participants was their intimate contact with a female animal caretaker. One of the participants (Manda) had been raised by her at home for a year and a half, but with daily interaction with the chimpanzee group. During this time the caretaker did not make any special attempts to provide "human enculturation", since the intention was that Manda be accepted by the group after the nursery period in a human environment. The caretaker has been a permanent figure in the lives of all the other chimpanzees as well and is at present trusted by them to the extent that it is easy for her to enter their enclosures and interact with them.

3.1.2. Materials

The same apparatus as that for the child study was used (Fig. 1), but without the plastic shield. Participants were tested while sitting in their enclosures and could only protrude a finger through the bars and not reach for the boxes. Participants sat on the floor and the apparatus board was presented to them on a raised surface at their chest/chin level. Since the chimpanzees were more mobile than the children, a larger cardboard blind was used than the one for the children, shielding from view also from the sides and the top. Pieces of fruit, shelled peanuts, meatballs and candy were used as rewards, depending on the participants' preferences.

3.1.3. Procedure

Each participant received the four sets of six trials per vehicle twice (once for the familiar, and once for the unfamiliar Helper), yielding a total of 48 trials (12 trials per vehicle type). The Helper was either the highly familiar caretaker or an unfamiliar female researcher (the same one that functioned as Helper in the children study), while the Hider was always a familiar male researcher. The first session with each individual started with a warm-up phase, in which the animal was acquainted with the apparatus and how to select a box when it was baited in full view. The order familiar/unfamiliar Helper was counterbalanced only in part; Linda and Penny had the order Fam > Unfam, while Santino had the order Unfam > Fam. Due to her poor performance and low attention span in the pretest, Manda had to be given the order Fam > Unfam for half of the stimuli, followed by Fam > Unfam for the second half.

Pretest phase. As with the children, the Hider first hid the reward in one of the boxes, while holding the blind in front of the apparatus. After the blind had been removed the Helper showed the location of the reward by opening the correct box, while the Hider looked away, by turning approximately 90° to the left and looking down at the protocol papers. Three correct consecutive trials were the criterion for continuing to the experimental conditions. Additional trials were sometimes presented if it appeared the participant had not fully understood the procedure.

Test phase. Trials were conducted in the same manner as for the children, with the Helper seeking the participant's attention by calling their name and then indicating the location of the reward by performing one of the four vehicles, while alternating gaze between box and participant (and vehicle, for Picture and Replica). The four vehicle types were presented in sets of three trials, followed by an "easy" trial in which the Helper showed the correct hiding place by opening the box. Presentation order was not counterbalanced because of the small number of participants. We used the order Point–Picture–Marker–Replica for all participants, intermixing indexical/directional and iconic vehicles to minimize the risk of transfer effects, beginning with Point (the condition expected to be easiest) to give participants an easier start and foster motivation. Breaks were taken when the participant's motivation appeared low or when the routines of the zoo interfered with testing.

