The crossroads of theory of mind and moral reasoning: intention†

Adam Cohen*

Department of Psychology and Center for Cognitive Science
Rutgers, the State University of New Jersey
New Brunswick, NJ 08901

* Rutgers Undergraduate Research Fellow

†Based closely on a thesis submitted in fulfillment of the requirements of The Henry Rutgers Scholars Program and written under the direction of Dr. Alan Leslie, April 2004.

Abstract

What is the relationship between moral judgment and the child's "theory of mind?" The present studies address whether preschool children can attribute the mental state, caring/not caring, and judge whether the consequences of an act were intended or not, as a function of the moral valence of the consequence. Pilot data suggested that younger children might not understand "not caring." Experiment 1 examined whether children, 3- to 5-years-old, could predict the affective state of someone who either cared or did not care about a food object or a person under different circumstances. Regarding objects, the results showed that 3-, 4-, and 5-year olds have adult-like intuitions regarding caring about objects and not caring about objects. The pattern of data was similar across the three age groups, but with strengthening effects across age. The "caring about people" results indicated that children were more likely to say that a person would feel happy or sad if someone they cared about was happy or sad. On the other hand, children tended to say that a person would feel neutral if someone else was happy or sad if the person did not care about the other person. These results demonstrate an understanding of both caring and not caring about people.

Experiment 2 explored how caring/not caring interacts with a judgment of "on purpose" in moral contexts. Adults have recently been shown to make asymmetric judgments of intention depending on whether a consequence of an action, that a protagonist does not care about, is beneficial or harmful to others. If the not-cared-about side effectside effect is beneficial, adults judge that the actor did not intend it. If the side effectside effect is harmful then the actor is judged to have brought it about intentionally. Preschool children, 3-, 4-, and 5-years-old were tested on scenarios in which an actor knew about, but did not care about, a side effectside effect of one of their actions. A control question screened for an understanding of "not caring that [someone would be harmed…]." A test question then required the subject to judge whether the side effectside effect of the action was brought about "on purpose." The results showed that for those children who failed the control "do not care that…" question, there was an overall bias to answer, "yes" to the "on purpose" question. However, the majority of children (aged 5 years), who passed the "not care" control question, showed the adult-like asymmetry, judging the side effectside effect to have been brought about intentionally when harmful, and to have been unintentional when beneficial. This is the first study to examine the relationship between developing 'theory of mind' and moral judgment in preschoolers.

Introduction

A wealth of studies have been conducted and a vast literature exists in both "theory of mind" and moral reasoning research, two areas of cognitive psychology. However, the relationship between the development of these two domains has received very little attention. One significant way in which these two domains are related to each other is through the concept of *intention*. Intentionality, in the sense of acting purposefully (as opposed to intentionality in the philosophical sense of "aboutness"), is central to acquiring a "theory of mind." Purposefulness in action is also crucially involved in moral (and legal) judgment.

Moral development

The history of moral reasoning research that still applies to today's theories can be traced back to the Swiss psychologist Jean Piaget and his landmark book, *The Moral Judgment of the Child* (1932). Piaget classified children into one of several categories based on their level of moral development.

Before two years of age, Piaget suggested that a child has no moral rules but rather their behavior is governed solely by motor abilities. They have no understanding of social cooperation and everything is viewed from their wants and desires.

During the third year, the child begins to develop a sense of morality. Piaget breaks this development up into two stages. In the first phase called morality of constraint, children between about two and seven years of age are very rigid in their beliefs of moral concepts. It is at this stage that the child is said to be egocentric because they are not capable, according to Piaget, of taking someone else's perspective into account, and so they are unable to use intention as a basis for understanding behavior. Children at this stage are "heteronomous," meaning they see moral rules as being inflexible and behaviors as being either right or wrong (Duska and Whelan, 1975). That children see moral rules this way is grounded in the unilateral respect children have for adult authority (Lapsley, 1996).

After the age of seven, children gradually move into the second phase, labeled the morality of cooperation. By developing equal relationships with peers, children transition from heteronomy to autonomy by being provided with a social environment that encourages them to work cooperatively together, changing rules through mutual consent. This stage is characterized by flexibility in that the child realizes that rules are not absolute and can be changed. Also, the child develops the ability to take multiple points of view, enabling them to make moral judgments based on the intent behind the behavior. This stage is grounded by the mutual respect that characterizes a society of equals (Lapsley, 1996).

Although Piaget was a meticulous observer and advanced entire fields of research in psychology, many of his studies including his moral reasoning work were experimentally flawed. Other psychologists picked up Piaget's work and produced more carefully designed studies. Even though these studies were also flawed, they were important in the sense that they "established a tradition of the empirical study of human judgments, as opposed to

behaviors…[D]evelopmental psychology came to the study of cognition far earlier than did many other fields of psychology" (Darley and Shultz, 1990).

Another early, major contributor to the area of moral development was Lawrence Kohlberg. Influenced by Piaget and moral philosophers, Kohlberg held that moral judgments progress through an unvarying, irreversible series of stages (for review see Kohlberg, 1981). Progress through the stages is moved forward by interactions with peers, but not everyone reaches the highest level of moral development, according to Kohlberg. By evaluating not the decision itself but the justification behind it, Kohlbergians have seemed to assume that all moral decisions are equally amenable to justification (Darley and Shultz, 1990).

More recent studies have challenged as well as modified the work of pioneers such as Piaget and Kohlberg in order to better fit empirical and theoretical anomalies. Turiel and his colleagues are responsible for an important line of research that draws a clear distinction between two domains of knowledge. They have demonstrated that children understand the distinction between moral rules (e.g. lying, stealing) and social conventions (e.g. dress code, forms of address) early in development (Turiel, 1983; Nucci and Nucci, 1982). They regard moral rules as obligatory, universal, and unalterable. They view social-conventional rules as alterable, based on society consensus. There may even be a cultural influence such that the distinction between "morality" and "convention" varies between cultures (Shweder et al, 1987). Kohlberg claimed a cross-cultural universality of his stages, but other findings challenge this claim (Gilligan, 1982).

Related work by Smetana (1990) looked at the relationship between moral and conventional judgments and the effects of language development on those judgments. Consistent with previous findings, children around 3.5-years-old distinguish morality and convention on all criteria, including permissibility, seriousness, generalizability, and rule and authority contingency. However, 34-month olds only judged moral transgressions to be more generalizable, and the youngest group, 26-month olds, did not distinguish morality and convention on any of the criteria. In addition, children who responded correctly to the language comprehension items differentiated between morality and conventional transgressions on the basis of generalizability, rule contingency, and authority contingency at earlier ages than when language ability was not considered (Smetana, 1990).

Piaget found a developmental trend that children in the egocentric stage of development do not morally judge the performer of a harmful act based on motives, that is, information provided by intention. Rather they used outcomes, that is, the amount of damage caused by the act, as the basis for moral judgments. Piaget's work showed in children under 9-10 years no clear preference for motives as the basis for judgment. However work by Nelson (1980) shows that children as young as 3-years-old use motive information in making a moral judgment when that information is explicit, salient and available. Further, Nelson demonstrated that when motive and outcomes are incongruent (e.g., a good motive such as intending to share a ball with another child and a bad outcome such as the ball hitting the other child in the head, causing him/her to cry), children tend to recall the story by making them congruent (Nelson, 1980).

Other theorists have devoted their attention to studying under what conditions moral responsibility and blame are assigned. Shultz et al (1986) discovered that children between 5-11

years of age employ a fairly sophisticated repertoire of moral concepts. They showed evidence of knowing that causal judgments are presupposed by judgments of moral responsibility, which in turn are presupposed by judgments about punishment. Children also used information on intention and negligence in making responsibility judgments and used information on restitution in assigning punishment. Developmental trends include an increasing sensitivity to these concepts, greater tolerance for harm doing, and more emphasis on restitution rather than punishment with increasing age (Shultz et al, 1986).

