Chapter Four

Representation and Desire: Case Study in How a Philosophical Error Can Have Consequences for Empirical Research

When Premack and Woodruff in 1978 asked whether the chimpanzee had a “theory of mind,” they prompted reactions not only from psychologists, but also from philosophers. Among those philosophers who responded to Premack and Woodruff were several who outlined a research paradigm for studying the understanding of false belief in primates and children (Bennett 1978; Dennett 1978; Harman 1978). This paradigm was later taken up by Wimmer and Perner (1983) and was instrumental in launching contemporary research on the child’s understanding of mental life.

Ever since, theory-of-mind research has shown how philosophical work can productively be employed by the practioners of other disciplines. There are risks, however; if the philosophy is genuinely being used, rather than merely tacked on as an afterthought, one would expect errors in philosophy to lead to further errors down the road. In this paper, I will examine one such error in theory-of-mind research, stemming from the misuse of the word ‘representation’.

What I shall argue, in particular, is the following. In contemporary philosophy, the word ‘representation’ is used with a variety of different meanings which are not always clearly distinguished even by the philosophers who discuss them. Some of

these meanings have found their way into the literature in developmental psychology, where they have been run together, resulting in equivocal arguments, misrepresentations of existing data, and even, I will assert, ill-fated research. I will begin by distinguishing two very different ways of viewing representation, and I will examine in detail how one philosopher conflated these different understandings. I will then describe the motivation and mistakes of the developmental research that is the focus of this paper. I will conclude with some suggestions about how certain experiments on the child's view of drawing might be of help confirming or disconfirming a popular hypothesis about the child's understanding of mind.

If this paper has any single effect on the reader, I hope it is this: That it entices her to acquire the (all too rare) habit of *clarifying* what is meant when the word 'representation' is employed, rather than simply invoking the word as though it had a single, univocal meaning on which everyone agreed.

Representation is a crucial concept in philosophy of mind and cognitive psychology, and trouble with its use is bound to strike to the roots of these disciplines. What I shall describe in this paper are only the troubles I know best.

1. Desire in Indicative and Contentive Accounts of Representation

The contemporary philosophers whose accounts of representation have had the most impact on the theory-of-mind literature in developmental psychology are probably Fred Dretske, John Searle, and Jerry Fodor. Although the differences between these philosophers' views of representation are enormous, this fact is not as widely recognized as it should be. (Even Fodor doesn't seem always to recognize the degree of difference between himself and Dretske; see Fodor 1984, 1987, 1990). I will focus on just one dimension of difference here, crucial yet typically ignored, and because ignored a source of unrecognized difficulties. The difference that interests me is the difference between *contentive* and *indicative* accounts of representation. Searle and Fodor offer contentive accounts, Dretske an indicative one.

I shall call an account of representation *contentive* just in case it treats as representational anything meeting the following condition:

(A.) It has propositional (alternatively: intentional or semantic) *content*.

The sense of 'propositional content' I mean to be invoking here is that now broadly used in philosophy of language and philosophy of mind. Although the notion of propositional content is notoriously unclear, my current project does not depend on any specific way of cashing out that concept. Accounts of the sort I want to label as 'contentive' are those that treat all the

following types of things as representational: beliefs, desires, and the other so-called “propositional attitudes”; sentences and linguistic acts; pictures, maps, and certain kinds of artistic objects perhaps. John Searle (1983), Jerry Fodor (1975, 1981, 1987, 1990, 1991), and Hartry Field (1978) offer -- at least to a first approximation -- contentive accounts of representation in the sense just described. Searle argues that anything with propositional content (everything listed above) is a representation. Fodor and Field argue that some things with propositional content, like beliefs and desires, while not themselves *representations* are nonetheless *representational states.* Belief and desires are “representational states,” on this view, because they are relations *between* people and internal representations. So, for example, John’s belief that it is raining is a relation between John and an internal representation with the content that it is raining (Fodor 1981, ch. 7; Field 1978).

Indicative accounts of representation require a further condition. Not only must any representation or representational state have “content” (condition (A.)), but also:

> (B.) The content of a representation *is supposed to match up* (alternatively, in “normal” conditions matches up) with the way things are in the world. If it does not, *misrepresentation* (itself a type of representation) has occurred.

On an indicative account, a representation's "job" is to reflect the way things stand in the world. All representations, on this view, have what Searle (1983) calls a "mind-to-world" or "word-to-world" (or "representation-to-world") *direction of fit.* This is in sharp contrast to things like desires and commands, which, though contentful, function *not* to reflect the way things are but (very roughly) the way things *should* be. Desires and commands have the opposite "direction of fit" -- they succeed by bringing the world into line with them, not by bringing themselves into line with the world. (For more on direction of fit see Searle 1983; Anscombe 1957; Humberstone 1992.) Fred Dretske (1988) espouses an indicative view of representation; so, for example, although he is happy to say that desires do have intentional content, he denies that they are representational (1988, p. 127).

Conditions (A.) and (B.) are meant to be approximate, not precise. Fodor, for example, though he accepts (A.) as a good "first approximation" of his view (1987, p. xi), suggests conditions in which he thinks having content is possible without representation (1987, p. 22). Searle seems to require that *mental* representations have not only a content but also a direction of fit (either direction), and a "psychological mode" (1983, p. 12). At the same time, Searle allows for "Intentional states" whose "representative content" is not a whole proposition. So, for example, Gernot might believe *that the stove is on* or desire *that Pauline arrive promptly,* but love *Sally* (1983, p. 6-7). Although belief and desire take entire

propositions as their contents, love does not. Since direction of fit is, for Searle, defined in terms of *propositional* content, Intentional states such as love, presumably, have no direction of fit, thus failing to fulfill one of Searle's apparent requirements for mental representations. Such details, however, are beside the point for my argument, so long as indicative and contentive accounts cluster *roughly* around the criteria I have given.

An essential point of agreement between those who subscribe to indicative and those who subscribe to contentive accounts of representation is that *beliefs* are representational. If I believe that yesterday it rained two inches, then I am in a mental state whose propositional content is that yesterday it rained two inches. If I believe that Rick will someday return my copy of Christopher Marlowe then I am in a state whose propositional content is that Rick will someday return my copy of Christopher Marlowe. Beliefs surely also satisfy condition (B.). My belief about yesterday's rain is supposed to reflect the way things actually are (or were) in the world. If it does not, it is my belief (not the world) that ought to be changed. Misrepresentation has occurred.

The crucial point of disagreement between the two accounts, for my purposes, is in the treatment of *desire*. On indicative views of representation (Dretske 1988, 1995; also Millikan's "indicative representation" 1984, 1993[1]) desire is not

[1] Millikan's distinction between "indicative" and "imperative" representations lines up nicely with my distinction between indicative and contentive accounts of representation

representational. Desires are not supposed to indicate how things are; they are dispositions or urges to bring things about that may not be the case, or hopes that events will transpire in one's favor regardless of one's input. We do not say of a person who desires an ice-cream sandwich but is not eating one that she *misrepresents* herself as eating an ice-cream sandwich. But if we regarded desires as representations in the indicative sense, we would be committed to saying that, by condition (B.) of representation: The content of that desire, "that I eat an ice-cream sandwich now," does not in fact match up with the world. Surely desires may be *based* upon false beliefs or misrepresentations -- perhaps I have forgotten what ice-cream sandwiches taste like and would be disappointed upon actually tasting one -- but that does not mean the desires themselves are misrepresentations. Rather, the beliefs that inform them are. Desires, then, are not representational states for those who subscribe to indicative accounts of representation. (For more on this, see Dretske 1988 and Millikan 1993.)

