

This article appeared in *Mind & Language* 16, 425-455, 2001.

Mindreading and the Cognitive Architecture underlying Altruistic Motivation*

Shaun Nichols

Abstract

In recent attempts to characterize the cognitive mechanisms underlying altruistic motivation, one central question is the extent to which the capacity for altruism depends on the capacity for understanding other minds, or 'mindreading'. Some theorists maintain that the capacity for altruism is independent of any capacity for mindreading; others maintain that the capacity for altruism depends on fairly sophisticated mindreading skills. I argue that none of the prevailing accounts is adequate. Rather, I argue that altruistic motivation depends on a basic affective system, a 'Concern Mechanism', which requires only a minimal capacity for mindreading.

Word count: 12,996 (without footnotes); 14,774 (with footnotes)

Throughout the last century, philosophers and psychologists have tried to explain features of our moral psychology by appealing to features of our capacity for understanding other minds, or 'mindreading'. Perhaps the most widely known treatment goes back to Piaget's early work in developmental moral psychology (Piaget, 1932). Piaget and his followers placed enormous weight on the ability for perspective taking, that is, imagining oneself to have the mental states of another (e.g., Kohlberg, 1984; Selman, 1980; Damon, 1977; Rawls, 1971, chapter 8; see Flanagan, 1991 for a useful review). Over the last two decades, there has been considerable empirical and conceptual progress in research on moral psychology and in research on mindreading. The moral psychology tradition has looked at the nature and development of two basic moral capacities: the capacity for altruistic motivation (e.g. Batson, 1991; Blum, 1994; Eisenberg, 1992; Hoffman, 1991; Sober and Wilson, 1998; Zahn-Waxler and Radke-Yarrow, 1982), and the capacity for moral judgment (e.g. Blair, 1995; Goldman, 1993; Nucci, 1986; Smetana and Braeges, 1990; Turiel et al., 1987). The mindreading tradition has explored the capacity for attributing mental states to others and predicting others' behavior (e.g. Baron-Cohen et al., 1985; Bartsch and Wellman, 1995; Currie and Ravenscroft, forthcoming; Goldman, 1989; Gopnik and Wellman, 1994; Gordon, 1986; Harris, 1992; Leslie, 1994; Nichols and Stich, forthcoming; Stich and Nichols, 1992). Although each tradition has flourished, work on moral psychology and work on mindreading has been pursued largely independently. Advances in both fields put us in an excellent position to begin charting the relations between these two capacities and to develop a more detailed picture of the core architecture of moral psychology.

Within the last decade, several philosophers and cognitive psychologists have begun to suggest cognitive accounts of altruism and moral judgment (e.g. Batson, 1991; Blair, 1995; Blum, 1994; Currie, 1995; Darwall, 1998; Deigh, 1995; Goldman, 1993; Gordon, 1995; Sober

* **Acknowledgements:** I would like to thank James Blair, John Darley, Justin D'Arms, Paul Harris, Dan Jacobson, Trisha Folds-Bennett, Todd Grantham, Chris Knapp, Jaime Leiser, Ron Mallon, Elizabeth Meny, Steve Stich, and an anonymous referee for discussion and comments on earlier versions of this paper. This research was supported by NIH grant PHST32MH19975. Address for correspondence: Department of Philosophy, College of Charleston, Charleston, SC 29424.

Email: nichols@cofc.edu

and Wilson, 1998). The role of mindreading is a central issue in all these accounts. Roughly, the proposals about mindreading and moral psychology fall into two camps. Some (e.g. Blair, 1995; Sober and Wilson, 1998) maintain that basic capacities of moral psychology do not require any mindreading ability at all. Others (e.g. Batson, 1991; Goldman, 1993) maintain that basic capacities of moral psychology depend on the capacity for perspective taking. Elsewhere, I try to develop an account of the basic capacity for moral judgment (Nichols, forthcoming). In this paper, I'll focus on the recent cognitive accounts of altruistic motivation. I'll argue that, contrary to the prevailing views, altruistic motivation depends on only a minimal capacity for mindreading and also on an affective system, a 'Concern Mechanism' that generates the motivation.

The literature on altruism is simply enormous, and it spans several disciplines including philosophy, social psychology, developmental psychology, and evolutionary biology. Although I'll draw on work from all of these areas, the goal of this paper is restricted to the project of determining the cognitive mechanisms underlying basic altruistic motivation. Since numerous cognitive mechanisms play an essential role in generating altruistic behavior – e.g., perceptual input systems, attentional mechanisms, motor control systems – it will be important to be a bit more explicit about my explanatory goals. I want to sketch an account of altruistic motivation that addresses two different questions. One question asks which mechanism produces the motivational state itself. In keeping with the prevailing views, I'll argue that the motivational state is an affective state, produced by an affective system, the 'Concern Mechanism'. The other question asks which mindreading mechanisms are required to activate the affective mechanism. For the most part, I'll defer discussion of the affective component of altruistic motivation until section 7. Until then, the focus will be on the extent to which mindreading is required for altruistic motivation. I'll consider in some detail recent proposals about the mindreading mechanisms underlying altruistic motivation. I'll argue against the radical view that mindreading capacities are unnecessary for altruistic motivation. Then I'll sketch the more prevalent proposal, that altruistic motivation depends on the capacity for perspective taking. I'll maintain that none of the arguments for the perspective taking account is convincing and that there is considerable evidence that altruistic motivation does not depend on such sophisticated mindreading capacities. Rather, I'll suggest that altruistic motivation depends on a Concern Mechanism that requires only minimal mindreading capacities, e.g., the capacity to attribute distress to another. But before we get to that, I need to get clearer about the operative notion of altruistic motivation.

1. Core Cases of Altruistic Motivation

To begin a discussion of altruistic motivation, it will be useful to set out some core cases of altruistic behavior. In science in general, it's not always clear at the outset what the core cases are, and new evidence and arguments might alter our conception of what should be included as core cases. The situation is no different in studying altruism, and we may want to revise our view about what the core cases are. Philosophical discussions in this area tend to rely on hypothetical cases of altruism. But since the present goal is to give an account of the cognitive mechanisms implicated in actual cases of altruistic behavior, it is important to begin with real cases. To his credit, philosopher Lawrence Blum takes this strategy and offers real examples of helping behaviors that he suggests need to be accommodated by an adequate theory of altruism

(Blum, 1994). Blum's cases all come from young children. For present purposes, it will suffice to recount just a few of the examples:

1. Sarah at 12 months retrieves a cup for a crying friend (Blum, 1994, p. 186).
2. Michael at 15 months brings his teddy bear and security blanket to a crying friend (Blum, 1994, p. 187).
3. A two-year old accidentally harms his friend (another two-year old) who begins to cry. The first child looks concerned and offers the other child a toy (Blum, 1994, p. 187).

The clearest real-life examples of altruistic behavior in adults come from work on helping behavior in social psychology. Perhaps the best known research on adults' helping behavior is the work on the 'bystander effect' by Bibb Latane and John Darley (1968). They found that when there are numerous bystanders, subjects are relatively unlikely to offer assistance to those in need. This finding is often used to draw a rather bleak picture of human altruism (e.g. Campbell, 1999). However, focusing on these studies obscures the pervasiveness of human altruism. For it turns out that if subjects perceive unambiguously serious distress cues and there are no bystanders, virtually *everyone* helps. For instance, in one study, Clark and Wood (1974) had each subject engage in a distracter task and as the subject left the experiment, he passed a room in which a man (the experimenter's accomplice) made a sharp cry of pain and then feigned unconsciousness apparently as a result of being shocked by an electronic probe. The researchers found that when the accomplice was no longer touching any of the electronic equipment, *all* of the subjects offered help. And even when the accomplice was still touching electronic equipment (thus presenting potential danger to the helper), over 90% of the subjects offered help (Clark and Wood, 1974, p. 282). An adequate account of altruistic motivation should explain the underpinnings of these kinds of helping behaviors.

This list of core cases is, admittedly, rather short. It excludes possible cases of altruistic motivation that do not involve helping others in need. Sometimes people are generous to strangers who aren't in need, and I don't mean to suggest that such behaviors can't be altruistically motivated. However, I think that by focusing on a more limited range of cases, we are more likely to make progress on cognitive theories of altruism. The cases of comforting or helping others in distress form a plausible core because such cases emerge so early in children and they appear to be pervasive among adults. Furthermore, although I'm focusing on a very short list of core cases, these cases already present a fairly daunting task. Devising an account of altruistic motivation that would capture both the child cases and the adult cases would be a considerable advance. Of course, it's possible that the examples from children and the examples from adults cannot be captured by a single account. But all else being equal, an account of altruistic motivation that can capture both of these cases would be preferable to an account that captures only one. I'll argue that a close look at the role of mindreading in these cases will provide us with a unified account.

2. Altruism without Mindreading?

One account of the role of mindreading in altruistic behavior is to deny that mindreading plays *any* essential role in altruistic motivation. There are two versions of the view that are discussed in the recent literature. However, as we'll see, neither account fits the evidence well.

2.1 Emotional Contagion

Perhaps the most common explanation of the basis for altruistic behavior is empathy. For instance, Goldman writes, 'empathy . . . seems to be a prime mechanism that disposes us toward

altruistic behavior' (1993, p. 358). However, it is important to distinguish between two different capacities that get labeled as 'empathy'. Most generally, empathy is regarded as a 'vicarious sharing of affect' or an emotional response in which the emotion is 'congruent with the other's emotional state or situation' (Eisenberg and Strayer, 1987, pp. 3,5). There are two rather different ways that one might arrive at a 'vicarious sharing of affect'. One way is by perspective taking, i.e., imagining oneself to have the other person's mental states. I will consider an empathy-based account of altruism along these lines in section 3. A quite different way that we arrive at the same affect is by emotional contagion, when we 'catch' another's affect. Some capacity for emotional contagion is present at birth as evidenced by the fact that infants will cry when they hear the cries of another infant (Simmner, 1971). The capacity for emotional contagion thus does not require the capacity for perspective taking. Indeed, since the capacity for emotional contagion is present at birth, this capacity is presumably completely independent of mindreading capacities. There is some dispute about when mindreading capacities become available, but all sides agree that newborn babies cannot engage in mindreading.