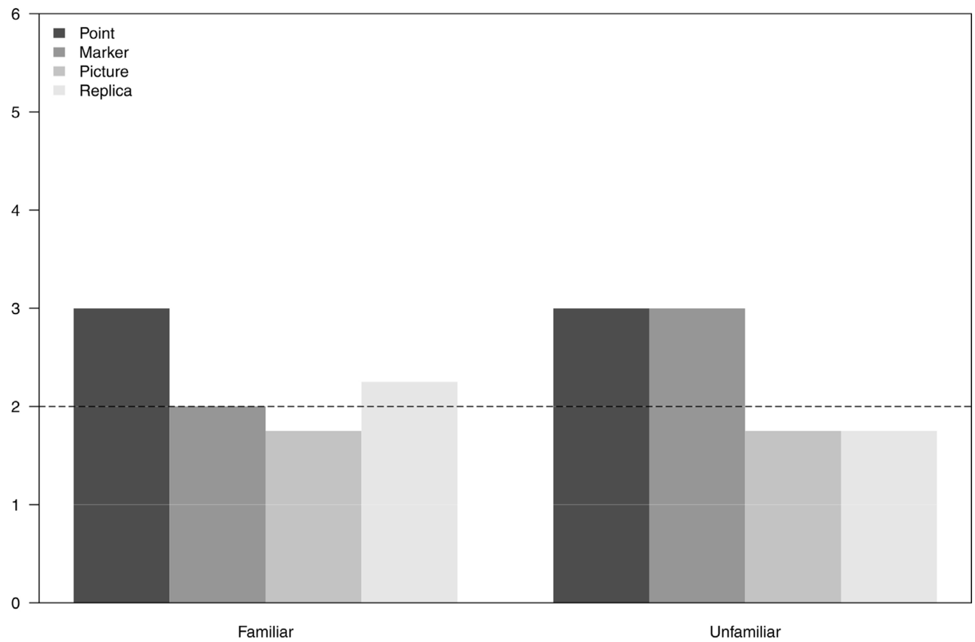

Fig. 4. Chimpanzees' mean scores for four vehicle types and for Familiar and Unfamiliar Helper.