On *contentive* accounts, however, desires are clear-cut, central cases of representational states. Desires, like beliefs, are "propositional attitudes" *par excellence.* If I desire that Tori watch the sunset, then I am in a state whose content is the proposition that Tori watches the sunset. If I desire an ice-

(1993, p. 98-99). On indicative accounts of representation, only what Millikan would call indicative representations are representations. On contentive accounts, both her indicative and imperative representations are regarded as representational. Most of Millikan's discussions of representation are discussions of indicative representations.

cream sandwich, then I am in a state whose content is that I eat an ice-cream sandwich. You get the idea.

Both types of account draw on certain of our pretheoretical intuitions. Indicative accounts pull heavily on the idea that there are always things out in the world that representations are supposed to be representations *of* and that if those things are portrayed inaccurately, or if there are no such things to begin with, then the representation must be a *mis*-representation. Contentive accounts depend more on recognition of the possibility of fictional or hypothetical representations -- paintings, for example, that are "representations" of unicorns or military sandboxes that are "representations" of hypothetical manuevers. These ordinary-language intuitions about representation conflict with each other: One cannot grant full credit simultaneously to the idea that all representations are meant to be portrayals of the way things are and to the idea that representations can be fictional or hypothetical.[2] Hence the divergence between the accounts.

[2] An interesting intermediate case is representations of the way things *would be*. Such representations leave room for accuracy or inaccuracy of a sort, although they are not about the way things are. So, for example, one might misrepresent a unicorn as having a second horn, or one might make inaccurate claims about how the interview would have gone had you only not spilled your coffee. This would seem to be a fertile field for further exploration in the literature on representation.

2. An Example from Philosophy

Perhaps my exposition in the previous section of the distinction between indicative and contentive accounts of representation will seem obvious. Nevertheless, people do not always make clear when using the word ‘representation’ exactly what it is they have in mind. Philosophers of mind and, to an even greater extent, cognitive psychologists tend to use the word ‘representation’ unqualifiedly, as though everyone were in perfect accord over the meaning of that term. The term is far more frequently invoked than explained. Since the word has no univocal meaning in philosophy and cognitive science, such behavior is ill-advised. Not only are indicative and contentive accounts quite different in nature, but the contentive accounts are themselves quite different -- Fodor (1975), for example, thinks representations must have a formal syntactic structure, while Searle (1983) denies this. Add aesthetically-motivated accounts of representation (e.g., Wollheim 1993) and “re-presentation” puns (the latter sometimes offered by the very same authors who give different accounts of representation when the latter is not being contrasted with presentation; Searle 1983; Dennett 1991a), and you have a recipe for disaster. Shortly I will describe the errors in developmental psychology that are the focus of this paper. In this section I warm up with a similar confusion in Dennis Stampe’s article “Toward a Causal Theory of Linguistic Representation” (1977). This article had a substantial impact on later philosophical work on the topic of

representation (especially Dretske 1988, 1995; Millikan 1984, 1993; Fodor 1984, 1987, 1990), but to my knowledge no one has noticed Stampe's conflation of an indicative with a contentive understanding of representation.

Stampe's (1977) article ambitiously takes up the task of offering a "causal theory of representation," stated as generally as possible and intended to unify the then (and still) popular causal theories of knowledge, memory, belief, evidence, perception, and reference. What all these phenomena have in common, Stampe says, is that they involve a representational "object" (1977, p. 81). Understanding representation in general should then be of use in understanding these phenomena in particular.

Since Stampe talks about representations as being the kinds of things with "contents" and "objects" in a fairly traditional sense, it seems likely that he would be willing to accept something like condition (A.) on representation as described above. But is he also willing to accept (B.), thereby making his account an indicative one? Most of the phenomena mentioned on Stampe's p. 81 (cited above) could plausibly be interpreted as having a mind-to-world direction of fit (although the case of reference is not clear). If S knows that P, believes that P, has evidence that P, remembers or perceives that P, then S's mental contents are supposed to match up in the right kind of way with the world; if they don't, misrepresentation has occurred. If these are the phenomena in which Stampe is interested, then an indicative account of representation may be appropriate.

Stampe, however, hopes to include in his analysis not only the above-mentioned phenomena but also, it becomes clear as he proceeds, intentions and desires, as well as speech acts like promises and orders (1977, p. 82, 85). These latter phenomena have a world-to-mind (or world-to-word) direction of fit and, as discussed above, absolutely are *not* representations on indicative accounts.[3] To make his commitment to including such phenomena clear, Stampe says that the causal relation he wishes to make criterial for representation "is one that holds between a set of properties $F(f_1 \ldots f_n)$ of the thing (O) represented, and a set of propositions $\Phi(\phi_1 \ldots \phi_n)$ of the representation (R)"; and, he continues, the causal route may run in a number of directions and the relation still be a "representational" one (1977, p. 85). O's having F may cause R's having Φ, as in the case of true belief, or R's having Φ may cause O's having F, as in the case of an intention acted on and thereby satisfied, or there may be some common cause for both of them. It looks, then, as though Stampe's account might be a contentive one after all. He seems happy to ascribe representational status at least roughly to the same broad range of phenomena that Searle and Fodor do. (However, since Stampe does not explicitly say that he regards all items with propositional content as representations, we cannot be certain whether Stampe might wish to add some criterion that might exclude some, such as fears or doubts.)

[3] Stampe argues in later articles that desires do have an indicative function: The desire that P is supposed to indicate that *it would be good* if P were the case (Stampe 1986, 1987). Nevertheless, since the actual content of the desire, P, is not supposed to

Having said all this, Stampe remarks that “for the sake of having a manageable form of expression” he will “indiscriminately” just speak of the object as causing the representation (1977, p. 85). This is not at all an atypical move in philosophy and the cognitive sciences: We set ourselves the task of talking about “propositional attitudes” or “intentional states” in general (i.e., belief, desire, intention, fear, doubt, etc.); for simplicity’s sake we decide to talk about just one of them in depth; the one chosen “at random” is always belief; and we end up saying very little, except perhaps as a special study, about how the other propositional attitudes or intentional states are supposed to fit into our “general” account. It is particularly striking that we should see Stampe following this pattern, given the complex and detailed treatment of desire he develops in other work (Stampe 1986, 1987, 1994). But rather than focus on this later work of Stampe’s, which does not exhibit the tendency or error in which I am interested, I want to focus on the seminal and general 1977 paper of Stampe’s, since it displays quite clearly and usefully just the kind of slippage that proves damaging in the psychological work I will be examining shortly.

