On the evolutionary origin of declarative pointing

Brinck, Ingar

2004

BACKGROUND
The distinction between imperative (IP) and declarative (DP) pointing is central to theories about language development and the evolution of language as concerns both apes and humans. In earlier work, I have presented an analysis of acts of DP in terms of their illocutionary function to indicate, classifying the other functions as their perlocutionary effects (Brinck, 2004a). I then became interested in DP’s broader and more fundamental role in evolution and development. I found that no proper distinction was made between the content of acts of DP and this fundamental role. I was also struck by the vague nature of well-known analyses of DP, in particular the one that characterises DP as a way of sharing experiences. How might such an analysis constitute the basis for a substantive account of DP’s ontogenetic and phylogenetic role? To put it crudely: What is the survival value of sharing experiences?

There is a general agreement among researchers in different disciplines that communication, theory of mind, and empathy are related both ontogenetically and phylogenetically. To me it seems reasonable to start uncovering the origin of DP by linking DP to intersubjectivity, since the normal development and proper functioning of DP is dependent on the normal development and proper functioning of intersubjectivity.

It seems plausible to assume that the interrelation between DP and intersubjectivity was forged because it has an ontogenetic and phylogenetic value for humans — DP being specific to humans. This assumption constitutes the background of the hypothesis that I will put forward in this talk. I will examine the hypothesis that the ontogenetic and phylogenetic roles of DP is to mutually evaluate the shared object of attention.

METHODOLOGY
The content of the article is based in an extensive reading of experimental data concerning pointing in human children and non-human primates. The material comes from cognitive psychology, developmental psychology, linguistics, neuroscience, and primatology.

ON THE (EVOLUTIONARY) ORIGIN OF DECLARATIVE POINTING*

Ingar Brinck, Lund University, Sweden
ingar.brinck@fil.lu.se

Condensed version 2004-06-01

Abstract
Imperative and declarative pointing are distinct kinds of communicative acts that rely on different cognitive capacities in the speakers. Declarative pointing is an important precursor to language, seen from both an evolutionary and a developmental perspective. Declarative pointing is functionally independent of affective intersubjectivity, yet it is intimately related to it in development. It is argued that declarative pointing once evolved because it allows for the mutual evaluation of joint objects of attention. Interaffectivity and joint attention to a distal object together constitute the prerequisites for using declarative pointing for the purpose of evaluation. Mutual evaluation has the benefits of enhancing co-operation and allowing for vicarious learning. It also makes possible the non-linguistic, active interrogation of others about their attitudes to jointly attended objects.

I. IMPERATIVE VS. DECLARATIVE POINTING: A COMPARISON

A distinction is made between imperative (IP) and declarative (DP) pointing (Bates, Camaioni, & Volterra, 1975). Producing and understanding acts of declarative pointing require more sophisticated cognitive capacities in the subjects than acts of imperative pointing do. Declarative pointing only occurs naturally in humans, although it has been reported in laboratory-reared apes as well (Leavens, Hopkins, & Bard, 1996; Leavens, Hopkins, & Thomas, 2004). It has a unique position among communicative gestures, seen both from an evolutionary and a developmental perspective, in being a major stepping stone to natural language. The distinction between the two kinds of pointing is rooted in the data reviewed below in the comparison of them in seven different respects.

(i) Age: Acts of both kinds of pointing emerge at about the same time, by 12 months. The production of declarative pointing occurs later than comprehension of it, by 15 months.

(ii) Function: Imperative pointing is used to make the adult do something for the child in order to achieve some end, such as fetching the object for the child (Bates et al., 1975; Bates, 1976). The adult is treated as the means and the object (event) as the goal to be attained by pointing. DP is used to direct the adult’s attention to some event or object in the world. It has an indicating function. The function constitutes the illocutionary force of the act (cf. Austin, 1962; Bates et al., 1975; Bates, 1976; Brinck, 2004a). Interaction with the adult is the ultimate goal. The object is used as a means to share experiences with the adult. The particular aim with the sharing of experiences is determined in the context of use. The aim is interpreted in terms of the perlocutionary effect of the act (cf. Austin, 1962; Brinck, 2004a).