"Theory of mind" development

There is a parallel literature to moral reasoning on another category of social cognition, namely, "theory of mind" research. As social creatures that depend on interactions with people in their environment for survival, it is important that humans be capable of reasoning about people and their behavior (Flavell, 1998). "Theory of mind" refers to our ability to explain, predict and interpret behavior in terms of mental states, such as pretending, believing, wanting, and intending (Leslie, 2000). The theory of mind approach is one of the most active areas of research in developmental psychology (Flavell, 1998). With the publication of a single paper in 1978 titled, "Does the chimpanzee have a theory of mind?" psychologists Guy Woodruff and David Premack introduced the field of "theory of mind" research. Premack and Woodruff stated, "An individual has a theory of mind if he imputes mental states to himself and others" (Premack and Woodruff, 1978). In other words, "theory of mind" is the cognitive ability to attribute mental states to self and others.

Being able to recognize mental states in others is one of the most sophisticated abilities young children have. One way to assess "theory of mind" in children is the false belief task. In one type of false belief task called the Sally-Ann Task (Baron-Cohen, Leslie, and Frith, 1985), a character named Sally, while in a room, puts a ball in a box. Sally leaves the room, and while gone, another character named Ann comes and moves the ball from the box to the basket. The subject then has to make a prediction about where Sally will go to look for the ball or predict where Sally thinks the ball is when she comes back. The key to the false belief task is that the subject's belief is different from Sally's belief and in order to correctly answer the test question, the subject has to attribute a mental state - different from his/her own - to Sally, which is how having a theory of mind is defined by Premack and Woodruff. It is important to note that an experiment designed to test a true belief would have been much less reliable because the subject's belief and the character's belief are the same, so there is no way to know if the subject is attributing a mental state or if he/she is simply reporting back to the experimenter his/her own belief. There were two major findings regarding the false belief task. First, using a slightly more complicated version, Wimmer and Perner (1983) found that the majority of 6-year- olds could pass the task. Second, Baron-Cohen, Leslie, and Frith (1985) showed that children as young as 4-years-old passed the version given above. What makes it even more remarkable that children as young as 4-years-old can attribute mental states is that mental states themselves are undetectable to the senses - you cannot, hear, taste, smell, feel, or see a mental state (Leslie, 2000).

Considering that at the latest, normally developed children come to know about mental states at 4-years of age, some cognitive scientists argue that the maturation of a specialized processing

mechanism, "theory of mind mechanism" (ToMM) describes the underlying mechanism for "theory of mind" (Surian and Leslie, 1999). This processing mechanism must deploy a system of representation capable of representing mental states. "Theory of mind" is intrinsically 'metarepresentational," since it requires one to not only utilize propositional attitudes but to employ them about propositional attitudes (e.g. having beliefs about beliefs) (Scholl and Leslie, 1999).

In a paper by Baron-Cohen, Leslie and Frith, titled, "Does the autistic child have a 'theory of mind'?" they showed that while the majority of subjects with autism were older and had higher IQs than normal and Down syndrome subjects, they failed the false-belief task while the majority of normal children and children with Down's Syndrome passed. They concluded that autistic children are impaired in their understanding about beliefs relative to normally developed children, relative to their own level of general intelligence, and relative to other disorders such as Down syndrome. The central symptoms of autism, a rare neurodevelopmental disorder with biological origins, include social incompetence, poor communication skills (verbal and nonverbal), and lack of pretend play. "Because 'theory of mind' abilities underlie human social competence, communication and pretending, the autistic triad might be the result of an impaired ToMM" (Leslie, 2000b).

Desires and emotions (along with perception) tend to be the earliest mental states young children talk about. Within our naïve theory of mind, desires and emotions are causally linked. For example, people are typically happy when they get what they want. Bartsch and Wellman (1985) propose a three step developmental sequence in understanding desire. First, around 2 years of age, children acquire a desire psychology. They have an elementary conception - although it is mentalistic it is non-representational rather than representational- of desire as well as of simple emotions, simple perceptual experiences, and attention. Second, around age 3, children reach a second level of understanding called a desire-belief psychology. Here children begin to talk about beliefs and thoughts as well as desires, however they still explain behavior in terms of desire only. Finally around the fourth birthday, children acquire the adult belief-desire psychology in which beliefs and desires are thought to determine actions jointly. An understanding of belief and desire is intrinsically linked to understanding intention.

Developing an understanding of "intention"

Developing the concept of intention is highly significant for at least two reasons. First it clarifies how people differ from other objects; humans, unlike other objects, are driven by intentions, motives, goals, and plans. That is, intention is intrinsic in the development of what is referred to as "theory of mind." Second, children must draw on intentions in order to understand responsibility and morality (Shantz, 1983).

Work by Flavell (1996) focused on 3- and 4-year-olds' ability to understand intentions, including prior intentions and intentions-in-action. Prior intentions are what philosopher John Searle means when "one is said to have an intention at the point when one resolves and plans to do something and believes that one will do it, that is, prior to initiating any relevant action." Intentions-in-action occurs when one "takes the further step of actually trying to achieve through action what they had previously formed the mental intention to achieve." The studies used story

lines where the character's intentions were different from both their desires as well as the outcome. The children were asked a series of questions, including "Try" (intention in action) and "Think" (prior intention) questions. Four-year-olds performed well in both studies demonstrating an understanding of both prior intention and intention in action. The results strongly support the notion that older preschoolers understand intentions as mental state representations that differ from desires. The 3-year-old's performance was inconclusive; however they did perform worse than would be expected by chance on the "Try" question. Another study looked more closely at the understanding of intentions, both prior and in-action, in only 3-year- olds. The results of the younger half of 3-year-olds compared with that of the older half revealed that the older half had significantly more understanding of intentions than the younger half. Thus Flavell concludes that "children probably begin to acquire conceptions of prior mental intention and intention-in-action, as differentiated from desire and outcome, some time between 3.5 and 4 years." Moses argues that the possession of "pure intention" around this age is closely connected with the understanding of true and false belief because the ability to understand belief is necessary to understand intention.

A newly developed technique, the "behavioral reenactment procedure" explores what children know about the mind but does not rely on language (Meltzoff, Gopnik, and Repacholi, 1999). Eighteen- month-olds seemed to interpret an adult's effortful behavior as going beyond the literal surface behavior itself, and as being about something else, about the unseen but inferred goal of the act. Also they use the whole pattern of behavior to indicate whether the adult is aiming to do what they are doing or something else. Meltzoff et al. believe 18-month-olds interpret the person's actions within a psychological framework that differentiates between the surface behavior of people and a deeper level involving goals and intentions. At least by 36 months, children know that the adult may desire/intend to do something different from what they do and that the person's emotional reactions after the event are a clue to the underlying desire/intention of the person. Children can first read the goal of the adult's act by 18 months and then later (36 months) can detect regularities in how they themselves and others emotionally react to the successes and failures of goal-directed actions (Meltzoff, Gopnik, and Repacholi, 1999).

Beliefs, desires, and intentions are alike in involving representational relations to reality, but the relationship differs for beliefs and desires and intentions in both directions of fit and causation (Astington and Gopnik, 1991). Beliefs are true if the representation matches reality, and if events in world give rise to our beliefs. Desires and intentions, respectively, are fulfilled if the world comes to match the representation, and the desires cause changes in the world. Young children, who are unable to remember and report their earlier false beliefs, can remember their unfulfilled desires. Three-year-olds are better at understanding the representational nature of desire than of belief. Yet, 3-year-olds seem to have more difficulty discerning the differences in desirability than they have in understanding the relation between desires and actions or outcomes, and this understanding improves between 3 and 4 years. Intentions are distinguished from desires by causal self-referentiality, where the propositional content of an intention is "I do it in order to carry out this intention." Three-year-olds distinguish between intentional and unintentional acts by matching goals and outcomes. Also, 3- year-olds are less likely than 5-year-olds to see intentions as means to ends, and to focus on the future directedness of intention. Three-year-olds seem to lack an understanding of the causal nature of intention. Astington and

Gopnik conclude by stating that 3-year-olds have a mentalistic but non-representational understanding of belief. Further, non-representational aspects of both belief and desire are understood before representational ones, and understanding of desire occurs before that of belief in both accounts. Lastly, an understanding of intention as a mediating factor between desire and action may develop only between four and five years (Astington and Gopnik, 1991).