If Stampe wants to talk only of the object’s causing the representation, for the sake of having a manageable form of expression, but nevertheless wants his claims to apply to cases in which the causation runs in the other direction as well, then

---

match up with the world, even on Stampe’s account of desire, desires cannot be indicative representations as I have described them, by the criteria stated on p. 4.

it should always be possible to adjust his claims to fit these other cases. If his claims cannot be so adjusted, then he will have not done what he has advertised -- he will not have presented a *general* account of representation applicable to representations running in all directions of causation. One way of thinking about this potential error is as a conflation of indicative and contentive accounts of representation. The class of representations would be viewed widely, i.e., contentively, while the properties attributed to representations in general would include properties that apply only to indicative, belief-like representations in which features of the represented object cause a representation of that object as having those features.

Before Stampe even leaves page 85, he shows signs of having made the error in question. He says, for example, that "the causal criterion requires that the relevant properties of the object represented cause the instantiation of the relevant properties in the putative representation of it" (1977, p. 85). This *may* be a reasonable criterion to apply to belief, especially if one spruces it up with an account of misrepresentation (Stampe does so in terms of "normal" or "fidelity" conditions). There may be something funny about a belief that X is F that is not causally hooked up in the right kind of way with X's being F (although even Stampe wants to modify this claim when applied to beliefs about the future). But we cannot, as I have just argued we must, generate from this description an even remotely plausible analogous condition for desire. If X is not yet F, X's

being F cannot possibly cause my desire that X be F, since what does not exist cannot be a cause. Nor can we get the results we want by turning the direction of causation around. There is *nothing* odd or wrong about a desire that X be F that does not cause it to be the case that X is F. Some desires simply are not satisfied. Other desires, about the weather for example, we may hope to be satisfied, but not as a result of a causal chain involving the desire in question. Nor is it plausible to think that there must be some common cause of both the desire that X be F and its eventually being the case that X is F. Stampe's claim that "the causal criterion requires that the relevant properties of the object represented cause the relevant properties of the putative representation of it" would not seem plausible had Stampe "indiscriminately" chosen to talk of the representation causing the state of affairs represented rather than the other way around. Stampe already appears to have slipped into treating representation indicatively, attributing to all representations properties that do not rightly apply to representations contentively understood.[4]

From this point onward, Stampe's account looks like an indicative account of representation. On page 86, he says that "the central fact about representations" is that they "provide *information* about what they represent" (my ital.). But in what sense do, for example, promises that P, orders that P, or

[4] Stampe later argues that although what a desire that P represents is P, what it represents P *as* is a state of affairs the obtaining of which would be good (1987, p. 355). The desire is then "ideally caused by the fact that it would be good were P to be the case" (1986, p. 167). This is importantly different from the desire's being caused by P

intentions that P provide any information about the state of affairs they represent, P? On pages 87-90, Stampe has a discussion of what it means to say that a representation is *accurate*. It makes no sense to turn the causal direction of these ideas around and apply his discussions to non-indicative representations.

Furthermore, Stampe says:

> There is nothing essentially mentalistic about [representation]; it *may* be a wholly physical relation. Neither is there anything essentially *semantic* about it, in the narrower (proper) sense of the term. It is the relationship that obtains between the moon and its image reflected on the surface of a pond, and it would do so were no minds ever to have existed; even if there had been nothing to count them, the number of rings in the stump of a tree represent the age of the tree (1977, p. 87, his ital.).

If representation is disconnected from the mental like this and really can run either direction for Stampe, then it ought to be just as legitimate to turn things around and say that the moon represents the reflection in the pond and that the age of the tree represents the number of rings in its stump. Stampe, I assume, doesn't want to say this -- if he did say it, he would have to abandon the idea of any good match between his usage of 'representation' and anyone else's -- but there is nothing in Stampe's account of representation that suggests that the moon *can't* be the representation of the reflection. It seems doubtful that Stampe would have made an analogous claim had he chosen to speak consistently of the representation's causing the object represented rather than the other way around. Perhaps Stampe

---

itself, as would be required on Stampe's criterion cited in the text, which requires a

would want to add conditions to the representational relation meant to apply specifically to cases in which the causer is the representation, thereby ruling out cases like the moon's representing its reflection, but in fact he discusses no such conditions.

In sum, Stampe focuses in his 1977 paper on features of representation that apply to belief-like mind-to-world cases; as a result, his account of representation looks very much like an indicative one. This may be fine for most of the phenomena he wants to discuss in this paper, but he cannot apply his account to desire, intention, or any of a number of other phenomena with a world-to-representation direction of fit that he does in fact claim to cover with his account. Although the paper begins as if it were going to offer a contentive account of representation, the account looks more indicative in the end.

Stampe is not unique among philosophers in running together indicative and contentive approaches to representation, and I have chosen his 1977 article as a focus not to single out him in particular, but rather because it is an influential and clear example, and it shows how even a philosopher like Stampe, who is generally attuned to the complexities of desire and other mind-to-world representations, can slip into a belief bias when speaking broadly about representation in general. In the airy heights of abstraction and generalization, the difference between contentive and indicative accounts of representation can sometimes go unnoticed.

---

relation between P *itself* and the desire that P.

Even Fodor, whose remarks about representation are usually clearly in the contentive camp, sometimes slips into thinking of representation indicatively. The clearest case of this is probably in his 1984: On the first page, Fodor says that "the point about propositional attitudes [belief, desire, etc.] is that they are *representational* states" (1984, p. 231, his ital.) -- i.e., they are relations between people and internal mental representations (Fodor 1981, 1991). Fodor then, as usual, focuses most of his attention on representations with a mind-to-world direction of fit. Finally, on the closing page of the article, Fodor remarks that if R represents S, "what R represents is its truth condition, and its truth condition is whatever causes its tokening in teleologically normal situations" (1984, p. 249). With the indicative/contentive distinction in hand, we can see the difficulty here. The first quotation insists that desire is a representational state, but the second does not allow desires to involve normally tokened mental representations. When Fodor later rejects the position endorsed in the second quotation, he finds it necessary to spend an entire chapter *arguing* against the claim that "Normally caused intentional states ipso facto mean whatever causes them" (1990, p. 82, 89) -- an argument he surely would have found unnecessary had he reflected sufficiently on the fact that both he and those he takes to be his opponents regard *desires* as intentional states.

3. The Error in Theory of Mind

I would now like to suggest that a number of developmental psychologists studying the child's theory of mind have also conflated contentive and indicative approaches to representation. I will focus on the work of two of the most prominent (and most philosophically-minded) researchers in the field: Alison Gopnik and Josef Perner. I will begin with textual evidence that the word 'representation' is being used sometimes contentively and sometimes indicatively by these two authors. I will then show how equivocation between the two meanings of 'representation' produces problems for their research on the child's understanding of desire.

Lynd Forguson and Alison Gopnik begin their 1988 paper with a very clear statement of a contentive account of representation:

> Accordingly, we will understand by the term *mental representation* a mental state consisting of (a.) a *representational attitude* (e.g. believing, wanting, wishing, regretting, fearing), and (b.) a *symbolic content* ... that differentiates one belief from another, one desire from another, and so on (1988, p. 228, ital. theirs).