(iii) Sign type: IP is a signal that is based in behaviourally motivated regularities. Acts of IP can be successful without involving a recognition of the other subject’s intentional states, by exploiting automatic, presumably innate, mechanisms such as the Eye Direction Detector and the Shared Attention Mechanism (cf. Baron-Cohen, 1985). Yet imperative pointing does not reduce to sheer ritualised behaviour, such as reaching, because it is intended to signal the reach to the addressee (Brinck, 2004a). It expresses the speaker’s communicative intention.

¹ Leavens, Hopkins, & Bard (1996) suggest that situations that require pointing (i.e., in which attention contact is needed to obtain an object distant to both subjects) do not present themselves to wild apes, which is why referential pointing does not in general occur among them. They furthermore suggest that the development of referential pointing is independent of linguistic competence, even in humans (p. 352). Call & Tomasello (1994) claim that using pointing flexibly to various targets such as food, locations, and objects depends on interacting with others in human-like ways during early ontogeny (p. 315). Leavens & Hopkins (1998) maintain that making the distinction between IP and DP may be inappropriate when talking about apes (p. 820), and contend that the referential pointing of apes blurs this distinction and involves elements of both kinds of pointing.

² DP is the paradigmatic act of nonverbal reference, according to Bates, O’Connell, & Shore (1987). Non-verbal reference may be defined as (i) the capacity of the sender to direct the attention of an observer through deictic and iconic manual gestures, (ii) the sender’s visual orienting behaviour, and (iii) attention-directing tactile or auditory signals (Leavens, Hopkins, & Thomas, 2004). Note that this definition does not presuppose symbol use or the so-called semiotic function (of differentiation between sign and object and substitution of sign for object) (Sonesson, 1989). To compare, Bard (1992, p.1187) defines intentional communication as “the ability to coordinate sequences of behavior involving objects with sequences of behavior involving social agents” in goal-directed activities, which are characterised by persistence until the goal is achieved. According to Leavens & Hopkins (1998, p.813), intentional pointing occurs when the sender points “with reference to the attentional status of the observer”. This is measured by the occurrence of gaze alternation in pointing individuals, in which the sender alternates his or her gaze between a distal object or location and the face of the recipient of a gesture (ibid.). N.B. In a later article, (Brinck, 2008), I give a detailed analysis of intentional communication.

³ The meaning of signals is either innate or acquired through ontogenetic ritualisation based on association learning (Zlatev, 2003, 273). Some signals are used non-intentionally, e.g., when a behaviour pattern is triggered by the context, and the sender cannot choose whether to exhibit it or not.
The following is an example of a rather common, merely instrumental use by an ape of another subject for the same function as in IP, i.e., getting hold of an object. The ape takes the careholder by the arm and leads him or her to the object. That means that the ape directs the attention of the other to the object by physically bringing the other to the object. This behaviour is similar to directly showing an object to the other, but reversed. In the one case, the child brings the object to the care-taker, in the other case, the ape brings the care-taker to the object. DP is symbolic in that it functions as a substitute, or proxy (Bates, 1976; Clark, 1978). It replaces the act of directly bringing the object to and showing it for the adult (or conversely bringing the adult to the object). The child differentiates between the expression (the pointing gesture) and its content (to indicate) Thus DP has a semiotic function (Sonesson, 1989; Zlatev, 2003). By being directed at and indicating the object, an act of DP is about the object. The act also has a socially recognised function and is in that sense conventional (Bates et al., 1979).

(iv) Communicative intention: The communicative intention (CI) of acts of pointing is double: there is the general intention to communicate, and there is the intention to communicate a certain kind of act. The CI of IP is constant between contexts of use. The CI of the act of DP is double, a constant, type CI (corresponding to the illocution to indicate) and a contextually determined, token CI (corresponding to the desired perlocutionary effects of the act as it is performed in the context of use).