In a study by Astington (1986), it was assumed that if children do not have the concept of intention than they will be unable to differentiate intentions and the appropriate actions that follow. It was hypothesized that they would not be able to understand the linguistic expressions that were used to make these distinctions. Children's comprehension of different expressions of intentions was tested using phrases in the present tense including *intend, mean, try, plan, will, is going to and wants to* among others and phrases in the past tense such as *intended, meant, on purpose* and their negations. Expressions in the present tense marked prior intentions and expressions in the past tense marked the intentional or unintentional nature of an action. *On purpose* was the easiest expression and *intended to* and *didn't intend* to were the most difficult to comprehend. Conclusions from the results are obscured by the inclusion of negation. 4- and 5-year-olds more frequently made the correct choice for the affirmative than for the negative sentences, although 7-and 9-year-olds did not show this trend. One conclusion Astington draws is that comprehension of expressions of prior intention on the one hand and comprehension of expressions distinguishing intentional and unintentional action, on the other, develop at the same time. The main conclusions to be drawn though are that children's comprehension of expressions of intention improves markedly between the ages of five and nine.

In 'theory of mind' research, judgments of whether an act was performed intentionally or not are assumed to be judgments about matters of fact. An agent either does or does not intend an act depending upon facts about their mental state at the time. However, recent findings with adults suggest that the situation is not so simple (Knobe, 2003). Adults were told two types of story, one in which there are morally good outcomes, and one with morally bad outcomes. When asked to judge whether the outcome was intended by an actor, adults' judgments varied according to the moral valence of the outcome. This shows that the judgments were not purely matter of fact judgments. Here are the stories:

> The chairman of the board of a company has decided to implement a new program. He believes
>
> (1) that the program will make a lot of money for his company
> and
> (2) that the program will also produce some other effect x.
>
> But the chairman doesn't care at all about effect x. His sole reason for implementing the new program is that he believes it will make a lot of money for the company. In the end, everything proceeds as anticipated: the program makes a lot of money for the company and also produces effect x.

The stories differed according to whether the side effect side effect x was good, for example, increased employment of the townspeople, or bad, for example, pollution of a local river. In case

the side effect was good, adults judged that it was *not* intended, and in case the side effect was harmful, they judged that it was intended.

The above effect with adults suggests that some judgments of intention are made outside of the theory of mind domain. The main question pursued here is, at what age do moral-based judgments of intention develop? A second question concerns the relation between moral-based intention judgment and theory of mind development. There are two key aspects of theory of mind that are involved in the Knobe (2003) stories. The first is that the actor *knows* ahead of time the side effect x will occur. The second is that the actor *does not care* that it will occur. Understanding of *knowing* is already fairly well developed around 3.5 years of age (Pratt & Bryant, 1990; Roth & Leslie, 1998). Nothing is known about the state of *not caring*. Because understanding *not caring* is a critical feature of the Knobe scenarios and because information on its development is lacking, the first experiment to be reported focuses on *not caring*.

Two sets of stories were constructed. In the first type, a protagonist can *like*, *hate*, or *not care* about a food object. There are two outcomes, either the protagonist is given or is not given the food item. Subjects are asked to predict how the protagonist will feel: happy, sad, or 'just ok'. In the second story type, a protagonist does or does not care about another person. There are two outcomes, the other person is sad or happy, and subjects are asked to predict how the protagonist will feel. In both story types, if subjects understand *not caring* then they should predict neutral affect in the protagonist, while predicting an appropriately valued reaction in the other conditions.

Experiment 1

Method

Subjects

Data were collected from 39 randomly selected 3, 4, and 5-year-olds (9 four-year-olds ranging from 41 to 48 months, with a mean age of 45.0 months and a standard deviation of 2.2 months; 13 four-year-olds ranging from 49 to 59 months, with a mean age of 54.1 months and a standard deviation of 3.4 months; 14 five-year-olds ranging from 61 to 72 months with a mean age of 66.9 months and a standard deviation of 3.4 months; 3 six-year-olds ranging from 76 to 79 months with a mean age of 77.3 months and a standard deviation of 1.5 months; six-year-olds were grouped together with the five-year-olds). An additional 13 children were tested but not included in the data due to one of several reasons, including not wanting to participate, a lack of verbal skills (generally due to English as a second language), failure of the control questions, or failure to understand the scales. The children who participated in the study attended daycare facilities and preschools in central New Jersey and came from varying ethnic and social-economic backgrounds. Parents of the children who were tested were required to fill out consent forms, and depending on the number of consent forms returned, a school was either visited once or multiple times. Regardless of the number of visits to any particular school, each child was tested only once.

Design

Every child was given 2 stories (see appendix), each with two conditions, to investigate what they thought about caring and not caring. In the first story there were three factors: getting/not getting, liking/hating/not caring, and age. There were two levels of getting since a character gets or does not get a food object. There were three levels of caring, namely, loving, hating and not caring about a food object. Finally, there were three age groups, 3-, 4-, and 5- year-olds. The first condition of the first story focused on children's intuitions about how someone feels when they get an object that they either love, hate or do not care about. Rather than have children give verbal responses, they had to point to a scale that ranged from a lot of happiness (symbolized with a happy face), to a lot of sadness (symbolized with a sad face), with neutral emotions, that is feeling neither happy nor sad (a neutral face) in the middle. The second condition of the first story was the same as the first except that instead of judging how someone feels when they get an object, it was concerned with how someone feels if they do not get an object that they either like, dislike, or do not care about. Again, a scale was used to measure children's responses.

The next story shifted attention away from caring and not caring about food objects and towards caring and not caring about people. Again, there were three factors: caring/not caring, affect (happy/sad), and age. There were two levels of caring since a character either cared or did not care about someone. There were two levels of affect where a character was either happy or sad. Lastly, the same age groups were used as in the first story. The first condition of the second story involved having children judge how a boy feels about a girl he cares about when the girl is happy or sad. Children's responses were measured by having them point to a new scale that was identical, except for its color, to the scale from the first story. The idea here was that the child would have more interest in seeing a new scale rather than work with the same scale as before. The second condition of the second story paralleled the first conditioned except that instead of determining how the boy feels about the girl being happy or sad when he cares about her, children had to decide how the boy would feel about the girl who was happy or sad when he does not care about her. Again, children's responses were measured on the new scale.

To familiarize the child with the scales, the test stories were preceded by three practice stories where the child was asked to show the experimenter where they should point if they think something tastes good, bad, or just ok.

By asking children about their intuitions regarding how people feel about some combination of getting/not getting some object they like/do not like/do not care about, the first part of this study attempts to gain insight into the child's understanding of caring about objects. The second part of this study is concerned with the child's understanding of caring about people. Children were asked how a person, who either cares or does not care about another person, would feel when that person is happy or sad. In order to learn more about the child's developing concepts of caring and not caring, the focus of Experiment 1 was to determine what relationships exists, if any, within and between caring/not caring about objects and caring/not caring about people.

Materials/Apparatus

A. *Stories*: Each of the stories was acted out using toy dolls. Both conditions of the first story used two boy dolls and one girl doll. Both conditions of the second story used a boy and a girl doll, but the second condition used a different set of dolls. The experimenter moved a doll forward and closer to the child if it was the doll being referenced. The stories were acted out on a tabletop.

B. *Scales*: To measure children's responses on how a character would feel in the context of the stories, a scale was used. The practice trials were run on the "pink" scale. The scale was an 8.5 x 11" piece of pink fluorescent paper. The scale had a straight line drawn horizontally across the center of the paper with 5 evenly spaced markings, indicating different degrees of judgment. Moving left to right, the 5 locations were measures of tasting very bad, testing a little bad, tasting OK, tasting a little good, and tasting very good. The locations were labeled with symbols to help the child remember what the locations were supposed to represent. Having a bad taste was represented by X's, having a good taste was represented by stars, and having an OK taste was left blank in the center of the scale. Depending on the degree of tasting good or bad, the symbols were either singular or multiple (e.g. the "very good" symbol had many stars compared to the "a little good" symbol which had only one star). The first two stories used the "green" scale, which was made by taping together two 8.5 x 11" pieces of fluorescent green paper along the 8.5" side. With a black marker, a line was drawn horizontally through the center of the papers with 5 evenly spaced markings, indicating different degrees of feelings. Moving left to right, the locations were measures of feeling very sad, feeling a little sad, feeling OK, feeling a little happy, and feeling very happy. Again the locations were labeled with symbols to help the child remember what the locations were supposed to represent. The 5 locations were designated with either a happy (two types of happy faces), sad (two types of sad faces) or plain (i.e. neutral; one type of neutral face) face. Depending on the degree of happiness or sadness, the mouths on the faces either had a slightly curved smile or frown, or a considerably curved smile or frown, respectively (e.g. the "very happy" symbol had a considerably curved smiley face compared to the "a little happy" symbol which had a slightly curved smiley face). The neutral face had a straight line for a mouth, indicating that this face was neither happy nor sad. For the last two stories, a new "blue" scale was introduced, however this scale was identical to the previously described "green" scale except that fluorescent blue paper was used.