Notice that desire is specifically included in the list of representational states, since it has "symbolic content". Nonetheless, a few pages later Forguson and Gopnik say

> However, these children do not seem to be able to distinguish between the different informational relationships that may hold between representations and reality. As we will see, they show little understanding of the principles of representational change, representational diversity, or the appearance-reality distinction.
>
> All these abilities require that the child simultaneously consider a particular representation as a representation *and* as an indicator of how the world

really stands (1988, p. 234-235, ital. theirs, underlining mine).

These latter remarks only make sense on the view that all representations have an indicative function; one does not need to understand indicator relationships to understand that *desires* may change (Forguson's and Gopnik's "representational change") or that different people may have different *desires* (Forguson's and Gopnik's "representational diversity").[5] Forguson's and Gopnik's main thesis, in fact, depends on the slide between contentive and indicative accounts. On the basis of experiments suggesting a shift between ages three and four in the child's understanding of *indicative* relationships and misrepresentation, they argue that the four-year-old but not the three-year-old understands representation in general. This claim would be warranted if Forguson and Gopnik consistently held an indicative account of representation; it is not warranted if their account of representation is a contentive one. I will shortly describe in more detail the role this error plays in Gopnik's research, but first I will examine the work of one other researcher to make the point clear and to show the prevalence of the mistake.

Josef Perner (1991a&b) also seems to conflate contentive and indicative accounts of representation. He says, for example, that the "scientifically satisfactory" way to view a person's -- Sue's -- desiring something requires that "an internal representation is posited in Sue's mind, which represents the

[5] One might argue that desires *are* indicators of how the world really stands, a desire for food, for example, indicating a need for food (something like is Stampe's later (1986) view). Even if this were true, it's hard to see how it would be necessary to understand

nonexisting situation she desires" (1991b, p. 116) and that "treating desires as mental representations becomes necessary for understanding how desires change and how they are controlled" (1991b, p. 205). Thus, he sometimes seems to treat desires as clear cases of representations. Elsewhere, however, he says that "for any representation it is possible to misrepresent" (1991b, p. 20) and

> the definition of representation should therefore contain two elements: (a.) there must be a correspondence between states of the representational medium and states of the represented world, and (b.) this correspondence must be exploited by an interpretive system so that the representation is used as a stand-in for the represented (1991a, p. 144).

Neither of these latter remarks is consistent with regarding desires as representations: It makes no sense to talk of a desire as a misrepresentation of something (though the beliefs on which a desire *depends* may be misrepresentations); desires do not correspond the way beliefs do (or are supposed to) to states of the external, represented world; desires do not (in any clear sense) function as "stand-ins" for what they are supposed to represent.

Perner later argues (contra his 1991b, p. 205, cited above) that desires are not themselves *representations,* but rather are *representational states* consisting of relations between people and representations (1995; see also Fodor 1981 and Field 1978). On this view, S's desire that P is a relation between S and an internal representation whose content is P. This account of

---

this fact about desire to understand change and diversity in desires as Forguson and Gopnik suggest.

desires as representational states is also not consistent with Perner's indicative-sounding remarks about representation cited above (1991b, p. 20; 1991a, p. 144). My desire, for instance, that I get some fresh air is not plausibly seen as a relation between me and some internal mental thing corresponding with, and possibly misrepresenting, the state of the world. If it were, we would have to say that this desire of mine involves a misrepresentation: I am *not* getting fresh air, so any mental representation with the content that I get fresh air and the task of corresponding to the world would have to be failing in its representational task. But of course there is no misrepresentation. The facts are clear: I know that want fresh air, and I know that I am not getting it.

Perner, I think, recognizes that there is a problem here and seeks to escape it by arguing that desires involve a "secondary" type of representation:

> The primary function [of a representation] is to reflect the represented environment faithfully so that the user can learn to use it as a reliable guide. This is primary because it establishes the meaning of representational elements.... But once this meaning has been established, a map of a fictional environment can be generated by combining representational elements established by the primary process. This allows representations to be positively employed to represent hypothetical, nonexisting states of the environment (1991b, p. 24-25).

Perner follows these remarks with an interesting discussion of the use of "models" (e.g., a military sandbox) for both indicative and fictional purposes. However, although these remarks do clarify his position in some ways, they don't get him out of the bind described above: Either secondary representations

are truly representational, in which case his account is contentive and he ought not regard correspondence to the world and the possibility of misrepresentation as necessary attendants of representation; or secondary representations are not genuine representations, in which case desire ought to be left off the list of representational states.

The consequences of not deciding this issue are serious, since they lead Perner to some fundamental errors -- very nearly the same errors that Forguson and Gopnik make. Perner, like Forguson and Gopnik, sees the child as shifting, between ages three and four, from a nonrepresentational to a representational understanding of mind. (The title of his 1991 book, in fact, is *Understanding the Representational Mind.*) His argument for this depends entirely on evidence for a transformation in the child's understanding of facts *unique to indicative representations* -- i.e., that beliefs may be false, that appearances may differ from reality, that photographs may fail to capture the present situation. The conclusions Perner wants to draw, however, are supposed to apply to representations *contentively* understood, including desires and other mental states with a world-to-mind direction of fit.

Gopnik and Perner both have enormous influence on research in the child's understanding of mind, and so it is interesting to see them making such a similar mistake. But this mistake might be of merely conceptual interest, had it not also led to misguided empirical research.

It does so via the following equivocal argument, which both Gopnik and Perner accept:

(1.) Children come to understand representation at four years.

(2.) Therefore, their understanding of all representational states must undergo transformation at this age.

(3.) Desire is a representational state.

(4.) Therefore, the children's understanding of desire must undergo some important transformation (presumably analogous to their transformation in belief understanding) at four years.

First, some caveats. Neither Gopnik nor Perner put the argument forward in precisely this form. Nor does Gopnik, at least, deny the possibility of some "décalage" (difference in timing) between belief understanding and desire understanding (Astington and Gopnik 1991). They also each admit that there is probably some less sophisticated, "nonrepresentational" understanding of desire available to younger children. Gopnik sees no such nonrepresentational correlate for belief (Astington and Gopnik 1991); Perner argues for the existence of such a correlate, which he calls "prelief" (Perner, Baker, and Hutton 1994; Perner 1995). Nonetheless, in the final analysis Gopnik and Perner are both clearly committed to the equivocal argument just mentioned. They explicitly include desires in their lists of representational states, and they explicitly -- prominently -- declare that the child comes to understand representational states at four years. Unless desire is to be treated as a special case, more difficult

to understand than representational states as a whole -- a view neither Gopnik nor Perner endorse and against which there seems to be good developmental evidence (see below) -- the conclusion that desire understanding should change between ages three and four follows naturally.[6]

Now the problem with this argument is, as you may have gathered, that premise (1.) is warranted only on one understanding of representation, while premise (3.) is warranted

[6] Even some who do not buy into the dominant view described here may be committed to an analogous argument. Henry Wellman, for instance, (1990; Bartsch and Wellman 1995) similarly puts desire on his list of representational states and then ignores it in his more detailed discussions of representation. Since Wellman has studied the child's understanding of desire more extensively than most and has even given it a central role in his developmental account, this fact is especially surprising. In his most abstract discussions of representation, Wellman characterizes representations contentively, as states with "internal mental content" (Bartsch and Wellman 1995, p. 14). Wellman writes,

> In adult understanding as philosophers treat it, a person's desires are typically construed as similar to beliefs. Thus, both desires and beliefs are called propositional attitudes. Beliefs are beliefs about a proposition: Joe believes that that is an apple. In this construal, beliefs are understood as representational. "Joe believes that that is an apple" means something like that Joe has a cognitive representation of the world and in that representation the designated object is an apple. A person's desires can be construed similarly, that is, as desires about propositions, about possible represented states of affairs. "Joe wants an apple," then, is understood as something like, "Joe wants that there be an apple and that he obtain it." ... Since a person's desires are also representational in this sense, it is feasible to talk of desires for not-real, nonexistent imaginary things. We say things like "Joe wants a unicorn" or "Joe wants to be the best ski jumper ever" (Wellman 1990, p. 210).