The capacity for emotional contagion suggests a natural and simple account of altruistic motivation. If the distress of another causes oneself to feel distress, this may provide a motivation to relieve the distress of the other – it will thereby relieve one's own distress. This view has a certain elegance, but it is not easy to find a prominent advocate for the view. Although Goldman maintains that altruistic behavior is generated by empathy, Goldman also maintains that emotional contagion is not genuine empathy (1993). Indeed Goldman's simulation account of empathy is quite implausible as an account of emotional contagion (see Nichols et al., 1996), so it's unlikely that Goldman thinks that altruism derives from emotional contagion. Martin Hoffman, one of the most influential figures in empathy research, has been read as proposing something like the simple emotional contagion view in the following passage: 'Empathic distress is unpleasant and helping the victim is usually the best way to get rid of the source. One can also accomplish this by directing one's attention elsewhere and avoiding the expressive and situational cues from the victim' (Hoffman, 1981, p. 52, quoted in Batson, 1991, p. 48). It's not clear that Hoffman is really committed to this view, but it is instructive to consider the account in any case.

Notice that on the emotional contagion account of altruistic motivation, mindreading isn't essential to altruistic motivation. As noted above, emotional contagion needn't implicate mindreading processes at all. The distress cues are like bad music that you try to turn off. It requires no knowledge of electronics to be motivated to figure out how to stop the offensive stimuli coming from a stereo – one simply experiments with the various knobs and switches. Failing that, one can just leave the room. Similarly, then, one might find the cries of an infant offensive, so one might try to figure out how to stop the stimuli. To be sure, mindreading can provide useful tools for stopping the unpleasant stimuli. But on this account, mindreading needn't be essential to the motivation to stop the crying.

This story has a *prima facie* virtue – we know that this capacity is well within the abilities of young children who provide some of our core cases of altruistic motivation. So, the emotional contagion account provides an extremely simple explanation of altruistic motivation, and we know that children have the capacity for emotional contagion. Hence, it would seem that our problem is solved. Altruistic motivation doesn't depend on mindreading at all. Rather, it depends on the rather primitive capacity for emotional contagion.

Things are not so simple, however. For consider that, at least in the core cases of altruism from adults, one way to rid oneself of the unpleasant cues is to *leave the situation*. But this is not what happened in the core cases noted above. Although the subjects could have

eliminated contagious distress by fleeing the situation, almost none of them did so (Clark and Word, 1974). The fact that adults often help when they could perfectly well escape has now been extensively explored in the work of C. Daniel Batson and his colleagues (Batson et al., 1981; Batson et al., 1983; Batson, 1990, 1991). This research provides powerful evidence that some core cases of altruistic motivation cannot be accommodated by the simple emotional contagion account.¹

Batson has the broader agenda of defending a perspective-taking account of altruism, which we will consider in section 3, but for present purposes, it will suffice to see how his data undermine the emotional contagion account. In classic social psychological fashion, Batson and his colleagues set up a mock shock methodology. Subjects were told that they would be in a study with another person and that one of them would be picked at random to be the worker and the other would be the observer. The worker would perform tasks while being given electric shock at irregular intervals, and the observer would watch the person performing the task under these aversive conditions. Of course, the real subjects always ended up in the observer condition, and the ‘worker’ was really a confederate. The subjects were then told that they would view the ‘worker’ via closed-circuit television (though it was really a videotape). The experiment manipulated the *ease of escape* for the subjects. In the easy-escape condition, subjects read ‘Although the worker will be completing between two and ten trials, it will be necessary for you to observe only the first two’; in the difficult escape condition, subjects read ‘The worker will be completing between two and ten trials, all of which you will observe’ (Batson, 1991, p. 114). The subjects subsequently viewed the worker endure two trials (of the ten trials that the worker had agreed to) in which the worker exhibited considerable discomfort. Subjects were given the opportunity to help out the worker by taking over some of her trials. Using this framework, Batson and colleagues also manipulated the degree of ‘empathy’ in the subjects (see section 5 for details). Across a wide range of studies, they found that subjects in low empathy conditions were much less likely to help when escape was easy. By contrast, subjects in the high empathy condition were equally likely to help whether it was easy to escape or not.

For our purposes, the crucial point is the following. On the emotional contagion model, one should only help when it’s easier to help than it is to escape. However, evidence from Batson and his colleagues suggests that there is an important kind of altruistic motivation that can’t be satisfied by escaping the situation. Hence, this kind of motivation can’t be captured by the emotional contagion model (see also Batson et al., 1981; Batson et al., 1983; Miller, Eisenberg, Fabes, and Shell, 1996, Eisenberg and Fabes, 1990). More generally, largely as a result of Batson’s work, it is now clear that an adequate account of altruistic motivation needs to accommodate the fact that in core cases of altruism, people often prefer to help even when it’s easy to escape.²

¹Actually, Batson addresses a somewhat broader category, the ‘aversive-arousal reduction’ model of altruism, according to which ‘becoming empathically aroused by witnessing someone in need is aversive and evokes motivation to reduce this aversive arousal’ (Batson, 1991, p. 109). Emotional contagion provides perhaps the most obvious mechanism for producing aversive arousal, and the emotional contagion model outlined here is a version of the aversive-arousal model.

²Darwall (1998) uses Batson’s evidence to shore up the view that sympathy presents a categorical justification for preventing another’s woe. Sympathy, according to Darwall, connects us to ‘person-neutral value’:

2.2. Sober and Wilson on Altruistic ‘Sympathy’

In Sober and Wilson’s recent book (1998), they propose an alternative path to altruism that doesn’t rely on mindreading or emotional contagion, but rather on a certain kind of sympathy. They suggest that both sympathy and empathy may motivate altruistic behavior (e.g., 1998, p. 232). They then try to distinguish sympathy from empathy in two ways.

First, Sober and Wilson maintain that there is a crucial difference between empathy and sympathy because in sympathy,

your heart can go out to someone without your experiencing anything like a similar emotion. This is clearest when people react to the situations of individuals who are not experiencing emotions at all. Suppose Walter discovers that Wendy is being deceived by her sexually promiscuous husband. Walter may sympathize with Wendy, but this is not because Wendy feels hurt and betrayed. Wendy feels nothing of the kind, because she is not aware of her husband’s behavior. It might be replied that Walter’s sympathy is based on his imaginative rehearsal of how Wendy would feel if she were to discover her husband’s infidelity. Perhaps so – but the fact remains that Walter and Wendy do not feel the same (or similar) emotions. Walter sympathizes; he does not empathize (1998, pp. 234-5).

But of course, this example does not really distinguish sympathy from empathy. As Sober and Wilson seem to anticipate, a sophisticated empathy account can easily accommodate their case by claiming that we use our imagination to empathize with what Wendy would feel if she were to discover the infidelity. Hence, as far as this example is concerned, ‘sympathy’ might merely be a special form of empathy.

The second, and more important, feature of their account is their claim that ‘sympathy’ doesn’t require mindreading. Sober and Wilson maintain that empathy requires that one be a psychologist, but that sympathy does not:

Empathy entails a belief about the emotions experienced by another person. Empathic individuals are “psychologists”...; they have beliefs about the mental states of others. Sympathy does not require this. You can sympathize with someone just by being moved

sympathetic concern presents itself as of, not just some harm or disvalue *to* another person, but also the *neutral disvalue* of this personal harm owing to the value of the person himself. In feeling sympathy for the child, we perceive the impending disaster as not just terrible for him, but as neutrally bad in a way that gives anyone a reason to prevent it. We experience the child’s plight as mattering categorically because we experience the child as mattering... sympathy’s emotional presentation is *as of* the neutral disvalue of another’s woe, and hence, as of a categorical justification for preventing it (p. 275).

However, whether sympathy presents itself this way is not shown by any of Batson’s data. What the social psychological evidence (and commonsense) suggests is that our emotional response to a child’s plight prompts us to help the child rather than flee. So *the person in need* does play a crucial role here. But one would need quite different evidence to show that people think that a child’s plight gives *anyone* reason to help the child. It’s important to distinguish here between what people *expect* others to do and what people think is *categorically* justified. It’s probably true that most people *expect* others to be moved by the plight of children (as we tend to expect others to share many of our attitudes [Nichols and Stich forthcoming]); however, it’s not clear that most people think that a child’s suffering matters categorically. As far as I know, there is simply no evidence on this issue, and it’s quite possible that there is considerable variation on the issue across individuals and across cultures.

by their objective situation; you need not consider their subjective state. Sympathetic individuals have minds, of course; but it is not part of our definition that sympathetic individuals must be psychologists (1998, p. 236).

Thus, Sober and Wilson apparently maintain that 'sympathy' does not require any capacity for mindreading.

Of course, Sober and Wilson are welcome to define a notion of 'sympathy' on which mindreading is not required for sympathy. However, they provide no evidence that this kind of sympathy exists. If we rely on traditional signs of sympathy, the evidence suggests that children only begin to exhibit the characteristic signs of sympathy after the first birthday (see section 7) and at this age, children probably have some rudimentary mindreading skills (see, e.g., Gergely et al., 1995; Woodward, 1998). So, it may well turn out that the capacity for sympathy exists only in creatures that have mindreading capacities and that the capacity for sympathy depends crucially on the capacity for mindreading. Furthermore, even if Sober and Wilson's 'sympathy' does exist, they provide no reason to think that it explains anything like the core cases of altruism with which we began. Again, as we'll see, children only begin exhibiting comforting behaviors after the first birthday, by which time they probably have some rudimentary mindreading skills. So, if we take Sober and Wilson's suggestion as an empirical claim about the cognitive underpinnings of core cases of altruistic motivation, it is utterly unsupported.