Procedure

Children were tested individually. An undergraduate assistant brought each child to a separate room or to quiet part of the classroom. People present in the room during the study were the author, a post-doctoral student, between one to three undergraduate assistants, and sometimes a faculty member of the daycare or school. A camera was set up before testing began such that the front of the child was being videotaped. Each child was videotaped during the test trials. The child sat next to the experimenter at a table where the pink scale was placed in front of the child. To make the child more comfortable talking, the experimenter introduced himself and asked the child easy questions such as if they liked stories, how old they were, and what their favorite story was. After some talking, the experimenter told the child that they were going to get started. The experimenter first introduced the child to the practice scale by pointing to and explaining three of the five different locations. The experimenter pointed to the "tastes very good" location first, the "tastes very bad" location second, and the "neutral taste" location third. Children were also told though that they could point to the two intermediate locations if they thought something tasted just a little good or just a little bad. To make it more transparent and interesting, the task was made interactive by giving the child an opportunity to practice using the scale. They were asked a series of questions about foods that they liked, disliked and did not care about. The experimenter gave the child examples (e.g. ice cream for "tastes good," bugs for "tastes bad," and water for "tastes ok") that they could use about what they liked, disliked and thought was OK tasting. However, children were allowed to pick their own examples if they did not want to use the experimenter's examples. If a child could not give 3 responses or use the suggestions given by the experimenter, the child was thanked for participating and was given a sticker as well as a certificate of participation. For children who successfully gave or used 3 responses, they were given the test stories next.

The practice scale was then removed and the child was told that he/she was now going to use a new scale. To introduce the green scale, the experimenter told the child what the locations meant. For example, children were told that the face on the right end of the scale, a happy face with a considerably curved smiley face, meant that someone feels happy, and that when someone feels happy, they should point there, as the experimenter placed his finger on the happy face. This was done for 3 of the locations, with the experimenter pointing to the very happy face first, the very sad face second, and the middle, neutral face third. The experimenter transitioned to the test stories where the child was introduced to three toy dolls, two boys and a girl, taken from a box on the ground near the experimenter. The characters were all initially standing next to each other. The child was then told how each character felt about peanut butter and jelly sandwiches: one liked, one disliked, and one did not care about peanut butter and jelly sandwiches. In the first story, the child is told that the characters are at a party and they get the sandwiches. The experimenter then, one at a time, moved one of the characters closer to the child and asked how the character feels about getting the sandwich. The child was then directed to report their answer by pointing to the scale. The characters were placed back so that they were standing next to each other again for the next condition. Next the second condition was given about the same three characters, except this time the child is told that they do not get the peanut butter and jelly sandwiches. The child is asked how each character feels about not getting the sandwiches after the experimenter moved the character out of line and closer to the subject. The experimenter recorded the child's responses for all tests on a score sheet as the child

reported his/her answers. After this story, the three characters were put away and the green scale was removed.

The blue scale was introduced along with two new characters, a boy and a girl, for the next two conditions of the second story. The child saw a scene acted out with the dolls as he/she was told the story. In both the first condition, where a boy cares about a girl, and the second condition, where a boy does not care about a girl, the child saw two scenes in each. In the first scene of each condition, the two dolls are just standing next to each other as the experimenter tells the child that the girl is happy. The child is then directed to point to the scale to describe how the boy feels about that. However in the second scene of both conditions, the experimenter acted out with the props the girl tripping and falling, to show that the doll was now sad. The child is again directed to the scale to show how the boy feels when the girl is sad.

Children were thanked for listening to the stories and were given both a sticker and a certificate of participation, and then were brought back to class by an undergraduate assistant.

Results

Data for the Object Caring stories were analyzed separately from the Person Caring stories. Subjects' judgments were given scores from -2 to 2 depending on whether they rated the story character's emotional state as sad (negative values) or happy (positive values), with zero representing neither happy nor sad but "just ok". Table 1 shows the mean scores by age group and condition with standard error in parentheses. These data were entered in a Get condition (repeated measures), (2) x Care condition (repeated measures), (3) x Age (3) ANOVA (see Table 1). The Get condition was highly significant as main effect ($F_{1, 36} = 17.0$, $p < 0.001$). The Care condition was also significant as a main effect ($F_{2, 72} = 3.8$, $p = 0.027$). There was no significant main effect of Age ($F_{2, 36} = 1.8$, $p = 0.19$). The Get x Care interaction was highly significant ($F_{2, 72} = 54.7$, $p < 0.001$). Figure 1 graphs children's scores on the Get/No Get stories by Care conditions (Like, Hate, Don't Care) collapsed across age. This shows the form of interaction to consist of the predicted switch in valence between Like and Hate according to whether the object was obtained or not, together with neutral valence for Don't Care in both Get and No Get stories. Finally, a significant 3-way interaction of Get x Care x Age was found ($F_{4, 72} = 2.66$, $p = 0.04$). No other effects were significant.

Table 1

Age	Gets	Cares	Mean (std error)	
3	1	1	1.44 (036)	
		2	-0.56 (036)	
		3	-0.11 (0.34)	
	2	1	-0.89 (0.35)	
		2	-0.44 (0.53)	
		3	-0.67 (0.33)	
4	1	1	1.31 (0.30)	
		2	-0.77 (0.30)	
		3	0.31 (0.29)	
	2	1	-1.31 (0.29)	
		2	0.23 (0.44)	
		3	0.00 (0.27)	
5	1	1	1.41 (0.26)	
		2	-1.53 (0.26)	
		3	0.18 (0.25)	
	2	1	-1.71 (0.25)	
		2	0.65 (0.38)	
		3	0.18 (0.24)	

Legend
Gets
1 - get object
2 - does not get object
Cares
1 - loves object
2 - hates object
3 - does not care about object

The 3-way interaction was investigated by way of a post-hoc Bonferroni analysis on Get/No Get difference scores. This showed that 3-year-olds had different scores on the Hate stories from the 5-year-olds ($p = 0.048$). No other differences were significant. Inspection of the means shows that judgments across age groups on the other stories were similar, with all ages showing positive scores for Like-Get, negative scores for Like-No Get, and neutral scores for both Don't Care-Get and No Get. Likewise, all groups produced negative scores for Hate-Get. However, the 3- year-olds also gave negative scores for Hate-No Get, whereas the older children gave this a positive score.

