Title
Explorations of the (Meta) Representational Status of Desire in the Theory-Theory of Mind Framework

Permalink
https://escholarship.org/uc/item/9vd5d69h

Journal

ISSN
1069-7977

Author
Ferres, Leo

Publication Date
2004

Peer reviewed
Explorations of the (Meta)Representational Status of Desire in the Theory-Theory of Mind Framework

Leo Ferres (lferres@ccs.carleton.ca)
Human-Oriented Technology Laboratory
Carleton University, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada

Abstract

Some researchers have proposed that what accounts for children’s earlier ability to reason by means of desire compared to reasoning by means of belief is the fact that desires do not necessarily invoke the ability to metarepresent. In this paper, I argue that this is a misconception stemming from the confusion between desire ascription and simple desire states. In other words, there would be no way to entertain a thought about someone’s desire without metarepresenting, in Leslie’s (1991) terms. I provide some empirical evidence in the fashion of Bartsch and Wellman (1995) that also points in this direction.

The problem

Although the concept of desire is at least as important as the concept of belief for describing, explaining and predicting the behavior of different entities (Fodor, 1987; Dennett, 1978b, 1987), substantially more attention has been paid to the development of the concept of belief in Theory-Theory of Mind research (henceforth, TToM) (Wellman, Cross, & Watson, 2001; Astington, 2001). In fact, quite often, metarepresentation (defined here as the internal representation of an epistemic relation (Leslie, 1991), a “second-order” representation (Sperber, 1999)) has been somewhat fused with the child’s ability to entertain/ascribe a belief that stands for a counterfactual state of affairs (Dennett, 1978a; Davies & Stone, 1995), without any reference to the explanatory power of the concept of desire or to its being an epistemic relation itself.

This incipient collapsing of metarepresentation and reasoning by false beliefs (beliefs that stand for counterfactual states of affairs) in TToM research is due, in part, to the assumption that the concept of belief, and especially that of false belief, taps the child’s metarepresentational capacities; while, arguably, the concept of desire does not. This is a common assumption despite the fact that beliefs and desires share several important characteristics (such as defining opaque contexts, being intentional in the philosophical sense, being subject and object specific (Wellman & Woolley, 1990), etc. to name but a few). In the developmental literature, more often than not, desires have been understood as a special case of mental state ascription; namely, one that does not demand of the agent doing the ascription (a child in our case), that he or she be able to metarepresent.

The previous assumption appears to have been brought about by a body of evidence that suggests that the concept of desire is acquired around a year before the concept of belief (Wellman, 1991; Tan & Harris, 1991; Astington & Gopnik, 1991; Harris, 1996; Bartsch & Wellman, 1995). Thus, in order to make it fit the evidence available from research on the concept of belief, it has been claimed that at least until three or four years of age—that is, until they acquire the concept of belief—desire ascriptions are to be thought of as non-metarepresentational. The reasoning behind this claim, I presume, goes along the following lines: on the one hand, a) reasoning by means of beliefs taps an organism’s (children’s, for example) metarepresentational capacities; on the other hand, b) the capacity to reason by means of beliefs is acquired at age X. Thus, if (a) and (b) hold, then c) metarepresentation should be acquired at age X, and not before X. Now, because of this conclusion, the rest of the argument tells us that if (c) is the case, then e) reasoning by means of desires taps an organism’s metarepresentational capacities if and only if f) the capacity to reason by means of beliefs is acquired at age X. However, it is not the case that (f). Therefore, the argument goes on to say, reasoning by means desire does not tap an organism’s metarepresentational capacities.

However, characterizing reasoning by means of the concept of desire as non-metarepresentational simply because it is acquired before the concept of belief and the latter is, only by assumption (as we saw a couple of paragraphs above), the flagship of the child’s metarepresentational capacities is, at all extents, an ad hoc solution. In these pages, I will take position against the conclusion of the argument that reasoning by means of desire does not tap on metarepresentational capacities. In other words, it is, I will argue, far from clear that reasoning by means of the concept of desire is non-metarepresentational in nature, even though it is acquired a year before the concept of belief.