In sum, then, neither emotional contagion nor Sober and Wilson's sympathy provides a promising explanation of altruistic motivation. It's particularly clear that neither proposal offers a unified account of the core cases of altruistic motivation with which we began. Hence, if we are to have a model of altruistic motivation that can accommodate our core cases, it cannot be one of these models that rejects outright the role of mindreading.

3. Perspective Taking Accounts of Altruistic Motivation

In the Piaget-Kohlberg tradition, the capacity for perspective taking is thought to be essential to a wide range of moral capacities, including altruistic behavior. Unlike the no-mindreading accounts of altruistic motivation, there is no shortage of advocates for the perspective taking account of mindreading and altruism. In the recent literature, the most prevalent account of mindreading and altruism continues to be that altruistic motivation depends on perspective taking. This view is suggested by several figures including Batson (1991), Blum (1994), Darwall (1998) and Goldman (1993).

Goldman (1992; 1993) is by far the most explicit about the cognitive architecture underlying perspective taking, so his work provides a useful starting point. As we've seen, Goldman maintains that empathy is central to altruism, and he maintains that genuine cases of empathy depend on perspective taking. His account of perspective taking draws on his earlier work on the off-line simulation account of folk psychology (Goldman, 1989; see also Gordon, 1986). Goldman maintains that the process of perspective taking is subserved by off-line simulation in the following way:

Paradigm cases of empathy... consist first of taking the perspective of another person, that is, imaginatively assuming one or more of the other person's mental states.... The initial "pretend" states are then operated upon (automatically) by psychological processes, which generate further states that (in favorable cases) are similar to, or homologous to, the target person's states. In central cases of empathy the output states are affective or emotional states (1993, p. 351).

Now, if we try to incorporate this account of empathy into an account of altruistic motivation, we get the following account of the processes underlying altruistic motivation when the agent sees another in distress. First, the agent determines the beliefs and desires of the person in distress. Then the agent pretends to have those beliefs and desires. These pretend-states are then operated on automatically, leading to affective states that are similar to the target's state, i.e., distress. These unpleasant affective states then motivate the agent to eliminate the problem at its source, viz., the other person's distress.

Batson's picture is less architecturally explicit, but is still clearly dependent on perspective taking. Batson claims that altruistic motivation derives from 'empathy' (1991, p. 83), and as Batson defines it, empathy requires perspective taking. He writes, 'Perception of the other as in need and perspective taking are both necessary for empathy to occur at all' (1991, p. 85). The empathic response to perceived need 'is a result of the perceiver adopting the perspective of the person in need' (1991, p. 83) and this involves 'imagining how that person is affected by his or her situation' (1991, p. 83).

Blum's (1994) view is somewhat more difficult to interpret. He maintains that altruistic behavior, or 'responsiveness' requires 'that the child understand the other child's state' (1994, p. 197). He rejects the idea that this understanding is limited to cases in which the subject infers 'the other's state of mind from a feeling the subject herself has, or has had, in similar circumstances' (1994, p. 192). Blum rejects this account because it is too 'egocentered' (1994, p. 193), and he argues that this can't be the sole cognitive process because 'such inference would not account for understanding states of mind different from those one is experiencing or has experienced oneself' (1994, p. 192). Rather, Blum maintains that 'understanding others means understanding them precisely *as other* than oneself – as having feelings and thoughts that might be different from what oneself would feel in the same situation' (1994, p. 193). So Blum apparently maintains that altruistic motivation depends on the understanding of others as potentially having different beliefs, desires, and emotions. But he doesn't offer an explicit explanation about how this understanding is achieved.

Although these accounts have important differences, they all share an assumption that altruistic motivation depends on some fairly sophisticated mindreading capacities. First, on Blum's account, and possibly Batson and Goldman's as well, the subject must be able to recognize that the other person might have different mental states than the subject herself would have in a similar situation. Second, for Goldman and Batson, perspective taking requires using the imagination to figure out someone else's mental states. As a result, in sharp contrast to the emotional contagion account, the perspective taking accounts of altruistic motivation invoke quite complex mindreading capacities.

4. A Minimal Mindreading Account of Altruistic Motivation

The accounts of altruistic motivation that make no appeal to mindreading have difficulty accommodating the psychological evidence and capturing the core cases of altruistic motivation. However, I think that we can accommodate the data with a much more austere proposal about the role of mindreading than the perspective taking accounts. I want to sketch an account of altruistic motivation that draws on as little mindreading as necessary to accommodate the core cases of altruism. Then in the subsequent sections, we'll consider the relative merits of the minimalist account and the perspective taking account.

The crucial finding in the core cases of altruism from social psychology is the fact that people often help even when it would be easy to escape (e.g., Batson, 1991). If the motivation is

caused strictly by an aversive response to immediate situational cues, as proposed by the simple emotional contagion model, then escape is a good alternative. For one can simply remove oneself from the source of discomfort. However, escape is not an adequate alternative if the motivation comes from an enduring *internal* cause. As a result, a natural first move is to suppose that subjects elect to help rather than escape because some aspect of the situation is preserved in an enduring mental representation, and this mental representation produces the motivation. One could conceivably try to use this move to extend the emotional contagion account. An emotional contagion theorist might continue to deny any role for mindreading and maintain that altruistic motivation comes from an enduring representation of the behavioral, acoustic, or physiognomic cues that cause emotional contagion. On this modified emotional contagion account, the reason subjects don't escape in the experiments is that the motivation comes not simply from the immediate situational cues, but also from the enduring representation of those cues. So, on this story, the subjects help because the emotional contagion can only be alleviated by eliminating the aversive cues. However, even this extended emotional contagion account is still quite inadequate. The problem is that superficial cues can produce emotional contagion, and if one knows that the cues leading to emotional contagion are merely superficial, this typically does not prevent one from experiencing emotional contagion, but it does undermine altruistic motivation.³ In the present context, the best way to see the problem is by considering what the account predicts about behavior in Batson-style scenarios with superficial distress cues. The extended emotional contagion account predicts that in these situations, subjects will be motivated to eliminate superficial distress cues rather than escape, and, while the relevant experiments haven't been conducted, this prediction seems most implausible. For instance, if a subject found herself in an empty classroom with a projector showing a computer-generated hologram of a baby crying convulsively, this stimulus would likely produce a negative affective response, and presumably the subject would have enduring representations of the cues that lead to this negative response. But in this case, so long as the subject realizes that the stimulus is a hologram and not a real crying baby, her aversive response will likely be relieved about equally well whether she turns off the projector or leaves the room. As a result, the extended emotional contagion account does not capture when escape will be an adequate solution for the subject.

Rather than opt for this implausible attempt to rescue the emotional contagion view, I think that we need to appeal to some capacity for mindreading to obtain an adequate account of altruistic motivation. A rough first proposal here is that altruistic motivation depends, not on a representation of superficial cues, but on a representation of the target's pain (or some other negative affective or hedonic mental state). Appealing to these kinds of representations will provide at least a partial explanation for why subjects help rather than escape in Batson-style scenarios. If altruistic motivation is triggered by a representation that the target is in pain, escape isn't an effective solution to the motivational problem since merely escaping the perceptual cues of pain won't eliminate the consequences of the enduring representation that another is in pain. Thus, this account provides some explanation for why escaping the situation

³ The notion of emotional contagion is defined as a 'vicarious sharing of affect'. As a result, the definition technically precludes merely superficial cues from producing emotional contagion, since one cannot *share* affect with something that doesn't have that affect. But obviously the mechanisms underlying emotional contagion can be activated by merely superficial cues, since one can synthetically produce many of the cues that lead to 'real' emotional contagion. Since the issue at hand is what mechanisms are in play, it is appropriate to use a broader category of response that also includes instances of 'ersatz' emotional contagion in which the affect is produced by merely superficial cues.

is not an adequate solution. Further, the account explains why the extended emotional contagion account is inadequate – if, in a Batson-scenario, you know that the aversive cues are merely superficial, then you don't have a representation that the target is in pain, so escape is an adequate solution. This account also fits well with Batson's finding that the motivation to help is relieved when the subject comes to think that the target's pain has been alleviated, even if the target's pain was alleviated by someone other than the subject (Batson, 1991).

I suggest, then, that altruistic motivation depends on the minimal mindreading capacity to attribute negative affective or hedonic states to others. On this view, a person can have the capacity for altruistic motivation even if the person doesn't have or doesn't exploit the capacity for imagining himself in the other's place and having different beliefs, desires or emotions than he himself would have in that situation. However, a person cannot have the capacity for altruistic motivation without some capacity to attribute negative affective or hedonic states to another. For the remainder of the paper, I will focus on distress as the exemplar mental state, but this is merely for ease of exposition. I don't mean to exclude the possibility that representations of other negative affective and hedonic states (e.g., grief, fear, sorrow) will produce altruistic motivation.

Appealing to the capacity to attribute distress helps explain why subjects are motivated to help even when they could more easily escape. Thus, the account seems, at least at this point, to accommodate the important cases promoted by social psychologists. However, I have thus far neglected to consider whether the account fits with the other class of core cases – comforting behaviors in young children. Is there reason to think that young children attribute distress? And are such attributions plausibly connected with their comforting behaviors? The answer to both questions is Yes. Henry Wellman and colleagues have explored emotion and pain attribution in the spontaneous speech of young children, using transcripts of children's speech from the CHILDES (Children's Language Data Exchange System) database (MacWhinney and Snow, 1990). Though this database was initially established to study children's language it has been an extremely valuable resource for studying the young child's understanding of the mind (see especially Bartsch and Wellman, 1995). Wellman and colleagues examined the spontaneous speech of 5 children, focusing on the transcripts collected for each child from the age of 2 until the age of 5. The researchers found that already at the age of 2, the children frequently make attributions of pain, usually using the word 'hurt' (Wellman et al., 1995, 130). Furthermore, in the cases analyzed by Bartsch and Wellman (1995), there are transcripts available for 4 children before the age of 2, and in each of these cases, the child is attributing pain well before the 2nd birthday (Sachs, 1983; Bloom, 1970; Bloom, 1973). So pain attribution apparently emerges very early indeed.