(v) The hearer’s response: An appropriate response to an act of IP can be triggered simply by the pointing gesture, which then functions as a perceptual cue (Brinck, 2004a). No alternative responses present themselves to the hearer, but the appropriate response is determined by the act that triggers it. The hearer has the choice only of either giving the appropriate response or refraining from responding. The primary response to an act of DP is determined by the illocutionary force of the act; the hearer is supposed to focus her attention on the indicated object. However, there may also be an optional secondary response, corresponding to the perlocutionary effects that an act of DP may have. A secondary response is in fact what is expected by the child. As her secondary response, the hearer will choose the action that appears to be the most appropriate or relevant one in the current situation, given the information that she has about the physical context and those of the speaker’s intentions that are manifest to her. Which the secondary response in the end will be may be negotiated with the speaker, by an exchange of emotional attitudes between speaker and hearer, signalled by facial expressions, body postures, vocalisations, gestures, etc.

Note that the hearer’s primary response to an act of DP may result from her attending to the speaker’s manifest attentional states, without exploiting explicit higher-order intentions (Brinck, 2004b). Acts of DP have a structural similarity with those communicative acts that Grice describes as conveying non-natural meaning (Grice, 1957). However, the similarity is superficial, since acts of DP do not necessarily include having intentional states about intentions.

Leavens, Hopkins, & Thomas (2004) favour an operational over a functional definition of pointing that can be objectively verified. They claim that also instrumental uses of gestures may be intentional, in the sense that the subjects exhibit the behaviour that constitutes nonverbal reference, as described in footnote 2. They argue against views of intentional communication that require that the subject is to influence the state of knowledge, or the internal, ‘mental’, states of the observer (cf. Baron-Cohen, 1999). This would require that the subject can “distinguish the behavioral correlates of different states of knowledge in an observer” (p. 55). As I see it, the fundamental difference between intentionally instrumental and communicative behaviour occurs when a gesture is intended to signal a behaviour (IP), instead of directly constituting the behaviour (as in the example above).
To sum up, when the hearer recognises the communicative intention behind an act of pointing, whether an act of IP or DP, she is thereby provided with a motivation for responding to the speaker in a particular way (cf. Grice, 1957). She is in control of her response, and may refrain from responding. In the case of DP she may also respond in another way than expected by the sender.

(vi) Developmental origin: IP develops from regularities that arise from ritualised behaviour (Brinck, 2004a; Zlatev, 2003). Learning by ritualisation does not automatically result in a reciprocal and bi-directional use of gestures. A gesture is reciprocal if, when an agent intentionally can produce the gesture, she also would understand somebody else’s use of it (and conversely). Learning by imitation results in a reciprocal use, because imitative learning requires understanding the imitated subject’s perspective (Tomasello, 1996). DP develops from imitation, a way of learning that requires the capacity to distinguish the means used to achieve a goal from the goal itself, and to vary means and goals independently.

(vii) Attention: In IP, the subject directs herself to the other as to a self-propelling causal agent that one can indirectly influence to do something for oneself by directing his or her attention. Attention is perceived as being related to intentional, goal-directed action. IP requires mutual-object focusing based in attention to the other subject’s attentional state as it is manifested in behaviour taken together with behavioural co-ordination (Baron-Cohen, 1997; Brinck, 2001). DP requires joint attention based in attention to the other subject’s goal-intended attention, by attention contact, gaze alternation and the following of gaze direction (Brinck, 2001). Goal-intention, as opposed to mere goal-direction, is characterised by the subject’s ability to direct his or her attention to an object that is not in his or her field of vision. It is based in the expectancy and anticipation of an object of attention. In an act where the object is occluded, the hearer is implicitly expected by the speaker to be able to attend to a merely anticipated object.

Attention contact between the speakers, such as eye-contact, displays the speaker’s intention to communicate with the hearer (Gómez, 1994). Achieving joint attention to an object requires the speakers’ having attentional states about attentional states as well as having intentional states about attentional states, but, crucially, not intentions about intentions (Brinck, 2004b). The message that acts of DP convey can be understood by attention-reading only; intention-reading is normally not necessary.