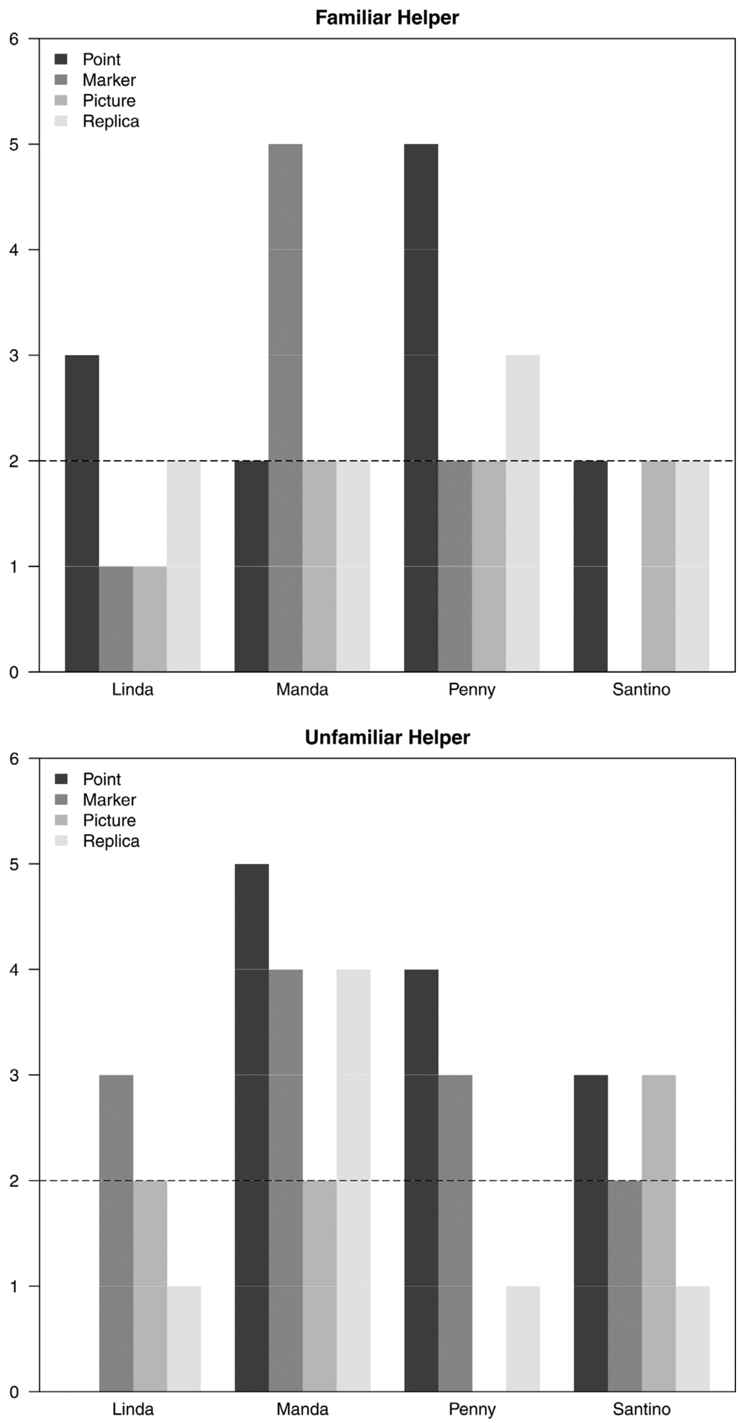

Fig. 5. Chimpanzees' individual scores with Familiar (a) and Unfamiliar (b) Helper.

3.1.4. Predictions

The first three predictions were essentially the same as those for children: (a) Point/Marker would yield better performance than Picture/Replica; (b) Point would be at some advantage compared to Marker, and (c) if the iconic vehicles were understood at all, rates for Picture would be higher than for Replica. In addition, we predicted (d) participants would perform better in the case of the familiar than the unfamiliar Helper.

3.2. Results and discussion

There was considerable variation in time taken to reach criterion in the pretest. It took four trials for Santino (starting with the unfamiliar Helper), as well as for Penny (with the familiar Helper). Linda (with familiar Helper) reached criterion in 13 trials and Manda (with familiar Helper) in 34 trials (divided into two sessions, with 17 trials in the last session).

Mean scores for each vehicle type are shown in Fig. 4. Consistent with the first prediction, mean scores for Point (familiar Helper) and Point and Marker (unfamiliar Helper) were higher than for Picture and Replica. However, none of these was significantly above chance, according to one-sample *t*-tests. For familiar Helper, Point = 3.00, $t(3) = 1.41$, $p = .252$, Marker = 2.00, $t(3) = .000$, $p = 1.000$, Picture = 1.75, $t(3) = -1.00$, $p = .391$, and Replica = 2.25, $t(3) = 1.00$, $p = .391$. For unfamiliar Helper, Point = 3.00, $t(3) = 0.93$, $p = .423$, Marker = 3.00, $t(3) = 2.45$, $p = .092$, Picture = 1.75, $t(3) = -.40$, $p = .718$, and Replica = 1.75, $t(3) = -.33$, $p = .761$. A repeated measures ANOVA, with vehicle type and Helper type as within-subject factors, no significant effects appeared.

Individual scores for the four participants are shown in Fig. 5. As with the children, a score of five of six was considered significantly above chance. In three cases a participant reached a score of five of six. Two of these three involved Point and the third involved Marker. Two occurred with the familiar, and one with the unfamiliar Helper.

In sum, the first prediction that Point/Marker would be easier than the other vehicles for the chimpanzees, as for the children, was partially confirmed by the tendency in the mean scores (Fig. 4) and by the fact that the only three cases in which a participant scored five of six trials were from this group of indexical, directional, non-representational vehicles. There was no support for any of the other predictions: that Point would be easier than Marker, Picture would be easier than Replica (performance for both was at chance), and better performance with the familiar Helper.[4]

4. General discussion and conclusions

The results provide at least tentative answers to our research questions. There was no clear evidence that Marker was easier than Point for either the children or the chimpanzees. We believe this is so due to the way it was used (i.e., involving a manual action directed toward the referent, with gaze alternation between participant and referent), the Marker vehicle does in fact function similarly to Point. The higher familiarity of Point played little, if any, positive role.