Desires as Metarepresentational

For reasoning by means of desires to be non-metarepresentational at ages younger than four years means that a child at those ages does not represent other people (or even themselves) as representing the desired object as part of the desire relation. Thus, these children are supposed to be merely in some sort of “connection”
to the desired object. According to this view, then, when a two-year-old child says something like “Peter wants a car”, he or she is not ascribing Peter (representing Peter as entertaining) a desire for a given car, but merely putting Peter in some sort of connection to either the car he wants (Wellman, 1990; Wellman & Bartsch, 1994; Bartsch & Wellman, 1995) or to a hypothetical situation in which Peter has the car (Perner, 1991). Consider the following proposition as a state of affairs in the world; that is, a proposition that holds:

Loreto wants to be in Chile. \[1\]

Within the general Theory of Mind (ToM) framework, there are two ways to interpret this proposition, and they have usually been confused. In a first, trivial interpretation, Loreto is just in a state such that she wants to be in Chile. That merely means that she has tokened the proposition “I am in Chile” in her desire box (Fodor, 1975) and will take action to bring it about that she is in Chile. This interpretation is useless at the time of explaining behavior because the agent trying to explain Loreto’s behavior (maybe Loreto herself) may not know that \textit{that} is the proposition she is tokening in her desire box. The second interpretation, however, is the non-trivial interpretation that to be able to explain someone else’s (even one’s own) behavior in terms of desires, one must entertain a belief about the organism’s desire state (Davies & Stone, 1995). To explain or predict behaviors and actions, it is not enough that we are able to be in desire states (unlike, for example, the case of Simulation-Theory of Mind, see Gordon (1995)). What is a conditio sine qua non is that we are able to entertain beliefs about an organism’s desire state (Dennett, 1987; Sperber, 1999). For example, one way to explain why Loreto is buying a ticket to Chile this morning is to entertain a belief with the embedded proposition in \[1\] above. Thus, in order to engage in folk psychological practice (Davies & Stone, 1995), we need to entertain a thought along the following lines:

\[B_u[D(\text{Loreto}, P) \land \neg P]\] \[2\]

where \(B_u\) stands for the agent’s belief state at the time of ascription of the desire state, \(D\) stands for the “desire” predicate which takes two arguments, the organism to which the agent is ascribing the desire (\textit{Loreto}, in \[2\]) and the organism’s desired state of affairs (the variable \(P\) in \[2\] or “I am in Chile” or any other proposition). For this quasi-formalization of desire ascription to work, \(P\) should also be part of the belief state as a proposition that does not hold; since, for something to be a desire, it is by definition that the conditions are false. Simplifying the issue slightly, it would indeed be a contradiction to desire something that one already has.

Notice further that there is in fact no way to formalize the first (trivial) interpretation of \[1\] above in the TToM framework. You may be able to formalize it for logical purposes as something like \(D(\text{Loreto}, P) \land \neg P\), but that will be of no use to someone trying to explain behavior.

The proposition in \[1\] is independent of any folk psychological theory-theory because it is not tokened as a belief about the world in the mind of a particular agent engaging in folk psychological practice. It is just a true proposition (of the external world) at time \(T\). Thus, part of the argument here is that if it is so difficult for us as adults to imagine a non-metarepresentational account of desire at early ages, then it might be the case that this non-metarepresentational characterization is wrong (Astington & Gopnik, 1991).