---

5 In the context of this article, an attentional state is defined as a state of vigilance directed at an object (or event) external to the subject of the state.

6 As evidenced by Umiltà et al. (2001), apes recognise grasping actions also when the object is occluded, when shown the object the moment before it is occluded. This experiment does not settle that apes understand goal-intended actions, in which the target of action is only anticipated and has not been observed in the present context.

7 Attention contact arises between social agents that recognise each other as social agents. Following Reddy (1999), I relate the concept of a social agent to the capacity for intersubjectivity, and especially interaffectivity, that develops during the first months. As noted by Calder et al. (2002, p.1130), gaze indicates attention in the subject who is looking and the goal of his or her attention. Gaze also triggers the mechanisms involved in the attribution of goals and intentions to others (ibid.) Calder et al. (2002) show that brain areas activated during the process of ToM tasks such as attribution of intentions (the medial prefrontal cortex), are as well engaged during the processing of another subject’s gaze. Calder et al. moreover suggest that the attribution of intentions to others by gaze processing operate in a mandatory, or automatic and reflexive, fashion (p. 1137).

8 An intentional state is here defined as a state that is directed at, and thereby is about, an object (or event) that not necessarily is present to the subject’s senses. At a representational level, intentional states are structured and have stable core representations as constituents, which pave the way for recursive operations on the contents of the intentional states as well as for productivity.

9 For an analysis of readings of attention and intention, see Brinck (2004 b).
These comparisons reveal several fundamental differences between IP and DP. I conclude that IP and DP are distinct communicative acts that require distinct cognitive capacities to be performed. DP does not develop, nor did it directly or continuously evolve, from IP. Why did DP evolve? Why was this behaviour once selected for?

II. INTERSUBJECTIVITY AND COMMUNICATION
The close relation between intersubjectivity and declarative pointing in development is often stressed. Which is the exact nature of this relation?

I distinguish between three kinds of intersubjectivity, which in principle are mutually independent. The first kind is interaffectivity (IAF) that is based in empathy. It consists in having attentional states directed at emotional states, and in the ability to emotionally respond to and reciprocate the emotions of other subjects. The second kind is interattentionality (IAT), which is based in the capacity for joint attention as this was described in section 1. The third kind is interintentionality (INT), a capacity that requires having intentions about (perceivable) intentional states. It thus supports the having of a fully-fledged theory of mind, although it is not intrinsically connected to theories of mind.

There is prima facie evidence that joint attention (as it occurs in DP), seen from a functional point of view, is distinct from interaffectivity. They can develop independently of each other. Some autistic children learn to point declaratively, in spite of their impairment, though they do so later than non-autistic children (cf. Camaioni, Perucchini, Muratori, & Milone, 1997). At least autistic children with a high IQ do develop joint attention behaviour, that is, the kind that requires mutual attention and eye contact (Mundy, Sigman, & Kasari, 1994). Garfield, Peterson & Perry (2001) maintain that the fact that high-functioning autists can develop a theory of mind when they have acquired linguistic skills (by 12 years), but not before, shows that the development of a theory of mind depends on language learning and (cognitive) social intelligence. Thus, IAF is not in principle necessary for either producing or understanding acts of DP.

Moreover, as evidenced by children with Williams syndrome, to develop IAF and IAT, INT is not necessary. Subjects with Williams syndrome have impaired language abilities and IQ, while hypersociable and affectively sensitive (Bellugi, Adolphs, Cassady, & Chiles, 1999). The capacity for joint attention, and thus for DP, does not seem to be intrinsically emotional or affective. Furthermore, the analysis of joint attention has made clear that having intentions about intentions are not necessary for achieving joint attention (Brinck, 2004b).

Yet, the ontogenesis of DP seems intimately related to IAF. Learning how to point declaratively is delayed by impairments that influence the capacity for interaffectivity. According to Franco & Butterworth, children who point declaratively naturally direct themselves to others as subjects of experience, that is, as subjects that can “feel interest, amusement or some emotion” (1996, 333). This view is shared by M. Tomasello (1999) and others, who maintain that the role of DP simply is to share experiences.