Figure 1

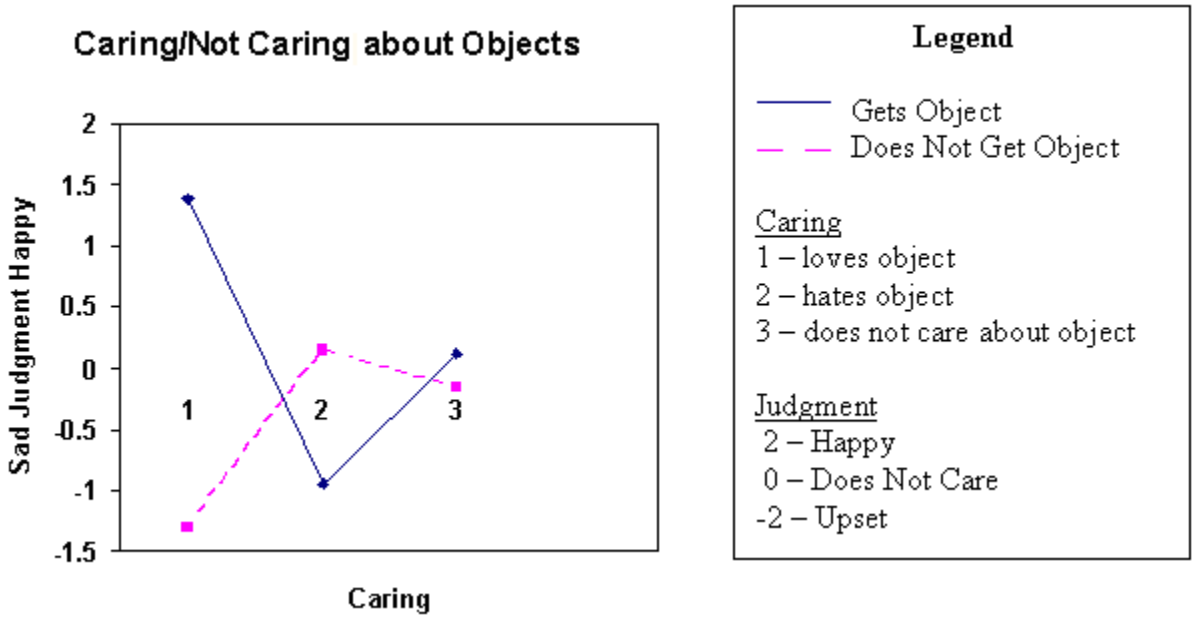


Table 2 shows mean scores by age and condition for the Person Caring stories, with standard errors in parentheses. These data were entered into a Care (repeated measures) (2) x Affect (repeated measures) (2) x Age ANOVA. There was no significant effect of age ($F_{2,36} = 2.2$, $p = 0.13$) and no significant age interactions (all F's < 1). Care was not significant as a main effect (F < 1). Affect had a significant main effect ($F_{1, 36} = 11.0$, $p = 0.002$) and Care x Affect interaction was also significant ($F_{1, 36} = 8.4$, $p = 0.006$). Figure 2 graphs subjects' scores by Affect and Care conditions collapsed across ages. It can readily be seen from Figure 2 that the Care x Affect interaction largely subsumes the Affect main effect. Children judgments showed contrasting valence across the Care stories and neutral valence in Don't Care stories.

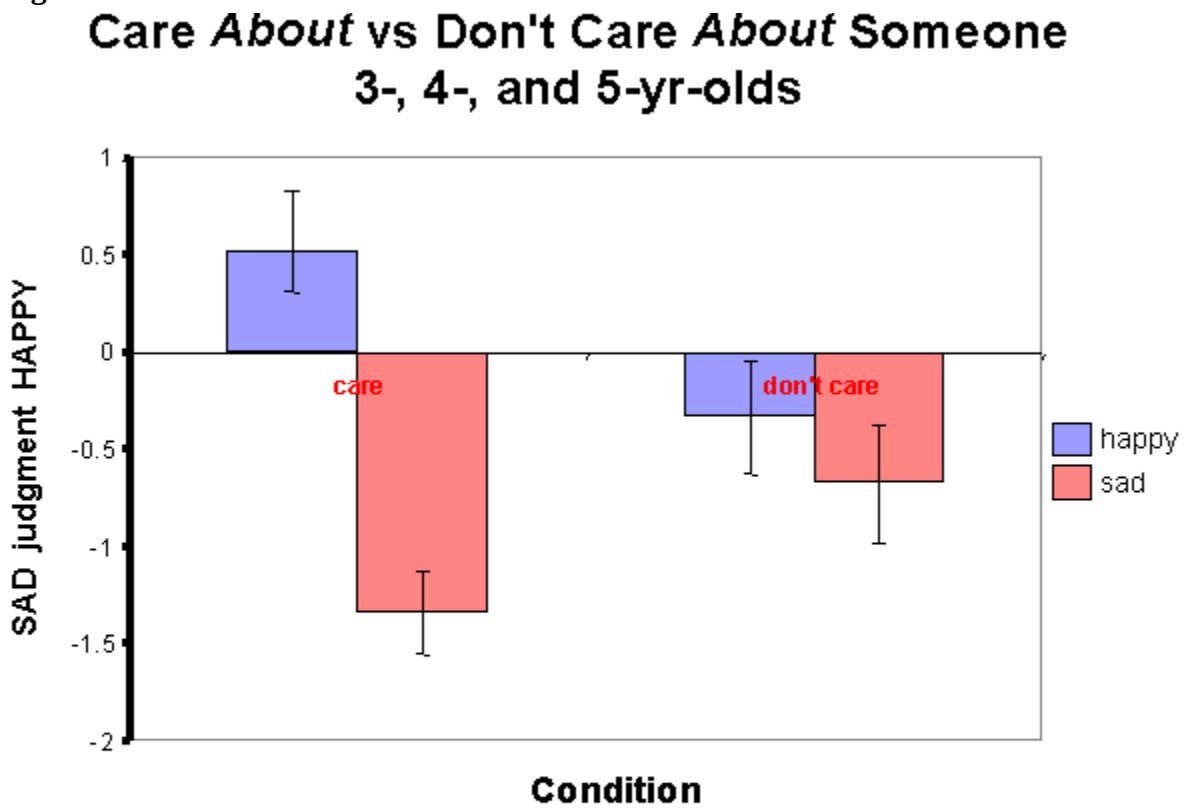
Table 2

Care	Affect	Mean (Std. Error)	
1	1	0.46 (0.26)	
	2		-1.28 (0.18)
2	1	-0.18 (0.24)	
	2		-0.51 (0.24)

Legend
Cares
1 - cares about person
2 - does not care about person

Affect
1 - Happy
2 - Sad

Figure 1



AFFECT: $F_{1,35} = 10.9, p < 0.01$
AFFECT x CARE: $F_{1,35} = 7.3, p = 0.01$
AGE: $F_{3,35} = 1.6, p = 0.2$, n.s.

Discussion

In the object caring stories of Experiment 1, there was a highly significant interaction between getting/not getting an object and caring (loving, hating, not caring) about an object. These results suggest that children take into consideration whether someone loves, hates or does not care about an object when deciding how someone feels about getting or not getting an object. As expected, children said that when someone loves an object, they will be happy when they get it, and upset when the do not get it. The opposite was true when someone did not get an object. Children said someone will be upset if he/she gets an object they do not like and happy if they do not get an object that they do not like. In the case where the child was told that someone does not care about an object, children said that a character would be neither happy nor sad whether or not they got the object.

There were also primary effects for both getting and caring about an object. The size of the primary effects seems to suggest that getting or not-getting alone influenced children's decisions

about how someone would feel. For example, someone will be happier if he gets something, even if he does not like or does not care about what he is getting. Similarly, caring alone appeared to effect children's decisions about how someone feels. Their responses suggested they believed that someone will be happy whether or not he/she gets something he/she like and upset whether or not he/she gets something he/she does not like. It is likely that these effects are merely superficial, and in fact are really just being carried by the interaction effect between caring and getting.

In addition, the data showed there was an age effect, namely that the care-getting interaction strengthened as children got older. There is most likely a larger degree of noise with the younger subjects, accounting for the weaker pattern.

Furthermore, the 3-year-olds showed an interesting response pattern that differed from responses of the older children. In the case where a person does not get something he/she does not like, one would expect subjects to say that the person would feel happy. However 3-year-olds said that the person would be slightly unhappy. One possible reason for this is that maybe they focus on only caring or getting, but not both. For example, the fact that the person hates something makes the person unhappy or the fact that he/she is not getting anything at all might make him/her unhappy. However, by not putting the two pieces together, the children are not coming up with the expected response that the person would feel happy. It might also be the case that these younger subjects have not developed an ability to think about counterfactuals. It could be argued that they do not say someone feels happy because they do not consider what the case would be if the person did in fact get the object that they do not like. Once considering this outcome, a person would be happier to not receive something they do not like.

One other interesting result in the data is younger children's response to when the person who does not care about the object does not get the object. Children responded by saying that the person would feel slightly upset, when the expected answer was that they feel neither happy nor sad, but just ok. It is possible that children again are relying only on the fact that the person is not getting something and will be upset, or it may reflect the fact that children are fatigued at this point in this story. The first part of this study supports the idea that children do understand caring and not caring about objects.