There was no indication that the actual marker, the “post-it” note placed on top of the indicated box, needed to be understood as a representation (i.e., fully-fledged sign), since this would have been indicated by developmental delay. Such a delay was obvious for Picture and Replica, compared to both Point and Marker. While Replica appeared somewhat more difficult than Picture for 30-month-olds (Fig. 2), it was remarkable how similar the two vehicles were, especially when considering individual above-chance performance: Both were reliably comprehended only by 30-month-olds, at a nearly identical proportion (Fig. 3). A similar result appeared for the chimpanzees: While the mean scores were not significantly above chance for any of the vehicles, and there were no main effects (possibly

[4] One of the two individuals who performed best with the indexical cues (Manda) was also the one who received most pretest trials (which may also be seen as indexical in nature). But since training occurred until a set criterion was met, most of this experience was made up of failures to indicate the baited box. That a subject succeeds with indexical cues after showing problems with indexical actions can hardly be taken as evidence for latent learning. Furthermore, in a meta-analysis of 28 object-choice studies involving experimenter-given cues, comprising 70 experiments, Mulcahy and Hedge (2012) found evidence for learning effects in only three cases (4.3%).

due to the small size of the sample), two participants (Manda and Penny) performed above chance only for Point and Marker, while all the chimpanzees failed on Picture and Replica alike (Fig. 5).

What is the reason for this divergence? Given that Point and Marker on the one hand, and Picture and Replica on the other, share a number of different properties (Table 1), we cannot provide a definite answer. First, there is the contrast in medium: Point and Marker involved bodily acts (once the act of placing the marker is taken into account), while Picture and Replica involve essentially "extra-bodily" vehicles. Second, the vehicles of the first pair showed the direction to the referent, while the latter did not. Third, there is the contrast in (dominant) semiotic ground: for Point/Marker this is indexicality (contiguity), while for Picture/Replica, iconicity (similarity). Fourth, Point and Marker do not need to be understood as representations (signs), while Picture and Replica do. To the differences in Table 1, we should add a fifth: Not only the representational relation of Picture/Replica to the respective box needed to be understood, but also that this was *used* with a particular communicative meaning ("the reward is there"). Indeed, there were a number of cases in which a child showed understanding of the first (e.g., commenting to the researcher, "Oh, it is the same," referring to the color photograph and the box), but then failed to make the correct choice. To tease apart these five factors, studies with new semiotic vehicles need to be performed, e.g., using iconic gestures "depicting" the shape of the correct box, or indexical vehicles "without bodily deixis" (Moore, Leibal, & Tomasello, 2013).

Still, our findings serve to safeguard against generalizing from some previous proposals. One cannot simply extend the hierarchy Icon < Index < Symbol, proposed by Peirce (1958) for reasons of semiotic theory, to the claim that iconicity should be more basic than indexicality in either evolution (Deacon, 1997) or child development (Lüdtke, 2012). Nor can one assume the opposite position, i.e., that iconicity in general is difficult for children and hence cannot have played any important role in the evolution of language (Cartmill & Goldin-Meadow, 2012). For one thing, iconicity as semiotic ground is not the same as *iconic signs*, where the expression is understood as standing for the referent (Persson, 2008; Sonesson, 1995, 2009). For another, iconic signs, even in the same modality such as gestures, vary among themselves, e.g., reenacted (pantomimic) gestures are produced by children earlier than those in which the child represents an object or an event by using the shape of the hand and its trajectory of movement (Andrén, 2010; Zlatev, submitted for publication).

Furthermore, the features shared by Point and Marker(-placing): Bodily, Indexical and Direction (Table 1), are not in themselves (nor in combination) sufficient to explain the results, since these features are also present in pointing without ostensive cues, which was not found to be helpful for children in an object-choice task (Behne et al., 2005). So while the (weak) positive tendency for Point and Marker in the chimpanzee study could possibly be explained in terms of indexicality (functioning as a discriminative cue), and/or directionality (the vehicles resemble an act of reaching/grabbing) alone, this cannot be the case for the children. Ostensive mutual gaze, serving to enact communicative intent (Csibra, 2010; Moore, submitted for publication), may be necessary (at least when sufficient common ground is not yet established) to help these features coalesce into a communicative act.