Leavens, Hopkins, & Thomas (2004) suggest that referential pointing is a synapomorphy, i.e., a shared, derived trait, or similarity, within Hominoidea, the common ancestor of humans and the great apes, more than 6 mya.

Baron-Cohen (1995, p.66) note that all sorts of declarative gestures, pointing, showing, etc, are missing in young children with autism, whereas they have no problem with behaviour based in individual attention. Baron-Cohen attributes this deficit to a failure in the shared attention mechanism (SAM). The key function of SAM is to provide a drive toward establishing what is of shared interest between the self and another person, independently of modality.
It seems that IAF is part of the ‘proper’, or normal, way of acquiring joint attention behaviour, as in DP, although this behaviour can be otherwise acquired by the help of linguistic skills and, perhaps, ‘cold’ interintentionality.

It is reasonable to assume that the close relation between IAF and DP exists because it either has a particular evolutionary force, or a specific developmental value, or both. Next I will present a hypothesis about the ontogenetic and phylogenetic roles of DP, which is based in the recognition of this intimate relation between DP and IAF.

III. HYPOTHESIS
The ontogenetic and phylogenetic roles of DP is to exchange information between the subjects about the value of the target (of the pointing act) for the purpose of evaluating the target. The purpose of evaluating the target can be to reach a joint evaluation, or to reach an individual evaluation of it by taking into account the other subject’s evaluative attitudes to the target.

IV. FLESHING OUT THE HYPOTHESIS
Evaluation of a common target by engaging in DP is described as follows. Information about the value that an object has for an individual in a given context is exchanged between subjects by behaviourally expressing and reading emotional states (i.e., evaluative attitudes) as these are displayed by the actions, body postures, facial expressions, vocalisations, etc, of the subjects. Communication about value ascription and evaluation consequently requires the capacity for interaffectivity. As noticed in section II, IAF is independent of INT. Interaffectivity consists in a ‘co-operative’ manner of communicating. It is co-operative in the sense that it occurs between subjects with similar properties and of equal value and who share a common goal (Reddy, 1999; Tomasello & Rakoscy, 2003). Declarative pointing is intrinsically cooperative.12

Emotional states can only have an evaluative function if they are intentional. Intentional states are relational: they hold between the subject and an object (or event). In the same manner as an intentional state, an emotion can be directed at and about some object. Ben Ze’ev (2000, 50) divides the intentionality of emotions into three components. The cognitive component consists of information about the given circumstances, the evaluative one assesses the personal significance of the information, and the motivational one concerns the subject’s desires, or action readiness, in those circumstances.

The object that the emotion is about is related to the cause that triggered it.13 By being about that object the emotion expresses a certain attitude to it, which corresponds to a particular kind of action readiness in the subject (Frijda, 1986).

IAF is physically grounded in the mirror neuron system (Gallese & Goldman, 1998; Gallese et al., 2002). IAF, as well as the capacity for reading the emotions of others, develop from infant imitation. The progress can be described stepwise with the following four types of affective intersubjectivity (cf. Nadel, 2004, 2005; Rochat, 2002).

i. Infant imitation is automatically triggered and biologically determined, that is, innate.

---

12 This claim has received independent support from recent research on pointing by esp. Ulf Liszkowski in Michael Tomasello’s research group, who underlines that communicative pointing is for sharing attitudes and exchanging information, and not merely for achieving joint visual focus. For a summary see Liszkowski (2011).