Although Wellman also emphasizes a simplified, non-representational understanding of desire he thinks is available even to two-year-olds (Wellman 1990, p. 210-211; Bartsch and Wellman 1995, p. 13-14), he clearly thinks that the adult understanding of desire is fully representational: Desires are mental states taking full propositions as their contents.

On Wellman's view, the child comes to understand representation at around three years of age (in this, Wellman deviates from the majority view). One would thus expect the child's understanding of desires to become representational like the adult's, thus enabling the child to talk of desires for "not-real, nonexistent imaginary things." In discussing the transition from a non-representational to a representational understanding of mind, however, Wellman leaves desires out of the picture altogether. He repeatedly emphasizes that there are two sorts of representation: reality-oriented representations like beliefs and fictional representations like imaginings and dreams (Wellman 1990, ch. 9). Desires do not fit into either of these categories and are not mentioned. Thus, for example, in discussing the child's understanding of representational diversity, Wellman remarks that "even three-year-olds understand representational diversity, but they understand only the diversity allowed by imaginings and by a hit-or-miss understanding of misrepresentation" (Wellman 1990, p. 255). He says this in spite of the fact that he earlier presented studies (Wellman 1990, ch. 8) that, he argued, showed that the two- or three-year-old child could understand that people can have and act on desires different from the child's own. His discussion of the acquisition of an "active, interpretive understanding" of representation at four years of age similarly ignores desire: Although the child's understanding of false belief and the appearance-reality distinction are discussed at length, no attempt is made to examine the child's understanding of the active, interpretative dimensions of desire or even to discuss what such dimensions might be.

only on the other understanding. The argument is thus an equivocal one and invalid.

Gopnik's and Perner's arguments for (1.) depend on several experiments well-known in the theory of mind literature, and which have received broad attention in both psychology and philosophy. One classic is Gopnik's and Astington's "Smarties box" experiment (1988; also Perner, Leekam, and Wimmer 1987). Children are shown the easily recognizable opaque candy container for the English confection "Smarties" and are asked what they believe is in the container. Naturally the children answer "Smarties." The container is then opened to reveal not Smarties, but a pencil. Children are then asked a series of questions, including "When you first saw the box, before we opened it, what did you think was inside it?" and (in the Wimmer and Hartl 1991 version) "What will [your friend] say is in the box?" Three-year-old children, but not four-year-old children, typically respond "pencils" to both these questions.

Leaving aside the interesting methodological and theoretical issues this experiment raises, suffice to say that it, and others like it, are generally taken to suggest that the following competencies emerge at about four years of age: (a.) an appreciation that other people may have false beliefs (Wimmer and Perner 1983; Perner, Leekam, and Wimmer 1987; Moses and Flavell 1990); (b.) an appreciation that one's own beliefs may have been false in the past (Gopnik and Astington 1988; Wimmer and Hartl 1991); and (c.) an appreciation that things may appear to be other than they are (Flavell, Flavell, and Green 1983; Flavell,

Green, and Flavell 1986; Gopnik and Astington 1988; Friend and Davis 1993). That these developments should occur at roughly the same time is not surprising: They all seem to tap a basic understanding of the possibility of misrepresentation (but see Vinden 1996 for another view); and for many researchers, indeed, the child's coming to understand misrepresentation at that age is seen as the surest sign of her coming to understand representation then (Perner 1991b; Moses and Flavell 1990; Astington 1993; Olson 1988; but see Hala, Chandler, and Fritz 1991).

The important thing to notice here is that all these experiments tap abilities associated *exclusively with indicative, mind-to-world representations.* Desires cannot be false; desires cannot be misrepresentations. This kind of evidence, then, only warrants the first step of the argument described above if 'representation' is construed indicatively. But for step (3.) to be plausible, 'representation' must be understood contentively; hence, the equivocation. The same problem may be put another way: The experiments cited show (at best) that the child comes to understand the nature of misrepresentation at around age four; but this understanding has *no* bearing on the child's understanding of desire; the evidence so far supplied provides no reason to suppose that the child's understanding of desire ought to be transformed at this age. And in fact it is not.

Gopnik performed a number of experiments aimed at discovering the expected 3-4 shift in desire.[7] Astington, Gopnik, and O'Neill (1989; reported in Astington and Gopnik 1991), for example, performed an experiment to see if children were as poor at recalling their past unsatisfied desires as they seem to be at recalling their past beliefs. (Searle (1983) regards false beliefs and unsatisfied desires as structurally similar in that they both involve unmet "conditions of satisfaction.") Children were shown two toys that looked very different but could not be distinguished by touch, and asked which toy they preferred. The toys were then dropped together into a bag and the child was allowed to withdraw only one. The child was then asked whether she got the toy she had wanted. While almost 80% of three-year-olds correctly described their unsatisfied desires, they performed no better than chance on the standard (Gopnik and Astington 1988) test for recollection of past false beliefs.

One might object that there is no good way, in this experiment, to tell that the children aren't simply reporting on their *present* desire for the toy they didn't get. In the standard false belief recollection tasks, the belief is shown to be false and thus *changed* before the child is asked to recall it. The child sees the Smarties box, and it opened to reveal a pencil; the child's belief about the contents is thereby changed. The children are then asked what they had (falsely) thought was in the container before it was opened. In Astington, Gopnik, and

[7] That this was her goal is not only evident from the experiments themselves, but also has been confirmed by personal communication.

O'Neill (1989), on the other hand, the child's desire is not necessarily changed when the unwanted toy is withdrawn, and thus reporting their present dissatisfaction would be a successful response strategy. One might argue that it is this disanalogy, and not a fundamental difference in their level of understanding desire and belief, that explains the three-year-old's good performance on the desire task and poor performance on the belief task.

Perhaps with the idea of addressing this problem, Gopnik and Slaughter (1991) actually worked to induce a change of desire in children -- for example, by presenting them with two books, allowing them to choose one, and then reading it to them so that they then desired to hear the other book. They found that three-year-olds have some difficulty with reporting their past desires in this task, but not as much difficulty as with the false belief tasks. Notice, however, that this is no longer a test of their recollection of an *unsatisfied* desire, so again the parallel to false belief is not complete.