There seems to be no obvious alternative formalizing of desire ascription (not desire states) except for \[2\] above. Thus, even at younger ages, every time children talk about their own or other people’s desires, they should be entertaining a thought along the lines of \[2\]. It is hard to characterize the thoughts the child is entertaining when explaining or reporting behaviors by means of desires when the latter are merely understood as “subjective connections” to objects. Suppose for the sake of argument, that children do in fact see desires as a “subjective connection” between the organism they are trying to explain the behavior of and the object that this organism desires. This could be relatively easy to see for desire ascription to organisms other than self. However, it would be hard to believe that when talking about their own desires and explaining their own behaviors by means of desires (“because I wanted to go to the park”), children think of \textit{themselves} as just holding a subjective connection to a state of affairs that does not hold. It is in fact very hard to believe that when reasoning about their own behaviors by means of desires, children are not representing themselves as wanting something in particular; to be in the park, for example. But suppose further that they do not representing themselves as wanting something in particular. It is undeniable that the very act of communicating those desires involve a metarepresentation of both the communicator and the person the speaker is talking to. When communicating, and more so when communicating mental states, there should be mutual metarepresentation of the communicator and the addressee (Sperber, 1999).

Given the arguments above, it is hard to take desire ascriptions to either other people (“Peter wants to have a car”) or to oneself (“I want to be in Chile”) as non-metarepresentational (at least in Leslie’s (1991) terms. Now, assuming desire talk stands proxy for desire reasoning about the behavior of other people (see, for example, Dennett (1978a), Bretherton and Beeghly (1982), Tager-Flusberg (1993), Bartsch and Wellman (1995) and the literature spawned by these studies), then we would expect desire talk to actually tap on the child’s metarepresentational capacities, albeit indirectly. One way to look at this is to follow Wellman and Bartsch (1994) and Bartsch and Wellman (1995). If we were able to tell genuine psychological references to desire apart from mere communicative uses of particular words associated to the expression of desire, then we should find difference between talk about desires and talk about other communicative uses of these words. Specifically, while genuine psychological references to desire should change
as a function of age (because they are presumably tapping on metarepresentational abilities), communicative uses should not. They should not because there is no need to be in a belief state about an epistemic state when using a mental state term for mere communicative purposes. If a child repeats an adult’s utterance, for example, it may be that he or she is just repeating it for the sake of not being silent during an interaction, but without having analyzed the utterance itself. More arguably, when a child says something like “I want a cookie” while the cookie is in plain view, the child might be actually saying something like “pass me the cookie”, without imputing a mental state either to self or to someone else. Instead of entertaining a thought such as \( B_a(D(self, I Have A Cookie) \land \neg I Have A Cookie) \), the child is simply in a desire state, maybe, such that \( D_a(self, I Have A Cookie) \land \neg I Have A Cookie \). But that does not qualify as a metarepresentational state in our terms here. The following study tests the prediction that there should be a difference between communicative uses of “want” (the desire term par excellence (Wellman & MacWhinney, 2000). These data come from the longitudinal observation of spontaneous speech production of 12 children (6 boys and 6 girls) whose ages ranged from 18 months at the time of the first observation to 60 months at the time of the last observation and who were acquiring English as their mother tongue. Each child was observed a total of 10 times, for about 40 minutes each, in 3-month intervals. Since the objective of Well’s (1981) research was to obtain spontaneous speech samples, a timing mechanism was devised to set off a tape recorder – connected to a wireless microphone in the child’s garment – at different times between 9am and 6pm, to prevent parents from planning activities, for example. Twenty-four 90-second samples were recorded in each observation. These were later transcribed into several files using normal English orthography. Table 1 in page 3 gives some general information about the samples, where Ages (mo.) means ages in months, \( N \) means number of participants in each age group (all twelve participants are the same children at different ages), \( \# \text{TotUtts} \) means the total number of utterances in the samples, \( MLU(\overline{x}) \) means the average mean length of utterance for that age group, and \( MLU(\overline{SD}) \) means the standard deviation of the mean for the MLU values for that group.