The data from the caring about people stories also revealed some significant effects. First, there was an affect main effect, where children's responses indicated that whether the girl was happy or sad influenced if they said the boy would be happy or sad, regardless of whether or not he cared about the girl. This effect however seemed to be carried by an interaction effect between affect and caring (or not caring) about a person.

The results suggest that children use both affect and caring information when determining how someone will feel. This was expected and supported the hypothesis that whether or not a person cares about someone will affect children's responses of how that person will feel based on the other person being happy or sad. The results indicated that when the boy does not care about the girl, children are likely to say that the boy is neither happy nor sad when the girl is either happy or sad. That is, affect only affects you if you care. Children responded by saying that the boy will be sad if the girl they care about is sad, but there was only a weak consensus

that the boy will be happy if the girl they care about is happy. This seems to reflect an 'honest' response of children that is perhaps less readily expressed by adults. When someone else is happy, adults act as if they are happy for that person, but generally it is probably more of an act of politeness then sincere happiness. Age collapsed across age groups, indicating no age effect. Even young 3-year-olds share similar feelings in regard to caring about people as adults do.

Given these positive results on understanding not caring, we can proceed to test preschooler's judgments of intention. Will preschool children make moral-based judgments of intention, as adults do in Knobe scenarios, or will preschoolers make only fact-based ('theory of mind') judgments of intention?

Experiment 2

Method

Subjects

The subjects were 120 children enrolled in preschools and daycare programs through out central New Jersey. There were 26 children that were three-years-old (ranging from 39 to 48 months, mean age 43.1 months, SD = 3.1 months), 47 children that were four-years-old (ranging from 48 to 59 months, mean age 53.6 months, SD = 3.6 months), 39 children that were five-years-old (ranging from 60 to 72 months, mean age 65.2 months, SD = 3.7 months) and 8 children that were six-years- old (ranging from 72 to 80 months, mean age 74.1 months, SD = 2.7 months). Another 8 children were tested but not included in the data for one of several reasons, most commonly because of failing the basic control questions (as opposed to the caring control question), but also because of shyness and in one case, a mistake by the experimenter. Subjects were randomly assigned to one of two conditions, either the upset condition or the happy condition. For three-year-olds, 13 were in the upset condition, and 14 were in the happy condition. There were 26 four-year-olds in the upset condition, and 23 four-year-olds in the happy condition. 19 five-year-olds were in the upset condition and 22 five-year-olds were in the happy condition. Finally, for six-year-olds, there were 4 in the upset condition, and there were 4 in the happy condition.

Design

Children were given one of two conditions (see appendix). In the upset condition, a boy does not care that a girl will get upset and performs an action that makes the girl upset. Children have to make a moral judgment about whether the boy makes the girl upset on purpose or not, and they have to successfully attribute the mental state of not caring to the boy. Control questions were given to confirm that children were remembering the story. The first set of control questions asked about how both the boy, who loves frogs, and girl, who hates frogs, each felt about a frog. Keeping the child involved in the story, they were then asked why they thought the girl would get upset if the boy brought the frog over. If a child did not give the correct answer or if they gave no answer, the experimenter told them the correct response. Before moving on with the story however, the experimenter went over this part of the story one more time with the child to make sure they understood why the girl would get upset. In anticipation of a control question

that was giving younger children a difficult time in pilot studies, the experimenter made a disclaimer to each subject to listen carefully to what the boy said next. To check if the child followed this complicated portion of the task, the experimenter slowly and clearly asked the child a control question about whether the boy cares that the girl will get upset. If the child answered this "care" control question incorrectly, the experimenter said, "ok, lets do the story again," and started the story over from the beginning. If a child failed this "care" control question a second time, the experimenter made a note on the score sheet. This "care" control question requires the child to use their "theory of mind" in that they have to attribute a mental state to the boy. The child is finally asked the test question at the end of the story, requiring them to make a moral judgment. Instead of using the word "intentionally," the child was asked if an action was done "on purpose," because children find the expression "on purpose" easier to comprehend than "intentionally."

In the happy condition, the boy does not care that the girl will be happy and does something that makes her happy. Children have to make a moral judgment in deciding if the boy makes the girl happy on purpose or not on purpose as well as attribute a mental state, namely, "not caring she will be happy" to the boy. Children were asked the same line of control questions, gauging how well they were following the details of the story. The experimenter slowly and clearly asked the child the "care" control question about whether the boy cares that the girl will be happy. Again, this "care" control question requires the child to use a theory of mind in that they he/she has to attribute a mental state to someone else. The child is finally asked the test question at the end of the story, requiring them to make a moral judgment.

Materials/Apparatus

Two child toy dolls were used to act out the stories. One of the dolls was a boy and the other was a girl. A toy frog was used when telling the story, which belonged to the boy. The stories were acted out on a tabletop.

Procedure

An undergraduate assistant individually brought children to the testing area, which was either a separate room or a quiet part of the classroom. The child was asked to have a seat next to the experimenter. A camera was set up facing towards the front of the child as an undergraduate assistant videotaped the story. The videotaping was for data purposes. Before jumping into the story, the experimenter talked to the child first, asking him/her questions such as if they liked hearing stories, how old they were, etc, just to get them talking. The experimenter then proceeded to tell the child that he/she they was going to hear a story. Using two toy dolls, a boy and a girl, and a toy frog, the experimenter told the story while acting it out with the props. At the start of the story, the boy and girl are standing apart from each other and the child is required to use their imagination as they are told that the boy and girl are at their respective houses. The experimenter than took out a frog from a basket on the ground that was not visible to the child, and asked the child what the boy had with him, as the frog was placed next to the boy. The experimenter proceeded with the story with the props in the same position. After the child answered the "care" control question, the child was then shown the boy bringing the frog over the girl's house by moving the boy doll and toy frog next to the girl doll. Finally, the

experimenter asked the test question, that is, was the action done on purpose. The experimenter recorded the subject's responses on a score sheet immediately after the response. The child was told that he/she did a great job and was asked if he/she would like a sticker. Each chiled was also given a certificate to bring home, recognizing participation in the study.

If a child failed the "care" control question the props were taken off the table, and the boy and the girl were re-introduced as the experimenter started from the beginning of the script.

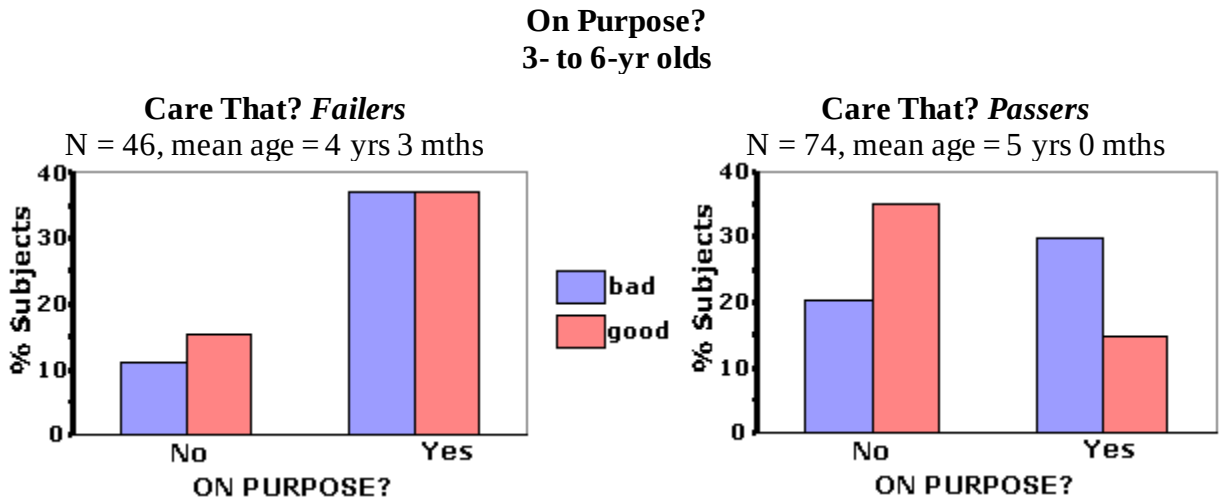
If a child failed the "care" control question a second time, the experimenter went on with the story as if the subject answered it correctly, and recorded their response to the test question. However, it was also recorded that the child failed the control question.