Not only do very young children make pain attributions, but in the work on comforting behavior, we find that young children respond to a variety of distress cues, and they direct their comforting behavior in ways that are appropriate to the target's distress. As we saw in the examples from Blum (1994), children exhibit comforting behavior in response to another's crying. In experimental studies on one-year olds, crying also elicited comforting behaviors; so did coughing and gagging (Zahn-Waxler and Radke-Yarrow, 1982, 116); and Zahn-Waxler and colleagues (1992a) found that children exhibited comforting behaviors in response to the target bumping her head, saying 'ow' and rubbing the injured part. Furthermore, in these studies, the children often comfort the target in appropriate ways. Zahn-Waxler and Radke-Yarrow (1982) conducted a longitudinal study in which a group of 15-month olds and 20-month olds were each studied for a 9 month period. The researchers report that during this period, every single child in these groups exhibited an instance of 'prosocial actions that focus on the specific distress cue'

(p. 124). For example, they describe one instance in which the mother of a 19 month-old child hurts her foot and the child witnesses the event. The child exhibited concern, ran over, said ‘hurt foot’ and rubbed the mother’s hurt foot (p. 124). In addition to showing that young children direct their comforting behaviors in appropriate ways, this example also indicates that young children actually make pain attributions in conjunction with their comforting behavior, and they seem to recognize what the target is distressed about.⁴ Thus, there is good reason to think that the minimal mindreading account I’ve proposed to explain the core cases of altruistic motivation in adults can also be extended to explain the comforting behaviors of young children.

As noted in section 1, this account is not intended to capture all instances of what we would consider altruistic motivation. We can be motivated to be altruistic to someone without attributing any negative affective states to them. For a dramatic example, we might be motivated to prevent the painless death of a peacefully sleeping stranger. However, one of the aims here is to develop an account of altruistic motivation that does not exceed the cognitive abilities of young children who exhibit comforting behavior. And although children display comforting behavior before the age of 2, they don’t have an understanding of death until much later (see, e.g., Carey, 1985). So if we try to develop an account of altruistic motivation that will capture cases like preventing painless death, the account might no longer be able to accommodate young children. As a result, I think that a promising initial strategy in developing a cognitive account of altruistic motivation is to focus on cases of altruistic motivation that are clearly within the repertoire of young children. This will leave open a number of interesting issues about the relation between ‘early’ altruistic motivation and ‘mature’ altruistic motivation. One possibility is that mature altruistic motivation develops out of the core system that I’m attempting to sketch in this paper. Another possibility is that there are independent systems subserving what we

⁴ One interesting question for future research concerns the extent to which altruistic motivation depends on the child’s appreciation that distress is an intentional mental state, a state that is (or can be) directed towards some object. For example, is the altruistic motivation mechanism (the ‘Concern Mechanism’ to be discussed below) activated by the attribution that Mommy is *sad that she hit her foot*? The evidence from Radke-Yarrow and Zahn-Waxler (1982) suggests that children often do appreciate what the target’s distress is about, and I mean for the minimalist account to be consistent with the possibility that the motivational system can be activated by attributions of distress as an intentional state. However, it is possible that one might try to develop an even more minimalist account on which altruistic motivation is activated by a general attribution that the target is in distress, with no specification of what the distress is about. (I am indebted to Paul Harris for raising this issue.)

Much of the available evidence seems to be compatible with both of these minimalist alternatives. Consider, for instance, the case from Blum in which the child retrieves a cup for a crying child. One possibility is that the child’s belief that the target is *sad that she lost her cup* activates the altruistic motivation system which then produces the motivation to relieve the target’s lost-cup-distress. An ultra-minimalist might maintain rather that the child’s motivation comes from the general attribution that the target is in distress. That is, the attribution that the target is sad (*simpliciter*) activates the motivation system which produces the motivation to relieve the target’s sadness. The child then uses other resources to determine a course of action for relieving the target’s sadness, and these resources might include the fuller intentional attribution of the target’s specific distress states. It is an open empirical matter which of these stories is right about the requisite mindreading underlying altruistic motivation, and of course it’s quite possible that the altruistic motivation system can process both kinds of distress attribution.

commonly group together as mature altruistic motivation, and the early emerging core system is just one of these independent systems. I won't try to address those issues here. There is, however, another fundamental way in which this account is only a partial account of altruistic motivation. Like the perspective-taking account, the minimal mindreading account doesn't yet explain the process that goes from mindreading to motivation. As will be discussed below (section 7), on both the perspective-taking account and the minimalist account, a natural assumption is that the representations generated by mindreading produce an affective response that motivates the agent to behavior altruistically. But first, we need to consider the relative merits of the minimal mindreading account and the perspective taking account.

5. Arguments for Perspective Taking: Batson's Evidence

Now that the two proposals are on the table, we can consider the arguments for each account. Although it's widely thought that altruistic motivation depends on perspective taking, it's not easy to find an argument for the view. The only systematic argument comes from Batson's data. Batson used various methods to manipulate the 'empathy' of subjects, creating conditions in which subjects would have either high empathy or low empathy. Batson is less architecturally explicit than one would like. But according to Batson, his evidence indicates that perspective taking is required for altruistic motivation since in the experiments high empathy subjects were much more likely than low empathy subjects to help in easy-escape conditions (e.g., Batson, 1991, p. 87; see also Darwall, 1998, p. 273). Batson's data do, I think, provide an important source of evidence against emotional contagion accounts, but they fall far short of establishing that perspective taking is required for altruistic motivation.

To begin, it's important to note that Batson's experiments cannot be decisive evidence for the perspective taking account. For the evidence does not show that altruistic motivation is absent among those with low empathy. A substantial minority of subjects in the low empathy conditions do help – averaging across studies, nearly a third of the low empathy subjects helped (Batson, 1991, chap. 8). And it's quite possible that most of the other low empathy subjects had some altruistic motivation, but not enough to outweigh the competing motivation to avoid the pain of electric shock. Submitting to painful electric shock to relieve a stranger is a rather high cost action, and it seems likely that if the altruistic option were low cost (e.g., returning an elderly person's books to the campus library), then the difference between high empathy and low empathy subjects might largely disappear.

Although Batson's evidence hardly counts as a decisive argument for the perspective taking account, it does seem that the perspective taking account provides a natural explanation for why high empathy would lead to higher altruistic motivation. For if altruistic motivation depends on taking the perspective of others, then increased perspective taking might increase the motivation. However, I think that the minimalist account can provide equally good explanations for Batson's findings. To see why, we need to consider in a bit more detail Batson's two central empathy manipulations: the perspective-taking manipulation (Batson, 1991, p. 120) and the similarity manipulation (Batson, 1991, p. 114). In the perspective-taking manipulation, subjects watched a videotape of a student with broken legs. The subjects were either told to 'attend carefully to the information presented on the tape' or to 'imagine how the person interviewed felt about what happened'. Subjects who were told to imagine the other's feelings were more likely than subjects in the other group to help in the easy-escape condition. Although the perspective taking account can explain these results, the minimalist account can explain the results equally well. For in the high perspective-taking conditions, subjects are more likely to focus on the

other's distress, and they are more likely to develop elaborate representations of the other's distress. Thus, on the minimalist account, it is hardly surprising that the perspective-taking manipulation facilitates altruistic motivation, since perspective taking implicates representations of the other's distress. In principle, it will be hard to undermine a minimalist account using this kind of manipulation since if you increase a subject's perspective taking of a distressed target, you will also increase the subject's representations of the target's distress.

In Batson's other important 'empathy' manipulation, subjects were shown a questionnaire purportedly filled out by the person who would later need help. One group of subjects saw questionnaires that expressed similar views to those expressed on the subject's own questionnaires. The other group saw questionnaires that expressed dissimilar views. Batson and colleagues found that subjects in their high-similarity group were more likely than subjects in the low-similarity group to help in the easy-escape condition. Batson notes that previous research by Stotland (1969) and Krebs (1975) shows that subjects in high-similarity conditions show increased empathy. But there is a crucial hedge on 'empathy' here. What Stotland (1969) and Krebs (1975) found was that subjects in high-similarity conditions showed heightened physiological response and expressed more concern for the other person. The level of *perspective taking* in these tasks was not measured. Nor do the researchers suggest that perspective taking is the crucial mechanism underlying the response of subjects in high-similarity conditions. There is, in fact, a large literature in social psychology suggesting that subjects are more *attracted* to people they think have similar attitudes (e.g., Newcombe, 1961; Byrne, 1971), and even that people are *repulsed* by those that they think have different attitudes (Rosenbaum, 1986). In light of this, it's hard to see how Batson's similarity manipulation could support the perspective taking account. What his findings do show is that we are more likely to help people who we think have similar attitudes (for a disturbing variation on this, see Tajfel, 1981). Coupled with the data on similarity and attraction, we might conclude from this that we are more prone to help people that we like. That's hardly surprising. More importantly, though, it is quite irrelevant to whether altruistic motivation requires perspective taking.

6. Developmental Evidence and Perspective Taking

Thus far, we have no reason to think that altruistic motivation depends on the kind of sophisticated mindreading suggested by perspective taking accounts. In this section, I'll argue that the empirical evidence actually weighs against the perspective taking account. In trying to determine the core architecture underlying a capacity, contemporary cognitive scientists pay close attention to two sources of evidence: evidence from development and evidence from psychopathologies. These sources give us a glimpse into which capacities might be independent from one another and which capacities seem to be inextricably linked. I will argue that evidence from development indicates that altruistic motivation is independent of sophisticated mindreading abilities like perspective taking. In section 8, I will take up evidence from psychopathologies and argue for a similar conclusion.