The same ostensive cues were present for Picture and Replica (and even elaborated, with gaze-alternation to the vehicle as well), but the performance of 18-month olds was at chance level and none of the 24-month olds performed above chance individually (Fig. 4). Thus, the higher degree of "semiotic complexity" (Andrén, 2010; Zlatev, submitted for publication) of Picture and Replica as semiotic vehicles, i.e., as fully-fledged signs (representations) clearly played a role. Still, the fact that only 30-month-olds reliably passed the task with Picture and Replica shows that it is not simply a matter of acquiring "the symbolic function" (Piaget, 1962): it is both much too late in development to appeal to this (even at 24 months, their average productive vocabulary surpassed 300 words, according to SECDI), and there were no correlations between SECDI vocabulary scores and object-choice performance. Thus, we can tentatively suggest that it may be the *combination* of the factors (a) bodily vs. external medium, (b) indexical/directional vs. iconic ground, and (c) non-representational vs. representational status that together contributed to the significantly better performance of the children for Point and Marker compared to Picture and Replica.

The question of a positive correlation between language development and task performance, and thus a possible causal role of language for the comprehension of other semiotic vehicles, can be given a tentative negative answer. Hence, our object-choice study with children has given some additional support to cognitive-semiotic approaches to development (and evolution), that do not overemphasize language, but rather see it as in part dependent on other semiotic skills (Donald, 1991; Piaget, 1962; Tomasello, 1999a, 2008; Zlatev, 2007, 2008a, 2009). At the same time, this should not be taken as implying that language development plays no significant role in semiotic development. After all, linguistic utterances are a particular form of semiotic vehicle and most (although not all, e.g., "hello", "yes", "this"...) words are fully-fledged signs/representations (verbs of actions/states, adjectives of properties, nouns of things, etc.). Thus, a (strong) effect of language on the establishment, if not the emergence, of the "symbolic function" cannot be excluded.

Indeed, while individual production and comprehension scores did not predict performance on the particular non-linguistic semiotic vehicles, a major vocabulary spurt between 18 and 24 months was noteworthy (also well-attested in the literature), according to the SECDI production list of 710 words, from 35 to 305 on average. And this spurt correlated with the developmental spurt in Point and Replica comprehension during the same period. This "correlation" might be interpreted as follows: It is not understanding of representations ("the symbolic function") that makes the big difference around 18 months (and even earlier), but rather the understanding of communicative intent together with semiotic vehicles that are not (yet) understood as representations. Later in development, the progressive understanding of communicative intentions can go hand in hand with development in the complexity of semiotic vehicles: from bodily to extra-bodily, from iconic to symbolic, from pictorial, 2-dimensional images to 3-dimensional replicas. Thus, as suggested earlier, the dimensions of semiotic complexity and communicative explicitness (Andrén, 2010; Zlatev, submitted for publication), or "vehicles of content" and "acts of ostension" (Csibra, 2010; Moore, submitted for publication) need to be both distinguished and considered in tandem. It is not *either* "speaker meaning" or "sentence meaning" (Grice, 1957), "inference" or "code" (Sperber & Wilson, 1995), "pragmatics" or "semantics", but rather both. We suggest that the gist of the Gricean approach focusing on communicative intentions, especially when taken in an embodied-enactive perspective (Moore, submitted for publication), can be combined with a semiotic approach focusing on meaning, in a productive synthesis that can describe how pragmatics (use and intent) pave the way for the development of semantics (representation and convention), within an overall semiotic and experiential horizon.

How can the results from the chimpanzee study be interpreted within such a cognitive-semiotic perspective? The expectation of better performance with a familiar than an unfamiliar Helper was not confirmed. While we should be careful in generalizing from one particular familiar Helper and four chimpanzee participants, this finding suggests that lack of emotional closeness with a (human) communicator is not among the major factors that prevents chimpanzees from understanding communicative intentions (Leavens et al., 2008). Conversely, the fact that the communicator was unfamiliar to the children did not prevent them from succeeding. Since this was an age in which their productive language was minimal (median score of 35 words for the Production score), and furthermore there were no clear correlations between vocabulary scores and task performance, our findings confirm conclusions from earlier research that the communicative object-choice task is tapping into a cognitive capacity that is well-developed in children, while underdeveloped in chimpanzees.