If a child failed any of the earlier control questions before the "care" control question, the experimenter helped the child by either repeating the part of the story relevant to the control question, or starting the story over. If a child failed an earlier control question twice, was too shy, or the child no longer wanted to hear the story, the story was stopped. The child was given a sticker and a certificate, and was thanked for participating in the study.

Results

Subjects were categorized as having said "yes" or "no" to the On Purpose question across Good effect/Bad effect stories. Preliminary analysis showed that a substantial number of subjects failed the Care That control question. Data were divided according to passing or failing this control question and analyzed separately. 46 out of 120 subjects (38%) failed the control question while 74 (61.6%) passed. Ages of passer/failers were compared by t-test, which showed a significant difference between the two groups ($t_{118} = 5.2$, $p < 0.001$, two-tailed). The average age of the failers was 4 years and 3 months while the average age of the passers was 5 years 0 months. For the On Purpose question, control failers responded similarly across the Happy/Upset conditions (&KHgr;$^2 < 0.1$, n.s.) with an overall bias to respond "yes" more often than "no", 74% vs. 26%, respectively (two-tailed Binomial probability = 0.002). Control passers showed a quite different pattern (see Figure 3). There was no over all bias for passers (44.6% "yes" vs. 55.4% "no") (two-tailed Binomial = 0.4, n.s.). Instead, passers tended to answer, "yes," to the On Purpose question when the side effect was bad (59.5% "yes" vs. 40.5% "no") and, "no," when the side effect was good (29.7% "yes" vs. 70.3% "no") (Upton's &KHgr;$^2 = 6.53$, $p = 0.005$, one-tailed).

Figure 3



Discussion

Subjects were split into two groups, those who passed the care control question and those who failed it. There was a significant difference between the ages of the passers and failers, with failers on average being younger by nine months. The difference in age may be due to the fact that the syntax of the control question is too complicated for some of the younger subjects. The care control question contains a propositional attitude with a complementizer that takes on a proposition, "…cares that p…" known as an embedded proposition. It is known that younger children have a hard time processing sentences with embedded propositions (Taylor, 1988).

Failers did not show the adult response pattern. Interestingly, there was a substantial effect where they showed an overall "yes" bias to both the happy and upset conditions. Considering that information from the care control question showed that these subjects believed that the boy cares about the relevant outcome, that is, he cared about making her upset, this could be interpreted as meaning that he wanted to upset the girl on purpose in the upset story. Similarly, in the happy story, children answer the control question believing that the boy cares that the girl will be happy. Therefore, it seems possible that they would say the boy makes the girl happy on purpose, simply because he cares about doing so.

Subjects who passed the care control question did show the adult response pattern. There was a predicted bias that there would be more "yes" answers in the upset condition and more "no" responses in the happy condition. This prediction was supported by the data. These subjects said that an action is done on purpose when a person does not care that the action will bring about a negative outcome. However, like adults, children switched their response when a person does not care that an action will cause a positive outcome. Here subjects say that the action was not

done on purpose. For children who understand caring and not caring, it appears that judgments of intention are being driven by moral judgments.

General Discussion

From this study, there is strong evidence that children begin to make moral-based intention judgments, at the latest, by the fifth birthday. The results suggest that children with an understanding of *not caring* consider an outcome as intentional or unintentional according to the moral valence of the outcome when making a moral judgment. Consistent with findings by Shultz (1986) that suggest children have a variety of moral concepts in place at 5-years of age, the 5 year mark in child development would seem to be the limiting factor in children's ability to make moral-based intention judgments since theory of mind ability is acquired at the latest, by the fourth birthday. A follow up study in this regard is to try to simplify the task and see if younger children also show moral-based intention judgments.

The task could be simplifed by substituting "*care that*" with "*care about*" in the screening care control question. Younger children tend to fail to understand *care that* but do understand *care about*. It is expected that children will have an easier time dealing with the care about phrase since it has no embedded propositions, which have previously been shown to be difficult for younger children to process. However one linguistic complication that arises with this substitution is whether or not the meaning changes if the phrases are swapped. In addition, this changing of phrases causes additional changes to the original care control question, since the proposition p in *care that* cannot be directly placed into the *care about* phrase without moving words around. Rather than get caught up in the linguistic debate over modulation, the idea that words in a sentence affect the meaning of other words in the sentence, the concern here focuses on cognitive development. Given that "care" is the prepositional attitude component of a mental state, understanding of "care" relies on theory of mind ability. Thus the difference in young children's understanding of *care that* and *care about* probably reflects a development in the theory of mind domain.

As soon as the appropriate theory of mind development regarding *care that* has taken place, children show the adult moral-based judgment of intention, without further development being necessary. Given that an actor does not care about an outcome, children show the crossover effect of saying that an outcome was intended if it was harmful but unintended if it was beneficial. The results of the study suggest that children begin to show the asymmetric, adult-like response patterns when making moral judgments once they have a hold of *care that*. Before then, they demonstrate an overall "yes" bias to an outcome being intentional regardless if it is harmful or beneficial, possibly because they interpret *cares that* to mean that an actor wants to upset or make another person happy. If this is the case, they may be employing fact-based, theory of mind judgments of intention. A noteworthy extension of this research would be running these studies on autistic populations. Considering that children with autism have a deficit in theory of mind, it would be interesting to see if they perform like the failers in this study, indicating the significance of a lack of mental state attribution, or if they would show the adult response patterns, indicating that they are still capable of making moral based judgments of intention even without a normal theory of mind. This study would give insights into the

degree to which fact-based intention judgments are used compared to moral based intention judgments in a moral context.

The findings here suggest that theory of mind and moral judgment are separate but interacting domains. *Caring/not caring* attributions occur independently from moral judgment decisions, but both pieces of information are used in determining if an outcome of an act was intentional. That is, children use their theory of mind ability to understand not caring which interacts with their judgments of moral valence of an outcome to decide whether an action was performed intentionally or unintentionally. The result is asymmetric intention judgments.

References

Astington, Janet W. Children's comprehension of expressions of intention. [Journal Article] *British Journal of Developmental Psychology. Vol 4(1) Mar 1986, 43-49. British Psychological Society, United Kingdom*

Astington, Janet W; Gopnik, Alison. Developing understanding of desire and intention. [Chapter] *Whiten, Andrew (Ed). (1991). Natural theories of mind: Evolution, development and simulation of everyday mindreading. (pp. 39-50). xiv, 362pp.*

Bartsch, Karen; Wellman, Henry M. Children talk about the mind. [Authored Book] (1995). *viii, 234pp.*

Baron-Cohen, S., Leslie, A.M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21, 37-46.

Darley, John M; Shultz, Thomas R. Moral rules: Their content and acquisition. [Journal Article] *Annual Review of Psychology. Vol 41 1990, 525-556. Annual Reviews, US*

Duska, Ronald; Whelan, Mariellen. Moral development: A guide to Piaget and Kohlberg. [Book] *(1975). 128pp.*

Feinfield, Kristin A., Patti P. Lee, Eleanor R. Flavell, Frances L. Green, John H. Flavell, Young Children's Understanding of Intention, Cognitive Development 14 (3) (1999) pp. 463-486.

Flavell, J.H., & Miller, P.H. (1998). Social Cognition. In W. Damon (Series Ed.) & D. Kuhn & R.S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp.851-898). New York: Wiley.

Gilligan, Carol. In a different voice: Psychological theory and women's development. [Authored Book] *(1982, reprinted 1993). xxx, 184pp.*

Kohlberg, Lawrence (1981) Essays on Moral Development: Vol. I. *The Philosophy of Moral Development* (San Francisco: Harper and Row).

Knobe, J. (2003). "Intentional Action and Side Effects in Ordinary Language." *Analysis*, 63, 190- 193.

Lapsley, Daniel K. Moral psychology. [Authored Book] (1996). *xiv, 289pp*.

Leslie, A.M. (2000b). 'Theory of mind' as a mechanism of selective attention. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences*, 2nd Edition. (pp. 1235-1247). Cambridge, MA: MIT Press.