In another experiment, Gopnik and Seager (1988; again reported in Astington and Gopnik 1991) showed children two books, a child's book and an adult's book, and asked which book an adult would choose. A slender majority (57%) of three-year-olds claimed that the adult would choose the child's book. Four- and five-year-olds, on the other hand, said this only 36% and 28% of the time, respectively. Gopnik and Seager draw a parallel between these percentages and similar percentages one sees on the

false-belief tasks. They take the experiment as evidence that young children don't understand that different people can have different desires. This conclusion, however, is contravened by the results of other studies suggesting that children do have an understanding of the diversity of desires (Flavell, Flavell, Green, and Moses 1990; Repacholi and Gopnik 1996; Bartsch and Wellman 1995), and one wonders whether the results might be an artifact of children's not having a very good idea (or all *too* good an idea?) of what kinds of *books,* specifically, adults might care to read. It is interesting to see how hard it is to get the right kind of symmetry between a false-belief task, like the Smarties task, and any kind of desire task.[8]

Perner did not as actively (or at least not as publicly) engage in experiments directed toward finding a 3-4 shift in the child's understanding of desire. One experiment he did perform suggests that three-year-olds generally understand that people are happy when they get what they want and unhappy when they

[8] Moore, Jarrold, et al. (1995) similarly try to construct a desire task parallel to the false belief task. In their task, children are placed in competition with a toy character ("Fat Cat") to complete a three-piece puzzle. Both the child and the character begin the game with a puzzle piece for the body of a frog. Each needs to acquire, next, a head piece and, finally, the eyes. In order to win pieces, players must draw cards from a pack: a white card indicates that no action is to be taken, a red card indicates that one may take a head if a head is not already possessed, and a blue card indicates that one may take the eyes if one already has a head. The children and Fat Cat draw cards, and the child earns a head, but the puppet does not. Now, presumably, the child wants a blue card so that he may complete the puzzle. At this point, the child is asked two test questions: (1.) Which color card does Fat Cat want now? and (2.) Which color card did you want last time? These questions are intended to test that the child can understand both another person's desire that is different from his own and that his own previous different desires were different. Three-year-old children are found to pass this test in approximately the same proportions that they pass false belief tests.

This experiment is no more supportive of the thesis of a 3-4 shift in understanding the representational nature of desire than are Gopnik's experiments (and Moore et al. do not regard it as supporting this thesis). First, the parallel with false belief is not complete. These are not tests of *unsatisfied* desires, and perhaps are better compared to the child's understanding that people can have different beliefs when the facts of the matter are unknown, which seems to develop earlier than their understanding of false belief and to be in place by three years (Wellman 1990). Second, the task seems sufficiently complicated that it might introduce extraneous task-specific difficulties that could mask the three-year-old's ability to understand conflicting desires (an

don't (Hadwin and Perner 1991; see also Yuill 1984; Wellman and Banerjee 1991; Wellman and Bartsch 1988; Harris et al. 1989). In fact, the bulk of studies on the child's understanding of desire have found no important shift between ages three and four. Besides the studies cited so far suggesting that by age three children understand (a.) people's diversity of desires and (b.) their emotional reactions to the satisfaction or dissatisfaction of their desires, other studies suggest that three-year-olds also understand (c.) that desires can fail to match up with the world (Lillard and Flavell 1992) and (d.) that desires prompt action to obtain the object desired (Wellman 1990; Wellman and Bartsch 1988; Bartsch and Wellman 1989). That children understand desire substantially earlier than they understand belief is also suggested by their natural speech patterns (Bartsch and Wellman 1995; Bretherton and Beeghly 1982).

Probably because of his treatment of representation, however, Perner (1991b) seems committed to discovering a 3-4 shift in the child's understanding of desire. The best he can find is the Gopnik and Seager (1988) criticized above and a couple of experiments on understanding intention (Shultz, Wells, and Sarda 1980; Astington 1991; Astington 1993 makes a case that understanding intention ought to be regarded as of a piece with understanding desire). Astington's (1991, 1993) argument that the child's understanding of intention undergoes important changes at around the same time as her understanding of belief

---

understanding suggested by Flavell et al. 1990; Repacholi and Gopnik 1996; and Bartsch and Wellman 1995).

may in fact stand up to scrutiny. Moore, Gilbert, and Sapp (1995) also find something like a 3-4 shift in the child's ability correctly to distinguish "want" from "need". Of course, a skeptic might reply that it's not surprising that *something* changes in the child's understanding of such world-to-mind states around age four; what is more surprising, perhaps, is how *little* change there is.

I would like to end this section with some positive remarks about the current potential for productive interaction between philosophers and psychologists on the topic of representation and the child's theory of mind. A view of representation that seems to be quite popular in theory-of-mind research since the failure in the early 1990's to find a convincing 3-4 shift in the understanding of desire (*pace* Astington 1993) is neither a contentive *nor* an indicative one, but something somewhere in the middle, on which beliefs, photographs, maps, and other contentive items with a mind-to-world direction of fit are regarded as representations as well as (at least some among) images, fantasies, pretenses, and dreams, but *desires* are either explicitly excluded from the list of representations or conspicuously unmentioned (Leslie 1987, 1988, 1994a&b; Lillard and Flavell 1992; Olson and Campbell 1994; and sometimes, apparently, Wellman 1990). This approach to representation has yet to be justified or spelled out in any detail. A little philosophical work might be useful in making explicit what exactly the commitments of such a view are -- and whether there

is really a coherent, workable view here at all. Influence may run in the other direction as well. If it turns out that there are important developmental symmetries between understanding mind-to-world representations and some of these other representations -- symmetries that *don't* hold between either of these types of representation and desire -- then perhaps there is a useful category here that philosophers have missed and ought to begin to incorporate in their own work on understanding the human mind.

4. Representational Art as a Test of a Hypothesis About the Child's Understanding of Mind

Those who interpret 'representation' contentively have insufficient evidence to warrant the conclusion that children come to understand representation at age four, given the breadth of the class of representations the narrowness of the evidence base, as I have argued. But what if we read 'representation' indicatively? Should we see children as coming to understand *indicative* representations at age four? In this final section I will review some of the evidence for this conclusion, and I will suggest in rough outline an experiment that may help decide the issue.

As I have remarked already, the preponderance of developmental psychologists writing on the child's theory of mind see the child as coming to understand false belief and the appearance-reality distinction at age four, or possibly a little before. Various objections have been raised against this claim (e.g., Hala, Chandler, and Fritz 1991; Fodor 1992; Leslie 1994a&b; Lewis and Osborne 1990), but I will not attempt to assess their merit here. What I would like to focus on instead is whether, even accepting these experiments at face value, we have sufficient warrant to conclude that the child at age four comes to understand indicative representation *generally*. I think that the evidence is slender at best.