Still, is not yet clear what this capacity is. Tomasello's (2008) explanation that non-human apes fail in such tasks because they are essentially more "selfish" in their interactions with each other than human beings, and therefore fail to understand communicative and cooperative intentions, remains a possibility. This explanation would be further strengthened if future studies were to confirm that to the extent that apes succeed in object-choice tasks, they do so by treating semiotic vehicles as discriminative cues, i.e., on the basis of indexicality alone. As suggested in earlier studies, indexicality can at least in some cases serve as a powerful cue for learning contingencies in non-human primates (Hermann et al., 2006; Persson, 2008). However, when raised in a "sociolinguistically rich environment", chimpanzees and bonobos predominantly succeed in understanding the communicative intent of pointing in an object-choice task (Lyn et al., 2010). Hence, an explanation in which apes are fully incapable of either enacting or comprehending communicative intentions seems problematic.

At the same time, it remains to be explained exactly how a high level of human "enculturation" influences the cognition of apes. Consistent with our current analysis, we may conclude by suggesting that this involves not only cognitive skills such as "intention reading", nor only "semiotic mediation" in a Vygotskyan analysis (Vygotsky, 1978), but their ability to combine the two aspects in a unified act of (bodily) communication, involving semiotic vehicles (of different complexity) stating *what* is being communicated and communicative intentions showing *that* one is communicating this (altruistically) to an addressee. An environment in which such acts are both routinely presented to them, and their own attempts in this direction rewarded, would naturally contribute to developing such a nascent capacity in non-human apes, as well as for its further strengthening in hominid evolution.

Acknowledgements

We thank The Swedish Research Council for financing the research project *Precursors of Sign Use in Intersubjectivity and Imitation* (PSUII), as well as *Riksbankens Jubileumsfond* (RJ) for its grant to the Centre for Cognitive Semiotics (CCS) at Lund University, which jointly provided the infrastructure and intellectual environment for conducting the present research. In addition, we express our gratitude to Ing-Marie Persson and the chimpanzee caretakers at Furuvik Zoo for their help with the animal study. Finally, we thank the editor and three anonymous reviewers for many constructive comments which led to substantial improvements of this article.