**Method**

**Data.** A total of 14,896 child utterances were taken from the Wells corpus (Wells, 1981) in the CHILDES database (MacWhinney, 2000). These data come from the longitudinal observation of spontaneous speech production of 12 children (6 boys and 6 girls) whose ages ranged from 18 months at the time of the first observation to 60 months at the time of the last observation and who were acquiring English as their mother tongue. Each child was observed a total of 10 times, for about 40 minutes each, in 3-month intervals. Since the objective of Well’s (1981) research was to obtain spontaneous speech samples, a timing mechanism was devised to set off a tape recorder – connected to a wireless microphone in the child’s garment – at different times between 9am and 6pm, to prevent parents from planning activities, for example. Twenty-four 90-second samples were recorded in each observation. These were later transcribed into several files using normal English orthography. Table 1 in page 3 gives some general information about the samples, where Ages (mo.) means ages in months, \( N \) means number of participants in each age group (all twelve participants are the same children at different ages), \( \# \text{TotUtts} \) means the total number of utterances in the samples, \( MLU(\overline{x}) \) means the average mean length of utterance for that age group, and \( MLU(\overline{SD}) \) means the standard deviation of the mean for the MLU values for that group.

**Procedure.** The first step was to identify all and only the instances of the term “want” in all and only the target child’s exchanges. Once the “want” utterances were identified and cleaned for false positives, they were coded as belonging to one of five mutually exclusive categories: genuine psychological references to desire (GPRDs), behavioral requests (BRs), direct repetitions (DRs), idiom expressions (IEs) and uncodable utterances (UUs). While GPRDs refer to mental states, the other three categories do not, they fulfill a mostly communicative function.

Genuine psychological references to desire (henceforth, GPRDs) are instances of children’s unequivocally referring to themselves or other people as being in a mental state of desire. Behavioral requests, in turn, (henceforth, BRs) are ‘unadorned’ instances in which the child uses a desire term to fulfill an immediate goal, like receiving something that is beyond her reach but in plain view. Bartsch and Wellman (1995) take these instances to mean nothing more than “give me x”. Direct repetitions (henceforth, DRs) are dialog turns in which the child merely repeats the adult (or his own) utterance. Idiomatic expressions (henceforth, IEs) are high-frequency collocation of words in the particular language. This is the case of Spanish “I don’t want to” or “I want more”, when they appear without an object. Uncodable utterances (henceforth, UUs) are instances of the desire term “want” for which categorization was impossible, due mainly to failure in retrieving contextual information from the dialog.

To code each of them into one of the five mutually exclusive categories, child utterances containing “want” were not taken in isolation, but embedded in a window of the four previous and the four following utterances of the whole sample. However, sometimes this short context did not help defining which category the utterance belonged to. Thus, the whole transcript had to be analyzed in order to assign a category to the utterance in question. An independent rater, unaware of the hypothesis of the study rated a subset of the data (10%=60 utterances, Cohen’s \( \kappa= .85 \)). Disagreements were resolved by discussion and, in the light of the discussions, there was a second coding pass to the whole data set.

**Results**

There were a total of 602 “want” utterances in the analyzed corpus. 347 (57.35%) were GPRDs, 145 (23.96) were other communicative uses of “want”. Of those 145 communicative uses, 37 (6.11% of total “want” utterances) were behavioral requests, 92 (15.20%) were direct repetitions and 16 (2.65%) were idiomatic expressions. Figure 1 below shows the average frequency of talk about genuine desires as a percentage of the total number of utterances for each particular child at each particular age. It is evident that the developmental picture I have obtained resembles the one in Bartsch and Wellman (1995) very closely, even the ranges of the percentages are similar (see Bartsch and Wellman (1995),