13 The cause and the object of the emotion can be related in several distinct ways. An analysis of which particular ways will be left out of the present discussion.
ii. Active reciprocation (1-3 months) consists in social smiling and turn-taking, and is sometimes described as a ‘proto-conversation’. Its aim is to express and share basic emotions.

iii. Dyadic social referencing (4-6 months) occurs when the child is able to decode and react to the adult’s behaviour. The child will not only imitate, but also respond to the emotions of the other in a way the seems appropriate to both child and adult (e.g., hesitation is a response to anger or irritation).

iv. Triadic social referencing (6-9 months) adds referential looking to the prior stage of dyadic social referencing. Emotions acquire intentional properties. Triadic social referencing is a mechanism of socialisation in children. By 1 year, emotional information influences infant behaviour towards the attended object.\(^\text{14}\)

The following is a necessary condition for DP’s ultimately serving to exchange evaluative attitudes around a joint object:

\(\text{(NC) Evaluation by DP requires that the subjects have the capacity to both express and read emotions.}\)

If the child could not both read and express emotions, and thereby get feedback on her reactions to the adults emotional behaviour, the following drawbacks would occur:
- the child will not be able to know if she expresses the emotions that the adult expects
- the child will not be able to know if she expresses the expected emotions in a correct way
- the adult will not be able to correct the child in case she has misunderstood the information
- unless the child changes her emotional behaviour in response to the evaluative information that she has gained, she will not be able to communicate an alternative point of view than the adult.

I conclude that making evaluations by engaging in acts of DP requires the capacity for triadic social referencing.

V. REASONS TO THINK THAT DECLARATIVE POINTING HAS EVOLVED FOR THE PURPOSE OF EVALUATION

Using DP for evaluation may be considered as a natural progress from triadic social referencing. Although a powerful mechanism of socialisation, and a useful tool for transmitting knowledge from adult to child about, to the child, unknown objects and events, triadic social referencing is limited in that the child’s role is passive. It allows for the child to seek information about the object to which the adult attends, but not for actively taking part in the process of information-seeking.

There are at least three individual reasons to think that DP has evolved for the purpose of evaluation.

A first reason pertains to the high processing cost of declarative pointing - high in comparison to those of imperative pointing. Learning how to evaluate entities and make joint evaluations of shared entities are valuable skills, the benefit of which would compensate for and motivate the cost of processing.

A second reason concerns the evolutionary advantage that a capacity for making evaluations by DP would have. When the indicating function of DP is combined with an exchange of information about how the speaker and hearer evaluate a joint object, DP has a strong

\(^{14}\) Russell, Adamson, & Bard (1997, p.192) claim that social referencing as a way of communicating about the affective significance of objects is common to humans and chimpanzees.
evolutionary force. Acts of DP will in this case allow for vicariously learning how to evaluate a certain (kind of) object. Such vicarious learning would increase individual, group, and species over-all fitness, and would reduce the need for dangerous trial and error learning. Exchanging attitudes to a joint object by engaging in acts of DP also constitutes a way of indirectly communicating knowledge acquired at earlier times, i.e., information about events that are not present during the actual act of DP. This knowledge is such that only one of the subjects may have experienced it in relation to the (same kind of) object as the one on which the act of DP is centred.\(^1\) Exchanging information about how one evaluates a joint object by manifest negative and positive attitudes also plays an important role for co-operation, in determining and co-ordinating common action. Exchange of information about how one values a common object is especially important to situations that require new ways of co-operating.\(^2\) Examples of such situations are ones that are not common to all of the agents of the group, and thus have not cemented into fixed patterns of behaviour, and also such situations that occur for the first time for the group (e.g., as a response to environmental changes). A third reason is related to the developmental advantages for the child that will result from using declarative pointing to evaluate. These advantages are that, first, the child can intentionally and actively learn about novel objects on her own terms, second, the child can learn about novel objects vicariously, and third, the child acquires a way of deliberately interrogating others about how they evaluate an object that she previously has developed a special interest in. This interrogation is performed by initiating acts of DP, and can consequently occur in the absence of language. A fourth advantage is that it becomes possible for child and adult to settle on a joint evaluation of the shared object. A further aim of this may be to co-ordinate their actions and co-operate.

References

\(^1\) This is an interesting fact in its own right, since, normally, (natural) language is considered necessary for intentionally communicating about that which is not present in the context of communication.

\(^2\) For an account of the cognitive mechanisms behind future-directed co-operation that requires new ways of co-operating, see Brinck & Gärdenfors (2003).