Meltzoff, Andrew N; Gopnik, Alison; Repacholi, Betty M. Toddlers' understanding of intentions, desires and emotions: Explorations of the dark ages. [Chapter] *Zelazo, Philip David (Ed); Astington, Janet Wilde (Ed); et al. (1999). Developing theories of intention: Social understanding and self-control. (pp. 17-41). xi, 358pp.*

Nelson, Sharon A. Factors influencing young children's use of motives and outcomes as moral criteria. [Journal Article] *Child Development. Vol 51(3) Sep 1980, 823-829. Blackwell Publishers, US*

Nucci, Larry P; Nucci, Maria S. Children's social interactions in the context of moral and conventional transgressions. [Journal Article] *Child Development. Vol 53(2) Apr 1982, 403-412. Blackwell Publishers, US*

Piaget, J. (1977) *The Moral Judgment of the Child*. London: Penguin (originally published 1932).

Pratt, Chris; Bryant, Peter. Young children understand that looking leads to knowing (so long as they are looking into a single barrel). [Journal Article] *Child Development. Vol 61(4) Aug 1990, 973-982. Blackwell Publishers, US*

Premack, David; Woodruff, Guy. Does the chimpanzee have a theory of mind? [Journal Article] *Behavioral & Brain Sciences. Vol 1(4) Dec 1978, 515-526. Cambridge Univ Press, US*

Roth, D., & Leslie, A.M. (1998). Solving belief problems: Toward a task analysis. *Cognition*, 66, 1-31.

Scholl, B.J., & Leslie, A.M. (1999b). Modularity, development and 'theory of mind'. *Mind & Language*, 14, 131-153

Shantz, C. (1983). Social Cognition. In J.H. Flavell & E. M. Markman (Eds.), *Handbook of Child Psychology, Vol. 3: Cognitive Development*, 4th edition. (pp.495-555). New York: Wiley.

Shultz, Thomas R; Wright, Kevin; Schleifer, Michael. Assignment of moral responsibility and punishment. [Journal Article] *Child Development. 57(1) Feb 1986, 177-184. Blackwell Publishers, US*

Shweder, Richard A; Mahapatra, Manamohan; Miller, Joan G. Culture and moral development. [Chapter] *Kagan, Jerome (Ed); Lamb, Sharon (Ed). (1987). The emergence of morality in young children. (pp. 1-83). xx, 354pp.*

Smetana, Judith G; Braeges, Judith L. The development of toddler's moral and conventional judgments. [Journal Article] *Merrill-Palmer Quarterly. Vol 36(3) Jul 1990, 329-346. Wayne State Univ Press, US*

Surian, L., & Leslie, A.M. (1999). Competence and performance in false belief understanding: A comparison of autistic and three-year-old children. *British Journal of Developmental Psychology*, 17, 141-155.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development*, **59**, 703-718.

Turiel, E. 1983. *The Development of Social Knowledge: Morality and Convention*. New York: Cambridge University Press.

Wimmer, Heinz; Perner, Josef. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. [Journal Article] *Cognition. Vol 13(1) Jan 1983, 103-128. Elsevier Science, Netherlands*

Appendix A

Experiment 1

Practice Story

Today I have a couple of stories for you. But before we get to the stories I want to show you something. I want to tell you about our pink scale. Now our pink scale tells us about things that are good, things that are bad, and things that are just ok. Over here we have stars. Stars mean something is good. So if you think something is good, I want you to point to the stars. Over here we have x's. X's mean something is bad. So if you think something is bad, I want you to point to the x's. Over here we have nothing. Nothing means something is just ok. So if you think something is ok, I want you to point to over here. Let's try it out.

What do you think about eating ice cream? Do you think ice cream tastes good to eat, do you think ice cream tastes bad to eat, or do you think that ice cream tastes just ok to eat? Do you think ice cream tastes a little good to eat, or do you think ice cream tastes very good to eat?

What do you think about eating bugs? Do you think bugs taste good to eat, do you think bugs taste bad to eat, or do you think that bugs taste just ok to eat? Do you think it tastes a little bad to eat bugs or do you think it tastes very bad to eat bugs?

What do you think about drinking water? Do you think water tastes good to drink, do you think water tastes bad to drink, or do you think that water tastes just ok to drink? (What do you think is something that tastes just ok to eat? Ok, so where would you point if I asked what do you think about ____?)

This is our other scale, our white scale. Our pink scale told us about things that are good, things that are bad, and things that are just ok. This one tells us about how people feel. Over here we have a happy face. A happy face means someone feels happy. So if you think someone feels happy, I want you to point here. Over here we have a sad face. A sad face means someone feels sad. So if you think someone feels sad, I want you to point here. Finally, over here we have a plain face. A plain face means someone feels just ok. They are not happy and they are not sad. They feel ok. So if you think someone feels just ok, I want you to point over here.

Story 1 - Object Task

Condition 1 - Gets food object

Look at these three kids! This is Ryan, this is Helen, and this is Pablo. They are at a birthday party.
And look what they get. They each get a peanut butter and jelly sandwich.

Now Ryan **loves** peanut butter and jelly sandwiches. So when he gets a sandwich, how does that make him feel?

And Helen **hates** peanut butter and jelly sandwiches. So when she gets a sandwich, how does that make her feel?

And Pablo **doesn't care** about peanut butter and jelly sandwiches. [He doesn't love them and he doesn't hate them]. So when he gets a sandwich, how does that make him feel?

[He'll feel OK, not happy, not sad, just OK]

Condition 2 - Does not get food object

Now they are at a different birthday party, and look what happens.

Now Ryan loves peanut butter and jelly sandwiches. So when he does not get a sandwich, how does that make him feel?

And Helen hates peanut butter and jelly sandwiches. So when she does not get a sandwich, how does that make her feel?

And Pablo doesn't care about peanut butter and jelly sandwiches. [He doesn't love them and he doesn't hate them]. So when he does not get a sandwich, how does that make him feel?

[He'll feel OK, not happy, not sad, just OK]

Story 2 - Person Task

Condition 1 - Cares about

Look this is Jon and this is Sarah. And Jon really cares about Sarah.

And right now Sarah is happy.
So how does Jon feel about that?

But then Sarah trips and falls, and now she is sad.
So how does Jon feel about that?

Condition 2 - Does not care about

Now here are two other kids. This is Justin and this is Nela. And Justin **doesn't care** about Nela. He doesn't like her and he doesn't dislike/hate her, he just doesn't care about her.

And right now Nela is happy.
So, how does Justin feel about that?

Remember, Justin doesn't care how Nela feels.

But then Nela trips and falls, and now she is sad.
So how does Justin feel about that?

Experiment 2

Story 1 - Upset Condition

Here is a boy named Andy, and he's over at his house (put him on one side of the table). And here is a girl named Janine, and she's over at her house (put her on the other side of the table). And look what Andy has with him, he has a [frog - let child answer]. Now Andy *loves* frogs, but Janine *hates* frogs.

Now can you remember, does Andy love frogs? [Yes.]
Does Janine love frogs? [No.]

Andy wants to bring the frog over to Janine's house, but she will get upset.

Why will she get upset? [Because she hates frogs.]

Now listen very carefully. Andy *does not care* that Janine will get upset. He is going to bring the frog over anyway.

Does Andy *care* that Janine will get upset? [No.]

So Andy brings the frog over to Janine's house and she gets upset.

Now I have a question for you. Does Andy make Janine upset on purpose?

Story 2 - Happy Condition

Here is a boy named Andy, and he's over at his house (put him on one side of the table). And here is a girl named Janine, and she's over at her house (put her on the other side of the table). And look what Andy has with him, he has a [frog - let child answer]. Now Andy *loves* frogs, and Janine *loves* frogs.

Now can you remember, does Andy love frogs? [Yes.]
Does Janine love frogs? [Yes.]

Andy wants to bring the frog over to Janine's house. If he brings the frog over, she will be happy.

Why will she be happy? [Because she loves frogs.]

Now listen very carefully. Andy *does not care* that Janine will be happy. He is going to bring the frog over for himself.

Does Andy *care* that Janine will be happy? [No.]

So Andy brings the frog over to Janine's house and she is happy.

Now I have a question for you. Does Andy make Janine happy on purpose?