The first point to note is that the claim that the child comes to understand indicative representations at age four is *broader* than the claim that the child comes to understand the

indicative nature of belief at age four. More things than beliefs have indicative content. Popular candidates include assertions, maps, models, fuel gauges, drawings, and photographs, to name a few. If the child comes to understand indicative representation *in general* at age four, and not simply something about the capacity for *minds* (or eyes) to be mistaken or tricked (what the false-belief and appearance-reality tasks seem to test), we should expect some analogous transformation in the child's understanding of at least some of these other things at around four years of age. Although Judy DeLoache and Deborah Zaitchik have performed experiments that are sometimes viewed as a test of this hypothesis, I do not believe that the data warrant a conclusion one way or another about the timing of the child's understanding of indicative representation in non-mental domains.

Judy DeLoache's work on this topic (1989a&b, 1991, 1995) has primarily been on the child's understanding of models. In her classic experiment, she showed children a full size room with various items of furniture and a scale model of the room with miniature versions of the same furniture, arranged analogously, and she pointed out the correspondences to the children. She then introduced the children to "Big Snoopy" and "Little Snoopy" who liked to do the same things: If Big Snoopy was on the chair in the big room, Little Snoopy would be on the chair in the little room, and so forth. This correspondence was demonstrated for the children several times, and they were asked to place Little Snoopy in the appropriate place, given Big Snoopy's

location. The crucial test was this: The children were shown Little Snoopy hiding somewhere in the little room, and were told Big Snoopy would hide in the same place in the other room. The children were then instructed to find Big Snoopy (and were then requested to retrieve Little Snoopy as a memory control). If a child went directly to the analogous hiding place in the full-size room, she passed the test. If she searched randomly, she failed the test. Children were able to pass the task right around their third birthday. DeLoache's conclusion: They understand that the model (indicatively) "represents" or "stands for" the room (1989b, 1995). Since the children are only 36-38 months old, this is seen as an argument against viewing the 3-4 shift as a shift in the understanding of indicative representations.

Perner (1991b) has pointed out the flaw in this reasoning: Understanding correspondence is not equivalent to understanding representation. Note, for instance, that correspondence between A and B is a symmetrical relationship, while A's representing B is an asymmetrical relationship. Adapting an example of Perner's: In the tract-home suburbs of California, all the houses in a neighborhood are generally built according to one of four or five floor-plans. If I live in one such house, and I visit my neighbor whose house is built from the same floor-plans, I know exactly where the bathroom is. The houses, like DeLoache's models, correspond, but they certainly do not represent each other. Children, then, quite conceivably could understand the

correspondence between the room and model without understanding that one *indicatively represents* the other.[9]

Deborah Zaitchik's work (1990; see also Perner and Leekam 1990 reported in Perner 1991b) on the child's understanding of photographs is often cited as evidence for the generality of the child's transformation in representational understanding at age four. Zaitchik first familiarized children with a Polaroid camera, allowing them to take a picture and letting them watch the photo come out of the camera and develop. She then performed a skit with Sesame Street characters. She laid Ernie out on a mat in the sun and had Bert take a picture of him, which was turned face down and allowed to develop without the child seeing it. While the photo was developing, Big Bird came by and sat down on the mat. The children were then asked, "In the picture, who is lying on the mat?" Four-year-olds did well on this task; three-year-olds did not. Zaitchik argues that this experiment shows that the child comes to understand pictorial representations at the same time she comes to understand false beliefs -- and thus that we can characterize the child as coming

[9] DeLoache has argued against a "mere correspondence" interpretation of her research in DeLoache and Smith (forthcoming). DeLoache's and Smith's criticism of this view does not, I believe, succeed. First, it treats the mere correspondence interpretation as asserting that the children are only detecting simple correspondences between individual objects within the model and the full-size room. This, however, the view need not take this approach: Children might still understand the complex relation between the model room, its parts, and full-size room and its parts, even without understanding that the model symbolizes or represents the full-size room (again, consider the case of the tract-homes). Thus, DeLoache's arguments that children understand fairly complex relations between the model and the full-size room does not touch the question of whether they understand that one *represents* the other. DeLoache and Smith also assert that the correspondence view cannot handle later (but still similar) experiments of DeLoache's, but they do not describe why they think this is the case, and it is far from obvious to me.

to understand the nature of indicative representations in general at around four years of age.[10]

Other interpretations of Zaitchik's results suggest themselves, however. Understanding the operations of a Polaroid camera is neither necessary nor sufficient for understanding the nature of indicative representations. That it is not necessary is obvious: People who live in cultures without cameras will not understand Polaroid photos, but it would be wild to assume that they therefore do not understand indicative representation. The child has been given only the most rudimentary instruction in how this machine works. She might think that the picture will update to portray the current state of its subjects, or she might think that the picture portrays the way things were when it was developed, as opposed to when it was taken. Nor does knowledge of the working of cameras require the knowledge of indicative representation: The child can understand the correspondence between the photograph and the state of affairs at the time the picture was taken without understanding its *representational* nature, by an argument similar to the one presented against the DeLoache studies. If the child comes to understand Polaroids at about the same time she comes to understand false belief, I see no reason to suppose this to be anything more than a coincidence. In fact, Parkin and Perner (1997) find only very small and insignificant correlations between the performance of three- to

[10] Zaitchik, however, later argues that three-year-old children do have some tentative and wavering representational understanding of false belief (Zaitchik 1991).

five-year-olds on false belief tasks and their performance on a Zaitchik-like photo task.

Setting aside Zaitchik and DeLoache, then, the evidence for or against the claim that children come at age four to understand indicative representation generally, as opposed to indicative mental representations in particular, has been quite slender. A good test of this hypothesis is needed.

Some initial questions we might consider are: When does the child come to understand that models, or model toys, or very simple maps are *supposed* to match up with the things they represent and thus can be inaccurate?[11] When does the child understand that gauges and thermometers can misregister the properties they are supposed to detect? Dretske (1988) and Perner (1991b) have rightly emphasized the understanding of misrepresentation as the *sine qua non* of understanding the normative component of indicative representation. Unless the child understands the possibility of misrepresentation, one could argue that the child is simply picking up on the correspondence between the representer and the represented, not the essential fact that the representer is supposed to match up with the represented.

Lindsay Parkin and Josef Perner (1997) have recently performed some experiments testing the ability of children to understand misrepresentation outside the domain of the mental. In these experiments, children are tested on their ability to

understand that a *sign* (an arrow) might misrepresent reality, and their performance is compared with their performance on a standard false-belief task. So, for example, a story is told in which a train can either be at an engine house or in a tunnel. The child is introduced to a sign that is supposed to point to where the train is and a driver who has seen the train. The child then observes the train move from one location (where the sign indicates and the driver has seen) to the other (where the sign does not indicate and the driver has not seen). The child is then asked (a.) where the train really is and either (b.) where the sign shows the train to be or (c.) where the driver thinks the train is. The child who answers (a.) and (b.) correctly -- i.e. says that although the train is really in the tunnel, the sign shows the train as being at the engine house -- is scored as having understood the misrepresentational capacity of signs. The child who answers (a.) and (c.) correctly is scored as understanding that beliefs can be false. Parkin and Perner not only find a 3-4 shift in the child's understanding of misrepresentation in signs, but also find a high correlation between children's performance on the sign task and their performance on the standard false-belief task, even when age and their performance on a Zaitchik-like photo task are factored out. That the false sign and the false belief tasks should be found to be equally difficult is a little surprising, since the direction the sign indicates can be read right off the sign, whereas what