The discussion of altruism began with Blum's cases of altruism in young children. Nor are his examples atypical. Blum draws some of his examples from a large body of literature in developmental psychology. This research claims that we start seeing the kind of behavior exemplified in Blum's cases early in the second year. Radke-Yarrow and colleagues (1983) found that at 10-12 months, children didn't respond like the kids in Blum's examples, but 'Over the next six to eight months the behavior changed. General agitation began to wane, concerned attention remained prominent, and positive initiations to others in distress began to appear'

(Radke-Yarrow et al., 1983, p. 481). And, as noted earlier, in Zahn-Waxler and Radke-Yarrow's (1982) study, they found that every single one of their young subjects performed a prosocial act directed at a specific distress cue.

Despite this impressive capacity for altruistic motivation, children under the age of two have severely limited mindreading abilities. Of particular significance, young children have severe deficiencies in their capacity to take the perspective of others. Before the age of 3 years, children are apparently incapable of recognizing that someone else might have a different belief than they do. The most famous result here is the young child's failure on the false belief task. In the classic version of this task, Wimmer and Perner (1983) had children watch a puppet show in which the puppet protagonist, Maxi, puts chocolate in a box and goes out to play. While Maxi is out, his puppet mother moves the chocolate to the cupboard. The children are asked where Maxi will look for the chocolate. Children under the age of 4 fail this and similar tasks (see also Wellman, 1990; Bartsch and Wellman, 1995). Furthermore, although children begin to pretend by around 18 months, they seem unable to use the imagination to understand other minds until much later (see, e.g. Nichols and Stich, 2000, forthcoming). Thus, since toddlers provide core cases of altruistic motivation and they lack the requisite perspective taking capacities, this provides a serious *prima facie* argument against the perspective taking accounts.⁵

In fact, young children's comforting behaviors offer a striking picture of both altruistic motivation and limited perspective taking. The comforting behaviors of young children tend to be 'egocentric'. Hoffman notes that young children's helping behaviors 'consist chiefly of giving the other person what they themselves find most comforting' (1982, p. 287). For example, young children will offer their own blanket to a person in distress. Hoffman offers an example of a 13-month old who 'responded with a distressed look to an adult who looked sad and then offered the adult his beloved doll' (1982, p. 287; see also Zahn-Waxler and Radke-Yarrow, 1982; Dunn, 1988, p. 97). Thus, toddlers' comforting behavior seems to be simultaneously altruistic in motivation and egocentric in perspective.

While much early altruistic behavior is guided by 'egocentric' considerations, this is perfectly compatible with the minimalist account. A common interpretation of the fact that toddlers offer their own comfort objects is that it shows that children don't really understand that it is the other person who is in distress. For instance, Hoffman (1982) claims that the fact that children tend to give their own comfort objects to help others indicates that 'Children cannot yet fully distinguish between their own and the other person's inner states... and are apt to confuse them with their own' (1982, p. 287). However, the examples of 'egocentric' comforting responses provide no reason to think that the child fails to distinguish her own states from the states of others. On the contrary, these responses provide evidence that the child recognizes that the other is in distress. After all, the child is offering the comfort object to the *other* person. Further, the fact that the child offers a comfort object suggests that the child does understand that *distress* is involved. Children don't try to relieve the other's distress by completely bizarre behavior like pretending that the banana is a telephone. And there's no reason to think that before 18 months, the child experimented with various means of eliminating crying in others (as one might experiment with an unfamiliar piece of electronics). However, the young child has limited

⁵ Of course, one might deny that toddler comforting behaviors count as core cases of altruism. Rather, one might claim that such cases should be construed as ersatz altruism. However, one would need an argument for excluding these cases. For if we focus on the underlying motivation, the evidence suggests that altruistic concern in toddlers is continuous with altruistic concern in later childhood and adulthood (e.g., Zahn-Waxler et al., 1992a; Eisenberg and Fabes, 1990; Eisenberg et al., 1989).

mindreading resources at hand and thus relies on egocentric mindreading strategies (Nichols and Stich forthcoming). As a result, the child's knowledge of how *his distress* is relieved guides his thinking about how to relieve the other person's distress. Thus, the toddler's egocentric comforting cases are not only *consistent* with the minimalist account, the cases provide further evidence that the child attributes distress to the other person.

Although there is strong evidence against the perspective-taking model, it would be derelict to claim a quick victory for the minimalist account that I've proposed. For there is a less austere alternative that is not excluded by the evidence. By the time toddlers exhibit comforting behaviors, they probably have the capacity to attribute desires that they don't have (see, e.g., Repacholi and Gopnik, 1997). So one might maintain that it is *this* mindreading capacity, the capacity to attribute discrepant desires, that is essential for altruistic motivation. This view has not been elaborated and defended in the literature, but it's possible that the view is close to Blum's (1994) account. Recall that Blum maintains that the understanding of others required for altruistic motivation depends on understanding that others might have thoughts and feelings that are 'different from what oneself would feel in the same situation' (p. 193). He rejects more austere accounts as too 'egocentered' (p. 193).

While this moderate 'discrepant desire' position doesn't contravene any of the data, it's unclear why the capacity to attribute discrepant desires should be essential to altruistic motivation. To see this, it's important to distinguish between three different kinds of egocentrism. One kind of egocentrism is just the view that an individual's basic motivations derive solely from that individual's own affective or hedonic states. We might call this view *psychological egoism*. Psychological egoism might be wrong, but the issue belongs primarily to the foundations of cognitive science, not to moral psychology. On the second kind of egocentrism, let's call it *ethical egocentrism*, a person is egocentric if none of the individual's desires are directed at another person's needs, except insofar as the individual thinks that addressing the other person's needs will help him. What's crucial about ethical egocentrism (and what distinguishes it from simple psychological egoism) is that if a person is ethically egocentric, he must go through a process of instrumental reasoning before arriving at a motivation to help another. For he must *think* that helping another will benefit himself. Both of these kinds of egocentrism need to be distinguished from a third kind of egocentrism – *mindreading egocentrism*. To say that someone is egocentric in this sense is to claim that the individual either can't or tends not to grasp that others have different likes and dislikes, different judgments, and different feelings than the individual himself. Notice that ethical egocentrism and mindreading egocentrism make quite independent claims. A person can perfectly well be ethically egocentric without being an egocentric mindreader. That is, a person might know that others have different interests and beliefs than he does, but at the same time, he might not care in the least about the interests of others, except insofar as he thinks it will affect him. Psychopaths seem to fit this characterization fairly well. Conversely, a person could be an egocentric mindreader without being ethically egocentric. That is, a person might be oblivious to the fact that others have different desires and thoughts than she does, but she might care about trying to help others in need, even if she doesn't think that doing so will serve her own interests. Of course, if she is an egocentric mindreader, she may not be very effective at helping others, because she won't be sensitive to the variation in desires, feelings, and thoughts that actually exist among those she tries to help. Now, finally, we can get to the point of drawing these distinctions – if someone is an egocentric mindreader, that provides no reason to conclude that she lacks altruistic motivation. The kind of egocentrism that undermines the claim for altruistic motivation is *ethical* egocentrism, not *mindreading* egocentrism. As we've seen, when toddlers

offer comfort, they often offer their own comfort objects to others. The fact that these children are using egocentric mindreading strategies does not undermine the claim that these children are altruistically motivated. Even if children turned out to be completely egocentric mindreaders, I see no reason to conclude that their attempts to comfort adults with their dolls and blankets would not be the product of altruistic motivation. Thus, although the discrepant desire view fits with the available evidence, it's not at all clear why we should prefer this account to the simpler minimalist theory.

7. Affect and Altruistic Motivation

I've argued that altruistic motivation depends only on the minimal mindreading capacity for distress attribution, but I've said nothing about how attributing distress to another leads to altruistic motivation. In keeping with most other accounts, I will assume that altruistic motivation is mediated by an affective response (see e.g., Eisenberg, 1992; Goldman, 1993; Hoffman, 1991).⁶ So, on the account I'm suggesting, the attribution of distress triggers an affective response that generates the motivation to help the person in distress. However, there are a couple of importantly different possibilities for the character of the affective response. One possibility is that the representation of the other's distress produces a distinctive emotion of sympathy or concern for the other person and this emotion is not homologous to the emotion of the person in need. The sympathy view has some support from an emerging body of research which ties altruistic behavior to a distinctive facial expression (Roberts and Strayer, 1996, 456; Eisenberg et al., 1989, p. 58; Miller et al., 1996, 213). There is also a bit of evidence that sympathy might have distinctive physiological characteristics (Eisenberg and Fabes, 1990, p. 140; Miller et al., 1996). Facial expression and physiological signs are two of the central features that have been used to delineate 'basic emotions' (e.g., Ekman, 1992). The exciting possibility here is that sympathy is a genuine, distinctive basic emotion with a distinctive facial expression and physiological profile and that this emotion is the motivation behind altruistic behavior. Darwin himself actually made a similar suggestion: 'Sympathy with the distresses of others, even with the imaginary distresses of a heroine in a pathetic story, for whom we feel no affection, readily excites tears.... Sympathy appears to constitute a separate or distinct emotion' (Darwin, 1872, p. 215). But Darwin seems to have had a somewhat different notion of sympathy in mind since he thinks that we can sympathize with the happiness of others.