<table>
<thead>
<tr>
<th>Ages (mo.)</th>
<th>( N )</th>
<th>#TotUtts</th>
<th>( MLU(\overline{x}) )</th>
<th>( MLU(\overline{SD}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-24</td>
<td>12</td>
<td>3483</td>
<td>1.480</td>
<td>0.252</td>
</tr>
<tr>
<td>25-28</td>
<td>12</td>
<td>1830</td>
<td>1.697</td>
<td>0.430</td>
</tr>
<tr>
<td>29-32</td>
<td>12</td>
<td>2694</td>
<td>2.246</td>
<td>0.566</td>
</tr>
<tr>
<td>33-36</td>
<td>12</td>
<td>2972</td>
<td>2.709</td>
<td>0.446</td>
</tr>
<tr>
<td>37-40</td>
<td>12</td>
<td>2041</td>
<td>2.981</td>
<td>0.390</td>
</tr>
<tr>
<td>41-44</td>
<td>12</td>
<td>1876</td>
<td>3.202</td>
<td>0.425</td>
</tr>
</tbody>
</table>
Talk about genuine desire seems to be present at the first age analyzed (AGE1, 18-24 months), to increase slowly by AGE2 (25-28 months of age) and then more drastically again at AGE3 (29-32 months). The frequency of genuine talk about desires seems to peak at around AGE4 (33-36 months) to drop and stabilize thereafter. In order to test for significant differences in children’s talk about desires at any given age group, a repeated measures ANOVA with age as a 6-level variable was used. There was an overall significant main effect $F(5, 55)=6.72, p < .000$. Post hoc analyses using the Bonferroni criterion for significance indicated that the average frequency of GPRDs for AGE1 ($M=0.51$, $SD=0.55$) and AGE2 ($M=0.96$, $SD=0.80$) were significantly lower than frequency of talk about GPRDs at AGE4 ($M=4.28$, $SD=2.24$) and AGE6 ($M=3.54$, $SD=2.10$).

Figure 1: GPRDs by age as a percentage of the total number of utterances for that child at that age, error bars are standard errors.

Figure 2 below shows the development of children’s communicative uses of “want” as a function of age. Except for direct repetitions, both idiomatic expressions and behavioral requests do not appear in the first age stage sampled (AGE1=18-24). By AGE2, all three categories of communicative uses are present, although not extremely different from the previous age. The frequency of DRs seems to grow and separate from the main trend at AGE3 and then again at AGE4, while at AGE3 both BRs and IEs are at the same level. Something indeed seems to happen at AGE4, when all three categories seem quite different in their frequencies, with DRs leading the frequency count, followed by BRs and IE in the last place. Both AGE5 and AGE6 seem to show the new convergence of these categories. It seems then that after AGE4, all three kinds of communicative uses of “want” stabilize. To test for significant effects, a repeated measures ANOVA with age as the withing-subject variable was carried out for each communicative use of “want”. The tests show that, taken one by one, there is no significant effects for age and each of the communicative uses of “want”, $p > .05$. However, a repeated measures ANOVA with age and communicative uses as within-subjects variables yielded a significant effect for both age $F(2.96, 32.61)^1=3.174$, $p=.038$ and communicative uses $F(2, 22)=12.708$, $p<.000$. No main effect was found for the interaction between age and communicative uses $F(3.36, 36.97)^1=0.682$, $p>.05$, n.s.

Figure 2: Communicative uses of “want” by age as a percentage of the total number of utterances for that child at that age.

Figure 3 shows quite clearly that although both GPRDs and all communicative uses start at roughly the same frequency, it is only GPRDs that increase the frequency significantly, while the other communicative uses of “want” stay roughly the same across ages. A repeated measures ANOVA with AGE as a 6-level variable (Ages 1 through 6) and communicative uses as a 4-level variable (GPRDs, BRs, IE, DRs) was used to test for differences. As expected from the previous analyses, there was an overall significant main effect of communicative uses, $F(3, 33)=59.545$, $p<.000$, a significant main effect for age, $F(5, 55)=7.444$, $p<.000$ and a significant interaction of Age and Communicative uses, $F(3, 33)=4.861$, $p<.000$.

Discussion

From the theoretical discussion above, we concluded that if children undergo some metarepresentational change of the concept of desire as a function of age, then while the developmental picture of GPRDs reëxect this change, communicative uses of “want” should stay relatively the same across ages.