[11]Liben and Downs (1989) have studied child's understanding of representation in maps. They don't find any noteworthy understanding of maps before the school years, perhaps

the driver believes cannot be read right off any of his expressions. Still, perhaps this only shows how inattentive three-year-olds are to data suggesting the existence of misrepresentation -- something also dramatically brought out by Gopnik's and Astington's (1988) data suggesting that children will not report previous false beliefs, even if those beliefs were verbally expressed only moments before.[12]

Another place in which it seems natural to look for an understanding of misrepresentation, outside the domain of the mind, is in the child's understanding of the pictures she draws. The child's first drawings tend to be simple scribbles, but by age three or four, most children begin to produce what are commonly called "representational" drawings (Golomb 1992; Winner 1982; Arnheim 1974; Freeman 1980). These drawings, often of people, have distinguishable limbs and facial features, which are verbally labelled by the child as such. Although talk of "representation" is just as common among those discussing child art as among those discussing the child's understanding of mind, there has been little effort to connect these two fields and see

---

because of domain-specific task demands.

[12] Martin Doherty and Josef Perner (1997) also have recently found evidence that children come at four years to be able to monitor the use of synonyms, and that performance on this metalinguistic (and so arguably metarepresentational) task correlates with performance on the false belief task; but a test of the ability to monitor the use of synonyms is not a test of the capacity to *misrepresent* that is characteristic of indicative representations specifically, and so is less relevant to the argument of this section than the Parkin and Perner (1997) experiments. If Doherty's and Perner's data are interpreted as showing that children come at age four to understand representation, construed contentively, then the results will have to be reconciled with other experiments seeming to show an earlier understanding of desire. Alternatively, in accord with the suggestion with which I concluded section three, it may be that there is an understanding of representation that does not include desire but does include beliefs and a number of other things that are not specifically indicative, like words.

what light they might shed on each other, even by those whose interests cross the two areas.[13]

If it is right that an indicative understanding of representation comes to the child at age four, then a transformation in the child's understanding of her artwork ought to take place at around that time. It may be no accident that theory-of-mind researchers interested in child art have tended to push for earlier competence, perhaps in light of the three-year-old's "representational" approach to art (Sullivan and Winner 1991, 1993; Freeman, Lewis, and Doherty 1991; Freeman and Lacohée 1995), but they have not to my knowledge pursued the connection in any detail.

It is possible that the three-year-old or young four-year-old who shows little sign of understanding indicative representation according to the traditional tests may create "representational" drawings yet not understand their representational nature, i.e., the fact that, if one draws Daddy, some features of the drawing ought to correspond with features of Daddy -- if Daddy has two eyes the drawing ought not to have three, on pain of being a misrepresentation of him. To my knowledge, the child's understanding of this fact about drawings has not been systematically tested.[14] Anecdotal remarks suggest that at least five-year-olds understand that drawings can be "wrong" if they

[13] Notably, Ellen Winner (Winner 1992; Sullivan and Winner 1991) and Norman Freeman (Freeman 1980; Freeman 1991 makes some abstract and very general connections; Freeman and Lacohée 1995 uses photographs and pre-fab drawings as cues in false-belief tasks but doesn't use the child's own drawings or use misrepresentational drawings). Tony Charman (Charman and Baron-Cohen 1992, 1993) is an exception, but his research has primarily been on autistic children.

don't match up in the right way with the things they depict, and a view of early school-age children as determined to get their drawings "right" is assumed in some theories of artistic development (e.g., Willats 1984; Gardner and Wolf 1987). Golomb and Winner both provide examples (though they mean to draw something different from the passages here quoted than the child's understanding of the duty of the picture to match up with reality):

> James, age 5;4, draws a tadpole man with arms extending from the head. He looks at it attentively and remarks: "Never seen hands coming from the head" (Golomb 1992, p. 55).
>
> Conversation between an adult and a five-year-old:
> Adult: "Which is prettier, a flower or a picture of a flower?"
> Child: "A flower."
> Adult: "Always?"
> Child: "Yes."
> Adult: "Why?"
> Child: "Because artists sometimes mess up" (Winner 1982, p. 112).

It might be useful, then, to see at what age it is possible to elicit such remarks from a child, at what age they begin to criticize drawings that "get it wrong" about the objects they depict. Were we to find a 3-4 shift in this domain, that would, I think, provide dramatic confirmation of the claim that children come at age four to understand indicative representations generally. Failure to find an appropriate 3-4 shift, on the other hand, would suggest that the 3-4 transition is, at best, confined to the domain of indicative mental representations.

---

[14] Annette Karmiloff-Smith's (1990) study of children's facility at intentionally distorting their drawings is a start, but it does not specifically address the children's view of their own distortions.

A few potential pitfalls should be noted. First, there is what might be called the "Picasso problem." It is hardly straightforward business to discern when an artistic representation is a misrepresentation and when it is merely a simplification, a convention, or a creative distortion. If Picasso puts both of his subject's eyes on one side of her head, do we want necessarily to say that he is *misrepresenting* his subject as having both eyes on the same side? Similarly, if the child draws a "tadpole" figure with legs and arms proceding directly from what would appear to an adult to be the head, we may not want to leap immediately to the conclusion that this is a misrepresentation and hold the child at fault for not admitting this. Although adult "stick figures" look nothing at all like people, it is simplistic to say that they are misrepresentations.

A less obvious pitfall lies in the distinction between the child's noticing a lack of correspondence and the child's noticing a genuine misrepresentation. DeLoache's tasks, described above, suggest that the child understands that one thing may correspond to another from at least the age three (earlier with photographs: DeLoache 1991), but as I argued, this ought not be viewed as tantamount to understanding representation. One must therefore be careful to sort out mere observations of a lack of correspondence from genuine criticisms of a drawing as misrepresentational. (The Golomb quote above, in fact, is ambiguous in this way.)

Yet another pitfall is suggested by the second quote above: Deviation from intention or from convention may be seen as “messing up” -- e.g. if a line goes off the page -- without being understood as misrepresentational. It therefore needs to be made clear exactly why the child criticizes any particular drawing. If the child criticizes a drawing of Daddy with three eyes, is this because the drawing doesn’t correspond as it should to Daddy’s features, or is it simply that a certain convention -- two eyes per head -- has been violated?

Avoiding all these pitfalls in coming to understand the child’s view of drawing would be no trivial task, but the rewards in understanding how the child thinks would, I believe, be enormous.

5. Conclusion

In this paper I argued that philosophical accounts of representation could be divided into two rough camps: broad or ‘contentive’ accounts on which desire is regarded as a representational state (Searle, Fodor) and narrow or ‘indicative’ accounts on which it is not (Dretske). These accounts have not always been clearly distinguished, even by philosophers instrumental in their development (Stampe, Fodor). I argued that influential researchers studying the child’s “theory of mind” (Gopnik, Perner) have conflated these two accounts and, as a result, have been lured into misguided research on the nature of desire. I concluded with a positive suggestion on how research on the child’s understanding of art might confirm or disconfirm a popular explanation of the apparent shift between ages three and four in the child’s theory of mind.