⁶Another possibility is that affect plays no role in altruistic motivation. Rather, perhaps altruistic motivation follows directly from an attribution of distress. Something like this might be Sober and Wilson's ultimate view (1998, pp. 312 ff.). They suggest that evolution built a mechanism for altruistic motivation that does not rely on hedonic or affective states. However, they do not explain how that mechanism might have evolved in the existing motivational systems of our ancestors. The standard models of motivation in psychology are 'hot' models, on which affect plays the central role in basic motivation (see, e.g., Kunda, 1999). To make a 'tepid' model of altruistic motivation plausible would require a broader defense of the idea of tepid basic motivation. Furthermore, there is some evidence suggesting important correlations between affect and altruistic behavior. As I'll elaborate shortly, the developmental data suggest a correlation between affective response and helping behavior in children, and the social psychological data suggest a similar correlation in adults. In addition, as we'll see in section 8, evidence on psychopathy indicates that psychopaths' lack of helping behavior might be correlated with a deficit to their affective response to others' distress.

The possibility that altruistic motivation derives from a distinctive basic emotion of sympathy is theoretically appealing, but it has turned out to be difficult to get unequivocal data correlating the postulated features of sympathy with altruistic behavior. There are several different measures – e.g., self-report, facial expressions, physiological measures. The findings suggest that some of these features are correlated with altruistic behavior some of the time. For example, Eisenberg and Fabes (1990) showed preschoolers a film of children who were injured and in the hospital, and the preschoolers were given the chance to help the hospitalized children by packing crayons for them rather than playing. Although children’s self-reports were unrelated to their helping behavior, the physiological measure of sympathy (heart-rate deceleration) was positively correlated with higher levels of helping (Eisenberg and Fabes, 1990, pp. 140-1). Further, facial expressions of concerned attention have been significantly correlated with greater helping in boys, but the findings are much weaker for girls (Eisenberg and Fabes, 1990, p. 141). And there is a bit of evidence that there is a correlation between these emotions and helping behavior in Batson-style experiments (Eisenberg et al., 1989).

Notice that if the above account of the affect is right, sympathetic motivation for altruism doesn’t count as empathy at all. Rather, altruistic behavior is motivated by a distinctive emotion that is not homologous to the emotion felt by the person in need, or indeed homologous to any other emotion.⁷ This would entail that a certain class of empathy-based accounts is thoroughly mistaken. If empathy is a vicarious feeling of the emotion that the target is feeling (caused by perspective taking or emotional contagion), then the empathy account is wrong not just about the mindreading required for altruistic motivation but also about the affect. For on the sympathy account, the emotion driving altruistic behavior does not parallel any other emotion. So, except in the iterative case of empathizing with someone feeling sympathy, empathy will not produce the emotion that generates altruistic behavior.

Although the idea that a distinctive emotion of sympathy underlies altruism is theoretically appealing, there is another possibility. The distress attribution might produce a kind of second order empathic distress in the subject. For example, representing the sorrow of the target might lead one to feel sorrow. This would provide a kind of empathic motivation for helping. And the motivation would be effective even when escape is easy. For the cause of the emotion is still the representation of the other’s mental state and as a result, one is motivated not simply to escape the situation since that would not rid one of the representation. As a result, this story would provide an equally effective explanation of Batson’s data. And some of the above research on sympathy actually provides support for this alternative story. For instance, Eisenberg and colleagues (1989) found that the strongest predictor of helping in adults was not facial sympathy, but facial sadness (Eisenberg et al., 1989, 61). The available evidence doesn’t

⁷As we saw in section 2, Sober and Wilson (1998, pp. 234-5) maintain that sympathy doesn’t require that the sympathizer and the target feel the same emotion simultaneously. But that doesn’t really distinguish sympathy from sophisticated accounts of empathy. The psychological work, however, really does raise the possibility of a profound distinction. Feelings of sympathy may not parallel *any* other feeling.

really decide between these two accounts of the affect underlying altruistic motivation.⁸ Indeed, perhaps both affective mechanisms are operative.⁹

8. The Concern Mechanism

For present purposes, what is really crucial is not the character of altruistic affect (whether it's a distinctive emotion or homologous to some other emotion) but the broader characterization of the cognitive mechanisms implicated in altruistic motivation. We are now in a position to state the proposal about the core architecture a bit more precisely. Altruistic motivation depends on a mechanism that takes as input representations that attribute distress, e.g., *John is experiencing painful shock*, and produces as output affect that *inter alia* motivates altruistic behavior. To avoid the terminological difficulties with 'sympathy' I'll use a slightly less problematic term and call this system the Concern Mechanism. In this section, I want to provide a somewhat fuller characterization of the Concern Mechanism, and I'll begin by revisiting the perspective taking account. For there seems to be a double dissociation between the capacity for perspective taking and the capacity for concern.

First, let us return to the developmental evidence. The comforting behaviors of toddlers suggest that the Concern Mechanism is intact and functioning in very young children. This is corroborated by a recent study in which Zahn-Waxler and colleagues traced the development of concern and comforting behaviors in one year old children. They trained mothers to record their child's emotional and behavioral responses to distress in others. Mothers were also trained to simulate various distress situations. Between 13-15 months, children were reported to respond with concern to 9% of the natural distress situations and 8% of the simulated distress situations. Between 18-20 months, children responded with sad facial expressions or concerned attention to 10% and 23% of natural and simulated distress situations. And by 23-25 months, children responded this way to 25% and 27% of natural and simulated distress situations (Zahn-Waxler et al., 1992a, p. 131). So it certainly appears that the capacity for concern or sympathy emerges before the age of 2. Furthermore, between 18-20 months, there is a marginally significant

⁸ I've sketched out two possible affective bases for altruistic motivation. In some interesting work, Batson suggests that there is a different, distinctive affective response that can arise from witnessing another in pain, 'personal distress'. Batson maintains that the feeling of personal distress is self-directed, and he has evidence indicating that the feeling of personal distress does not motivate altruistic behavior (1990).

⁹As Paul Harris has reminded me, there is a great deal of individual variation among young children in their response to distress in others (e.g., Cummings et al., 1986; Zahn-Waxler et al., 1979). This might be thought to undermine the suggestion that a basic emotion underlies altruistic behavior. However even in emotions that are widely accepted to be basic emotions, one finds considerable individual variation (see, e.g., Haidt et al., 1994 on individual variations in disgust). More importantly, as is often the case in cognitive science, it is difficult to know what to make of the individual variation in behaviors. There is an abundance of factors that seem to contribute to the individual variation in children's responses to another's distress. Some of the variation is attributed to differences in child-rearing practices (Zahn-Waxler et al., 1979); some of it is attributed to other family environmental features (e.g., Klimes-Dougan and Kistner, 1990); some of the variation seems to be genetically based (Zahn-Waxler et al., 1992b), and, in older children, some of the variation might be due to differences in perspective taking abilities (Stewart and Marvin, 1984). In light of the complex interaction of these and other features, the individual variation seems quite consistent with the proposal that altruistic motivation depends on a basic affective mechanism.

correlation between concern and comforting behavior, and by 23-25 months, there is a very significant correlation between concern and comforting behavior. The developmental pattern charted by these results suggests, perhaps not surprisingly, that the coordination of the concern response and altruistic behavior is a complicated developmental process. This developmental process no doubt depends on a suite of conditions, environmental and otherwise, that we don't understand. Nonetheless, the broad pattern indicates that the Concern Mechanism is up and running well before the capacity for perspective taking has developed, which suggests that the Concern Mechanism is dissociable from the capacity for perspective taking.

The possibility of a dissociation between the Concern Mechanism and the capacity for perspective taking is further supported by evidence on children with autism. Researchers in the mindreading tradition have explored in some detail the capacities of people with autism, and on a wide range of mindreading tasks, children with autism tend to perform much worse than their mental age peers (see e.g., Baron-Cohen, 1995; Frith, 1989). For instance, most autistic children fail false belief tasks long after their mental age peers can pass such tasks (e.g., Baron-Cohen et al., 1985). In addition to their difficulties with false belief, autistic children fail classic perspective-taking tasks like determining which gifts would be appropriate for which person (Dawson and Fernald, 1987). Further, one of the central characteristics of autism is lack of imaginative activities and spontaneous pretend play (Wing and Gould, 1979). Thus, there is considerable evidence that the capacity for perspective taking is seriously compromised in autism.

Despite their difficulties with perspective taking and imagination, recent studies show that autistic children *are* responsive to distress in others (Blair, 1999; Yirmiya et al., 1992). For instance, in one recent experiment, autistic children were shown pictures of threatening faces and distressed faces, and the autistic children showed the normal pattern of heightened physiological response to both sets of stimuli (Blair, 1999). Thus, although autistic children have a deficit in perspective taking, they do respond to the distress of others. More importantly for our purposes, a recent study suggests that autistic individuals will engage in comforting behaviors. Sigman and colleagues (1992) explored responses to distress in autistic, Downs Syndrome and normally developing children. In one task, the distress was made as salient as possible. The parent was seated next to her child at a small table, and while showing the child how to use a hammer with a pounding toy, the parent pretended to hurt her finger by hitting it with the hammer. The parent then made facial and vocal expressions of distress but didn't utter any words (Sigman et al., 1992, p. 798). Researchers found that autistic children were much less likely than other children to *attend* to the distress. This fits with a broader pattern of inattentiveness to social cues in autism. For instance, autistic children are much less likely than Down Syndrome children to orient to someone clapping or calling their name (Dawson et al., 1998). Despite the fact that autistic children were less likely to notice or attend to the distress, several autistic children provided comfort to the parent in this experiment. Overall, few children helped, but autistic children helped as often as the children in the other groups.¹⁰

The fact that autistic children show normal physiological response to distress in others and the finding that autistic children do engage in comforting behaviors suggests that the core architecture for altruistic motivation is intact in autism. This poses a serious problem for the

¹⁰6 out of 29 autistic children helped; 7 out of 30 mentally retarded children helped; and 3 out of 30 normally developing children helped (Sigman et al., 1992, 800).

perspective taking account since that account predicts that individuals with serious deficits to imagination and perspective taking would show corollary deficits to altruistic motivation.¹¹