References

Andrén, M. (2010). *Children's gestures between 18 and 30 months*. Department of Linguistics, Lund University (published Doctoral dissertation).
Aureli, T., Perucchini, P., & Genco, M. (2009). Children's understanding of communicative intentions in the middle of the second year of life. *Cognitive Development*, *24*, 1–12.
Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year olds comprehend the communicative intention behind gestures in a hiding game. *Developmental Science*, *8*, 492–499.
Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, *30*(3), 359–375.
Berglund, E., & Eriksson, M. (2000). Communicative development in Swedish children 16–28 months old. The Swedish early communicative development inventory – Words and sentences. *Scandinavian Journal of Psychology*, *41*(2), 133–144.
Call, J. (2001). Body imitation in an enculturated orangutan (*Pongo pygmaeus*). *Cybernetics and Systems*, *32*, 97–119.
Call, J., Hare, B., Carpenter, M., & Tomasello, M. (2004). 'Unwilling' versus 'unable': Chimpanzees' understanding of human intentional action. *Developmental science*, *7*(4), 488–498.
Call, J., & Tomasello, M. (1996). The effect of humans on the cognitive development of apes. In A. E. Russon, K. A. Bard, & S. T. Parker (Eds.), *Reaching into thought* (pp. 371–403). New York: Cambridge University Press.
Call, J., & Tomasello, M. (2005). What chimpanzees know about seeing revisited: An explanation of the third kind. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (pp. 45–64). Oxford: Oxford University Press.
Campbell, M., & de Waal, F. (2011). Ingroup–outgoup bias in contagious yawning by chimpanzees supports link to empathy. *PLoS ONE*, *6*(4), e18283. http://dx.doi.org/10.1371/journal.pone.0018283
Carrasco, L., Posada, S., & Colell, M. (2009). New evidence on imitation in an enculturated chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, *123*(4), 385–390.
Cartmill, E. A., & Goldin-Meadow, S. (2012). Is pantomime a likely stage in language evolution? Evidence from human and primate gesture. In Th. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, & J. Hurford (Eds.), *The evolution of language: Proceedings of the 9th international conference*. Singapore: World Scientific.
Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141–168.
Custance, D., Whiten, A., & Bard, K. (1995). Can young chimpanzees (*Pan troglodytes*) imitate arbitrary actions? *Behaviour*, *132*, 837–859.
Deacon, T. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.
DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences*, *8*(2), 66–70.
Donald, M. (1991). *Origins of the modern mind. Three stages in the evolution of culture and cognition*. Harvard: Harvard University Press.
Eriksson, M., & Berglund, E. (2002). *Instruments, scoring manual and percentile levels of the Swedish Early Communicative Development Inventory, SECDI. FoU-rapport Nr 17, FoU-nämnden*.
Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N. J. Enfield, & S. C. Levenson (Eds.), *Roots of human sociality: Culture, cognition, and human interaction* (pp. 229–255). Oxford: Berg Publishers.
Gómez, J. C. (1994). Mutual awareness in primate communication: a Gricean approach. In S. Parker, M. Boccia, & R. Mitchell (Eds.), *Self-recognition and awareness in apes, monkeys and children* (pp. 61–80). Cambridge, MA: Cambridge University Press.
Grice, H. P. (1957). Meaning. *Philosophical Review*, *LXVI*, 377–388.
Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785.
Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151.
Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495–514.
Hare, B., & Tomasello, M. (2004). Chimpanzees are more skillful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*, 571–581.
Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction*. Oxford: Oxford University Press.
Hermann, E., Melis, A. P., & Tomasello, M. (2006). Apes' use of iconic cues in the object-choice task. *Animal Cognition*, *9*, 118–130.