The analyses carried out yield some results that point towards this direction. Although there are differences among the communicative uses themselves (that is, there are differences between DRs and IEs, for instance, at 33 months, see Figure 2), there is no main effect for

1 $F$ corrected for sphericity by Greenhouse-Geisser.
age for each of these communicative uses taken in isolation. However, there is a main effect for GPRDs between the first age analyzed (18-24 months) and AGE4 (33-36 months) and AGE6 (41-44 months). This, obviously, draws a difference between the communicative uses of “want” and genuine psychological references to desire. This difference, I am inclined to say, may be related to metarepresentational issues during the acquisition of the concept of desire. If it were not about matters metarepresentational, it would be very difficult to explain why the other communicative uses (BRs, in particular), which are morphosyntactically very similar to GPRDs, do not provide a main effect for age.

This paper is not calling into question the hypothesis that belief (and particularly false belief) taps an organism’s representational capacities, nor that belief is acquired at whatever age (probably, at 4 years of age). What is being questioned here is the working assumption that if belief taps on representation and belief is acquired at 4 years of age, then metarepresentation is acquired at 4 years of age.

The argument that uses the premise above seems to be, at least, enthymematic. It could be said that reasoning by means of belief is tapping certain kinds of metarepresentational abilities, the kinds for which many computational resources have to be in place (Wimmer, Hogreve, & Perner, 1988; Leslie, 1988; Davies & Stone, 1995). As an analogy, you can take the difference that exists between a belief attribution such as “Loreto believes there’s a blue car outside school” and “Loreto believes Namic thinks there’s a blue car outside school”. Terminological differences aside, children seem to acquire the ability to solve problems like the latter by around 6 years of age (Perner & Wimmer, 1985), two years after they have allegedly acquired the ability to metarepresent (metarepresent by false belief, that is). However, just because of this empirical fact, one would not argue that by passing this more complicated task, the child has now acquired another ability, one different from the metarepresentational abilities acquired two years earlier. The same argument holds for desires: just because children talk and reason by means of desired a year before they do so with belief, that does not mean that a new ability has been acquired. That children are able to calculate this double-embedding of the same belief concept a couple of years later than 4 years of age points in the direction of problems with some of the computational mechanisms that help children deal with metarepresentation (Fodor, 1992), but not with the ability to metarepresent itself.

General conclusions and future work

The main point of this paper is that it is extremely hard to consider reasoning by means of desire (at any age stage) as a non-metarepresentational endeavor. This hypothesis has been analyzed in two ways: a) by means of a logical analysis of what is involved during desire reasoning and communication and b) by providing some preliminary empirical evidence that even talk (as a proxy for reasoning) about desire shows a clear developmental trend when compared to communicative uses of the same words used to talk about desire (“want”, in this case).

If the main point of this paper is right, then the most pressing issue to deal with is the lag between the acquisition of the concept of belief and that of desire. In other words, if metarepresentation lies in the nature of both belief and desire but children have less difficulty understanding the representational nature of the latter while failing to understand the equivalent metarepresentational character of the former (Astington & Gopnik, 1991), then again it may be the case that something other than metarepresentation is at stake. Of course, much more work is needed in this area. Nonetheless, I would like to propose that the answer to this riddle lies in the computational mechanisms dealing with metarepresentation at the different stages. Not with metarepresentational abilities themselves. In other words, I would like to propose that the ability to metarepresent is acquired as soon as children start talking about and reliably communicating their own and other people’s mental states, starting with desire at around the 30th month of life. This is somewhat earlier than previously thought, but it would help explain and make sense of the whole philosophical tradition of belief and desires as belonging to roughly the same theoretical arena as the rest of the propositional attitudes.

Acknowledgments

This paper was written while the author was holding a postdoctoral fellowship at the Human-Oriented Technology Laboratory at Carleton University. I would like to thank, particularly, John Logan, Andrew Brook and Gitte Lindgaard for insightful comments and support.

References