So, even though autistic children have a profound deficit in perspective taking, the available evidence indicates that they have no correspondingly serious deficit to the Concern Mechanism. The complementary question is whether there are individuals who show a deficit to the Concern Mechanism but no serious deficit to perspective taking. In fact, it's plausible that psychopaths fit this description. The standard diagnostic tool used in the United States, the DSM IV, uses the diagnostic category of Antisocial Personality Disorder, and the DSM IV suggests that psychopathy is the same condition (p. 645). People with Antisocial Personality Disorder 'frequently ... tend to be callous, cynical, and contemptuous of the feelings, rights, and sufferings of others' (p. 647). 'Persons with this disorder disregard the wishes, rights, or feelings of others. They are frequently deceitful and manipulative in order to gain personal profit or pleasure (e.g., to obtain money, sex, or power).... They may believe that everyone is out to 'help number one' and that one should stop at nothing to avoid being pushed around' (p. 646). A number of researchers characterize psychopathy somewhat differently from the Antisocial Personality Disorder (e.g., Hare, 1991), but the alternative diagnostic criteria tend to present a similarly disturbing portrait of psychopaths. For instance, psychopathy is characterized by a lack of remorse and empathy, being deceitful and manipulative, and a tendency to adult antisocial behavior (Hare, 1991). These characterizations certainly suggest that psychopaths are significantly less likely than non-psychopaths to exhibit altruistic behavior. Recent evidence provides an explanation for this – unlike autistic and normal children and adults, psychopaths show relatively little physiological response to the distress of others (Blair et al., 1997). Blair and colleagues found that non-psychopathic criminals show about the same amount of Skin Conductance Response to another's distress cues as they do to threatening stimuli. Psychopaths, on the other hand, show normal response to threatening stimuli but significantly lower response to the distress cues of another (Blair et al., 1997). Coupled with their apparent lack of altruistic behavior, this suggests that the Concern Mechanism is defective in psychopathy. Nonetheless, evidence indicates that psychopaths are quite capable of perspective taking, and that they perform as well as normal adults on standard perspective taking tasks (Blair et al., 1996).

Hence, while the evidence is still quite preliminary, there seems to be a double dissociation between perspective taking and the Concern Mechanism. Young children and autistic children have immature or impaired perspective taking abilities, yet they seem to have the Concern Mechanism intact and functioning. Psychopaths, by contrast seem to have a normal capacity for perspective taking but a deficit to the Concern Mechanism. The evidence from development and psychopathologies thus counts heavily against the perspective taking account. It seems that altruistic motivation does not require sophisticated mindreading or perspective taking abilities. And it doesn't take any imagination to be an altruist.¹²

¹¹ Although the evidence on autistic children poses a serious problem for perspective taking accounts, it is perfectly compatible with the minimal mindreading account. For despite their deficits in perspective taking, children with autism are capable of attributing simple negative emotions (e.g., Yirmiya et al., 1992; Baron-Cohen, 1995; Tager-Flusberg, 1993).

¹² The Concern Mechanism has many of the features of modules as set out by Fodor (1983). The evidence on development and psychopathology indicates that the Concern Mechanism has a characteristic ontogeny and a characteristic pattern of breakdown. It's also plausible that the mechanism is fast. It is somewhat more difficult to evaluate whether the Concern Mechanism is 'encapsulated' (Fodor, 1983, 2000) since the relationship between affective systems and encapsulation is far from clear in the current literature. But the Concern Mechanism plausibly

If the basic story about mindreading and the Concern Mechanism is right, it has a particularly interesting implication for the possibility of altruism in non-human animals. For if human altruism requires so little mindreading, it becomes quite possible that the mechanisms underlying helping-behavior in some non-human animals are analogous to the mechanisms underlying altruistic motivation in humans. Although it's hotly debated at present, some non-human animals may well have the mindreading capacity to attribute distress to another. There is some evidence, for instance, that chimpanzees can attribute goals (Premack and Woodruff, 1978; Call and Tomasello, 1998). Research also suggests that non-human primates are sensitive to a conspecific's distress signals (e.g., Miller et al., 1963).

Apart from its intrinsic interest, the possibility that altruistic motivation might be present in non-humans is of some importance to an evolutionary approach to altruism. If altruistic motivation in humans is an adaptation that depends on sophisticated mindreading abilities like perspective taking, then the altruistic motivation system must have been shaped after the evolution of our sophisticated mindreading abilities. If so, the mechanisms for altruistic motivation presumably emerged relatively recently in evolutionary time since, by most accounts, humans are the only primates with sophisticated mindreading abilities. The Concern Mechanism account of altruistic motivation, on the other hand, needn't be committed to the view that altruistic motivation is a recent adaptation since on this view the requisite mindreading mechanisms are minimal and may well have been present in our more distant phylogenetic ancestors.

At this point, the research on mindreading and altruism suggests a strikingly simple model of the core architecture underlying basic altruistic motivation. The evidence suggests that basic altruistic motivation requires only a minimal capacity for mindreading, the capacity to attribute negative affective or hedonic mental states like distress. These attributions, I've suggested, produce altruistic motivation by activating an affective system, the Concern Mechanism. Of course, the account I've offered in this paper is hardly a full account of the cognitive mechanisms implicated in mature altruistic behavior, for the altruistic capacities of adult humans far outstrip those provided by the primitive mindreading and Concern mechanisms. Nonetheless, the empirical work suggests that the Concern Mechanism and a minimal capacity for mindreading form the core of our capacity for altruism.

Department of Philosophy
College of Charleston

possesses at least one feature of encapsulated systems. A cognitive mechanism is encapsulated if it has little or no access to information outside of its own proprietary database, and one of the central features of an encapsulated system is that such systems resist our preferences (Fodor 2000, p. 63): You can't make the Muller-Lyer illusion disappear by wanting it to go away. It's likely that the Concern Mechanism is similarly resistant to our preferences and to the dictates of practical reason. For instance, I might think it's best, all things considered, not to feel concern when my daughter gets inoculated because any show of concern on my part might intensify her anxiety about inoculations. Nonetheless, it can be extremely difficult to suppress concern in these circumstances. In this sense at least, the Concern Mechanism resembles encapsulated systems.

References:

- American Psychiatric Association 1994: *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. Washington, D.C.: American Psychiatric Association. (DSM-IV).
- Baron-Cohen, S. 1995: *Mindblindness*. Cambridge, MA: MIT Press, Bradford Books.
- Baron-Cohen, S., Leslie, A.M., and Frith, U. 1985: Does the autistic child have a 'Theory Of Mind'? *Cognition*, 21, 37-46.
- Bartsch, K. and Wellman, H. 1995: *Children Talk about the Mind*. New York: Oxford University Press.
- Batson, C. 1990: How social an animal?: The human capacity for caring. *American Psychologist*, 45, 336-346.
- Batson, C. 1991: *The Altruism Question*. Hillsdale, NJ: LEA.
- Batson, C., Duncan, B., Ackerman, P., Buckley, T., and Birch, K. 1981: Is empathic emotion a source of altruistic motivation?. *Journal of Personality and Social Psychology*, 40, 290-302.
- Batson, C., O'Quin, K., Fultz, J., Vanderplas, M. and Isen, A. 1983: Self-reported distress and empathy and egoistic versus altruistic motivation for helping. *Journal of Personality and Social Psychology*, 45, 706-718.
- Blair, R. 1995: A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, 57, 1-29.
- Blair, R. 1999: Psychophysiological responsiveness to the distress of others in children with autism. *Personality and Individual Differences*, 26, 477-485.
- Blair, R., Jones, L., Clark, F., Smith, M. 1997: The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology*, 34, 192-198.
- Blair, R., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M. and Jones, L. 1996: Theory of mind in the psychopath. *Journal of Forensic Psychiatry*, 7, 15-25.
- Bloom, L. 1970: *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.
- Bloom, L. 1973: *One Word at a Time: The Use of Single Word Utterances before Syntax*. The Hague, Netherlands: Mouton.
- Blum, L. 1994: Moral development and conceptions of morality. In *Moral Perception and Particularity*, L. Blum. Cambridge University Press.

- Byrne, D. 1971: *The Attraction Paradigm*. NY: Academic Press.
- Call, J. and Tomasello, M. 1998: Distinguishing intentional from accidental actions in orangutans, chimpanzees, and human children. *Journal of Comparative Psychology*, 112, 192-206.
- Campbell, J. 1999: Can philosophical accounts of altruism accommodate experimental data on helping behavior? *Australasian Journal of Philosophy*, 77, 26-45.
- Carey, S. 1985: *Conceptual Change in Childhood*. Cambridge, MA: MIT Press.
- Clark, R. and Wood, L. 1974: Where is the apathetic bystander? Situational characteristics of the emergency. *Journal of Personality and Social Psychology*, 29.
- Cummins, E., Hollenbeck, B., Iannotti, R., Radke-Yarrow, M., and Zahn-Waxler, C. 1986: Early organization of altruism and aggression: developmental patterns and individual differences. In C. Zahn-Waxler, E. Cummings, and R. Iannotti (eds.) *Altruism and Aggression: Biological and Social Origins*. New York: Cambridge University Press.
- Currie, G. 1995: The moral psychology of fiction. *Australasian Journal of Philosophy*, 73, 250-259.
- Currie, G. and Ravenscroft, I. forthcoming: *Recreative Minds: Image and Imagination in Philosophy and Psychology*. Oxford: Oxford University Press.
- Damon, W. 1977: *The Social World of the Child*. San Francisco: Jossey-Bass.
- Darwall, S. 1998: Empathy, sympathy, care. *Philosophical Studies*, 89, 261-282.
- Darwall, S., Gibbard, A., and Railton, P. (eds.) 1997: *Moral Discourse and Practice*. Oxford University Press.
- Darwin, C. 1872: *The Expression of the Emotions in Man and Animals*. University of Chicago Press. [Reprinted 1965.]
- Dawson, G. and Fernald, M. 1987: Perspective-taking ability and its relationship to the social behavior of autistic children. *Journal of Autism and Developmental Disorders*, 17, 487-498.
- Deigh, J. 1995: Empathy and universalizability. *Ethics*, 105, 743-763.
- Dunn, J. 1988: *The Beginnings of Social Understanding*. Cambridge, MA: Harvard University Press.
- Eisenberg, N. 1992: *The Caring Child*. Harvard.