Hribar, A., Sonesson, G., & Call, J. (2013). From sign to action. Studies in chimpanzee pictorial competence. *Semiotica* (in press).
Hurely, S., & Chater, N. (2005). . *Perspectives on imitation. From neuroscience to social science* (Vols. I & II) Cambridge, MA: MIT Press.
Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*, 224–234.
Kuhlmeier, V. A., & Boysen, S. (2002). Chimpanzees (*Pan troglodytes*) recognize spatial and object correspondences between a scale model and its referent. *Psychological Science*, *13*(1), 60–63.
Leavens, D. A., Hopkins, W. D., & Bard, K. A. (2008). The heterochronic origins of explicit reference. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 187–214). Amsterdam: Benjamins.
Leekam, S. R., Solomon, T. L., & Teoh, Y. (2010). Adults' social cues facilitate young children's use of signs and symbols. *Developmental Science*, *13*(1), 108–119.
Lipps, T. (1903). Einfuhlung, innere Nachahmung und Organenempfindung. *Archiv fur die gesamte Psychologie*, *I*(2/3), 185–204.
Lüdtke, U. (2012). Relational emotions in semiotic and linguistic development: Towards an intersubjective theory of language learning and language therapy. In A. Foolen, U. Lüdtke, T. Racine, & J. Zlatev (Eds.), *Moving ourselves, moving others: Motion and emotion in intersubjectivity, consciousness and language* (pp. 305–346). Amsterdam: Benjamins.
Lyn, H., Russell, J. L., & Hopkins, W. D. (2010). The impact of environment on the comprehension of declarative communication in apes. *Psychological Science*, *21*(3), 360–365.
Melis, A., Call, J., & Tomasello, M. (2006). Chimpanzees (*Pan troglodytes*) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154–162.
Moore, R. (2013) Enacting and understanding communicative intent. (submitted for publication).
Moore, R., Liebal, K., & Tomasello, M. (2013). Three-year-olds understand communicative intentions without language, gestures, or gaze. *Interaction Studies*, *14*(1), 62–80.
Mulcahy, N. J., & Hedge, V. (2012). Are great apes tested with an abject object-choice task? *Animal Behaviour*, *83*(2), 313–321.
Myowa-Yamakoshi, M., & Matsuzawa, T. (2000). Imitation of intentional manipulatory actions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *114*, 381–391.
Over, H., & Carpenter, M. (2012). Putting the social into social learning: Explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*, 182–192.
Peirce, C. S. (1958). *Collected papers of Charles Sanders Peirce*. Cambridge: Harvard University Press.
Persson, T. (2008). *Pictorial primates. A search for iconic abilities in great apes*. Lund: Media-Tryck.
Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: Norton.
Segerdahl, P., Fields, W., & Savage-Rumbaugh, S. (2005). *Kanzi's primal language. The cultural initiation of primates into language*. Basingstoke: Palgrave Macmillan.
Silva, K., Bessa, J., & de Soussa, L. (2012). Auditory contagious yawning in domestic dogs (*Canis familiaris*): First evidence for social modulation. *Animal Cognition*, *15*(4), 721–724.
Sonesson, G. (1989). *Pictorial concepts. Inquiries into the semiotic heritage and its relevance for the analysis of the visual world*. Lund: Aris/Lund University Press.
Sonesson, G. (1995). On pictorality. The impact of the perceptual model in the development of visual semiotics. In Th. Sebeok, & J. Umiker-Sebeok (Eds.), *The semiotic web 1992/93: Advances in visual semiotics* (pp. 67–108). Berlin/New York: Mouton de Gruyter.
Sonesson, G. (1998). Index/indexicality. In P. Bouissac (Ed.), *Encyclopedia of semiotics*. Oxford: Oxford University Press.
Sonesson, G. (2009). The view from Husserl's lectern: Considerations on the role of phenomenology in cognitive semiotics. *Cybernetics and Human Knowing*, *16*(3/4), 107–148.
Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell.
Strawson, P. F. (1964). Intention and convention in speech acts. *Philosophical Review*, *73*(4), 439–460.
Suddendorf, T. (2003). Early representational insight: Twenty-four-month-olds can use a photo to find an object in the world. *Child Development*, *74*(3), 896–904.
Tomasello, M. (1999a). *The cultural origins of human cognition*. London: Harvard University Press.
Tomasello, M. (1999b). Having intentions, understanding intentions, and understanding communicative intentions. In P. Zelazo, D. R. Olson, & J. W. Astington (Eds.), *Developing theories of intention* (pp. 63–75). Cambridge: Cambridge University Press.
Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
Tomasello, M., & Call, J. (2004). The role of humans in the cognitive development of apes revisited. *Animal Cognition*, *7*, 213–215.
Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, *68*(6), 1067–1080.
Tomasello, M., & Carpenter, M. (2005). Intention reading and imitative learning. In S. Hurley, & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science: Vol. 2. Imitation, human development, and culture* (pp. 133–148). Cambridge, MA: MIT Press.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–691.
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Zlatev, J. (2007). Intersubjectivity, mimetic schemas and the emergence of language. *Intellectica*, *2/3*(46/47), 123–152.
Zlatev, J. (2008a). The coevolution of intersubjectivity and bodily mimesis. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 215–244). Amsterdam: Benjamins.
Zlatev, J. (2008b). From proto-mimesis to language: Evidence from primatology and social neuroscience. *Journal of Physiology – Paris*, *102*, 137–152.
Zlatev, J. (2009). The semiotic hierarchy: Life, consciousness, signs and language. *Cognitive Semiotics*, *4*, 169–200.
Zlatev, J. (2013). Image schemas, mimetic schemas and children's gestures. *Gesture* (submitted for publication).
Zlatev, J., Persson, T., & Gärdenfors, P. (2005). Triadic bodily mimesis is the difference. Commentary to M. Tomasello, M. Carpenter, J. Call, T. Behne and H. Moll "Understanding and sharing intentions: The origins of cultural cognition". *Behavioral and Brain Sciences*, *28*, 720–721.