- Eisenberg, N., Fabes, R., Miller, P., Fultz, J., Shell, R., Mathy, R. and Reno, R. 1989: Relation of sympathy and personal distress to prosocial behavior: a multimethod study. *Journal of Personality and Social Psychology*, 57, 55-66.
- Eisenberg, N. and Fabes, R. 1990: Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14, 131-149.
- Eisenberg, N and Strayer, N. 1987: Critical issues in the study of empathy. In N. Eisenberg and J. Strayer (eds.), *Empathy and its development*. New York: Cambridge University Press.
- Ekman, P. 1992: An argument for basic emotions. In N. Stein and K. Oatley (eds.), *Basic Emotions*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Flanagan, O. 1991: *Varieties of Moral Personality: Ethics and Psychological Realism*. Harvard University Press.
- Fodor, J. 1983: *Modularity of Mind*. Cambridge, MA: MIT Press, Bradford Books.
- Fodor, J. 2000: *The Mind Doesn't Work That Way*. MIT Press.
- Frith, U. 1989: *Autism: Explaining the Enigma*. Oxford: Blackwell.
- Gergely, G., Nadasdy, Z., Csibra, G., and Biro, S. 1995: Taking the intentional stance at 12 months of age. *Cognition*, 56, 165-193.
- Goldman, A. 1989: Interpretation psychologized. *Mind and Language*, 4, 161-185.
- Goldman, A. 1992: Empathy, mind, and morals. *Proceedings and Addresses of the American Philosophical Association*, 66, No. 3, pp. 17-41.
- Goldman, A. 1993: Ethics and cognitive science. *Ethics* 103, pp. 337-360.
- Gopnik, A. and Wellman, H. 1994: The theory-theory. In L. Hirschfeld and S. Gelman (eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press. Pp. 257-293.
- Gordon, R. 1986: Folk psychology as simulation. *Mind and Language*, 1, 158-171.
- Gordon, R. 1995: Sympathy, simulation, and the impartial spectator. *Ethics* 105, pp. 727-742.
- Haidt, J., McCauley, C., and Rozin, P. 1994: Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences*, 16, 701-713.
- Hare, R. 1991: *The Hare-Psychopathy Checklist-Revised*. Toronto: Multi-Health Systems.

- Harris, P. 1992: From simulation to folk psychology: The case for development. *Mind and Language*, 7, 120-144.
- Hoffman, M. 1976: Empathy, role-taking, guilt and development of altruistic motives. In T. Lickona (ed.), *Moral Development and Behavior: Theory, Research and Social Issues*. New York: Holt, Rinehart and Winston.
- Hoffman, M. 1981: The development of empathy. In J. Rushton and R. Sorrentino (eds.), *Altruism and helping behavior: Social, personality, and developmental perspectives* (pp. 41-63). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hoffman, M. 1982: Development of prosocial motivation: Empathy and guilt. In N. Eisenberg, ed., *Development of prosocial behavior*, pp. 281-313. New York: Academic Press.
- Hoffman, M. 1991: Empathy, social cognition, and moral action. In W. Kurtines and J. Gewirtz (Eds.), *Handbook of Moral Behavior and Development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kohlberg, L. 1984: *The Psychology of Moral Development: The Nature and Validity of Moral Stages*. Harper and Row.
- Krebs, D. 1975: Empathy and altruism. *Journal of Personality and Social Psychology*, 32, 1134-1146.
- Kunda, Z. 1999: *Social Cognition: Making Sense of People*. Cambridge, MA: MIT Press, Bradford Books.
- Latane, B. and Darley, J. 1968: Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10, 215-221.
- Leslie, A. 1994: ToMM, ToBY and agency: Core architecture and domain specificity. In L. Hirschfeld and S. Gelman (eds.) *Mapping the mind*. Cambridge University Press, 119-148.
- MacWhinney, B. and Snow, C. 1990: The child language data exchange system: An update. *Journal of Child Language*, 17, 457-472.
- Miller, R., Banks, J., and Ogawa, N. 1963: Role of facial expression in 'cooperative-avoidance conditioning' in monkeys. *Journal of Abnormal and Social Psychology*, 67, 24-30.
- Miller, P., Eisenberg, N., Fabes, R., and Shell, R. 1996: Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Developmental Psychology*, 32, 210-219.
- Newcombe, T. 1961: *The Acquaintance Process*. New York: Holt, Rinehart and Winston.

- Nichols, S. forthcoming: Norms with feeling: Towards a cognitive account of moral judgment.
- Nichols, S. and Stich, S. 2000: A cognitive theory of pretense. *Cognition*, 74, 115-47.
- Nichols, S. and Stich, S. forthcoming: *Mindreading*. Oxford: Oxford University Press.
- Nichols, S., Stich, S., Leslie, A., and Klein, D. 1996: Varieties of off-line simulation. In P. Carruthers and P. Smith (eds.) *Theories of Theories of Mind*. Cambridge University Press.
- Nucci, L. 1986: Children's conceptions of morality, social conventions and religious prescription. In C. Harding, (ed.), *Moral Dilemmas: Philosophical and Psychological Reconsiderations of the Development of Moral Reasoning*. Chicago: Precedent Press.
- Piaget, J. 1932: *The Psychology of Moral Development: The Nature and Validity of Moral Stages*. M. Gabain, trans. New York: Free Press. (Translation published 1965.)
- Premack, D. and Woodruff, G. 1978: Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 516-526.
- Radke-Yarrow, M., Zahn-Waxler, C., and Chapman, M. 1983: Children's prosocial dispositions and behavior. In P. Mussen (ed.), *Handbook of Child Psychology, Vol. 4, Socialization, Personality, and Social Development*. New York: Wiley.
- Rawls, J. 1971: *A Theory of Justice*. Harvard University Press.
- Repacholi, B and Gopnik, A. 1997: Early understanding of desires: Evidence from 14 and 18 month olds, *Developmental Psychology*, 33, 12-21.
- Roberts, W. and Strayer, J. 1996: Empathy, emotional expressiveness, and prosocial behavior. *Child Development*, 67, 449-470.
- Rosenbaum, M. 1986: The repulsion hypothesis: On the nondevelopment of relationships. *Journal of Personality and Social Psychology* 51, 1156-1166.
- Sachs, J. 1983: Talking about there and then: The emergence of displaced reference in parent-child discourse. In K. E. Nelson (ed...) *Children's Language*, Vol. 4, 1-28. Hillsdale, NJ: Earlbaum.
- Selman, R. 1980: *The Growth of Interpersonal Understanding*. New York: Academic Press.
- Sigman, M., Kasari, C., Kwon, J., and Yirmiya, N. 1992: Responses to the negative emotions of others by autistic, mentally retarded, and normal children. *Child Development*, 63, 796-807.
- Simner, M. 1971: Newborn's response to the cry of another infant. *Developmental Psychology* 5, 136-150.

- Smetana, J. and Braeges, J. 1990: The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly*, 36, 329-346.
- Sober, E. and Wilson, D. 1998: *Unto Others*. Harvard University Press.
- Stewart, R. and Marvin, R. 1984: Sibling relations: The role of conceptual perspective-taking in the ontogeny of sibling care-giving. *Child Development*, 55, 1322-32.
- Stich, S. and Nichols, S. 1992: Folk psychology: simulation or tacit theory. *Mind & Language*, v. 7, no. 1, 35-71.
- Stotland, E. 1969: Exploratory investigations of empathy. In L. Berkowitz, ed., *Advances in experimental social psychology*, 4. New York: Academic Press.
- Tager-Flusberg, H. 1993: What language reveals about the understanding of minds in children with autism. In S. Baron-Cohen, H. Tager-Flusberg and Donald Cohen (eds.) *Understanding Other Minds: Perspectives from Autism*, 138-157.
- Tajfel, H. 1981: *Human Groups and Social Categories*. Cambridge: Cambridge University Press.
- Turiel, E., Killen, M., and Helwig, C. 1987: Morality: Its structure, functions, and vagaries. In J. Kagan and S. Lamb (eds.) *The Emergence of Morality in Young Children*. Chicago: University of Chicago Press, 155-244.
- Wellman, H., Harris, P., Banerjee, M., and Sinclair, A. 1995: Early understanding of emotion: Evidence from natural language. *Cognition and Emotion*, 9, 117-179.
- Wing, L., and Gould, J. 1979: Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9, 11-29.
- Woodward, A. 1998: Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1-34.
- Yirmiya, N., Sigman, M., Kasari, C., and Mundy, P. 1992: Empathy and cognition in high-functioning children with autism. *Child Development*, 63, 150-160.
- Zahn-Waxler, C., and Radke-Yarrow, M. 1982: The development of altruism: Alternative research strategies. In N. Eisenberg-Berg (ed.), *The Development of Prosocial Behavior*. New York: Academic Press.
- Zahn-Waxler, C., Radke-Yarrow, M., and Brady-Smith, J. 1977: Perspective-taking and prosocial behavior. *Developmental Psychology*, 12, 87-88.

Zahn-Waxler, C., Radke-Yarrow, M., and King, R. 1979: Child rearing and children's prosocial initiations toward victims of distress. *Child Development*, 50, 319-330.

Zahn-Waxler, C., Radke-Yarrow, M. , Wagner, E. and Chapman, M. 1992a: Development of concern for others. *Developmental Psychology*, 28, 126-136.

Zahn-Waxler, C., Robinson, J., and Emde, R. 1992b: The development of empathy in twins. *Developmental Psychology*, 28, 1038-1047.