Preschoolers' appreciation of uncommon desires and subsequent emotions

Carolien Rieffe*, Mark Meerum Terwogt, Willem Koops, Hedy Stegge and **Annemiek Oomen**

Department of Developmental Psychology, Free University, Amsterdam, The Netherlands

Many theory of mind researchers have argued that even preschoolers understand the causal relationship between desires and emotions: the fulfilment of a desire results in a positive emotion, whereas its frustration elicits a negative emotion. Children can acknowledge this simple link between desires and emotions, even when their own desire differs from that of the story protagonist. However, in this paper we argue that under some conditions preschool children will not base their emotion predictions for another person on the basis of the other person's desires. In the first experiment, 3- to 5-year-old children were tested for their understanding of desires, when the protagonist's desire for a snack increasingly conflicted with their own preference for a snack. Only the 4-year-olds performed as expected: they gave more accurate emotion predictions when the distance between their own preference and the protagonist's desire was reduced. When the snacks were replaced by toys, however, preschool children showed a bias in their emotion predictions that seemed gender related. The second experiment confirmed that sex-stereotyped beliefs about desirability biased children's predictions of others' emotions: 4- and 5-year-old children were more accurate in their predictions when the protagonist had a traditional desire (a girl wanting to play with a doll), than when the protagonist had a non-traditional desire (a boy wanting to play with a doll), irrespective of children's own preferences for one toy over the other. In sum, evidence was found for two biasing influences in children's understanding of others' emotions: (1) an increased distance between the protagonist's desire and participants' own desires, and (2) beliefs about desirability based on, for example, cultural-norms for gender related preferences, which increases with age.

'Emotions are nonrational insofar as the desires that underlie them are nonrational, which can be argued for all desires' Frijda (1995, p. 1) argued. Nevertheless, the causal relationship between desires and the emotions is rational and, thus, emotions that follow desires can be derived by means of reason. Consequently, an open-minded assessment of others' subjective desires is essential for understanding emotions.

The aim of desires is not to represent reality, as is the case with beliefs. On the contrary, 'the world has to fit the mind' (Searle, 1983, p. 8), and when the situation

* Requests for reprints should be addressed to Dr Carolien Rieffe, Department of Developmental Psychology, Free University, vd Boechorststraat 1, 1081 BT Amsterdam, The Netherlands (e-mail: c.rieffe@psy.vu.nl).

matches or seems to match the desired state, the desire is met. Perner (1991) points out that 26-month-olds can pursue a goal, such as their wish to clean a black board, and then stop their activities when this goal is met. Thus, 2-year-olds understand when the world fits their mind, in other words, when reality meets their desire. Bartsch and Wellman's (1995) claim that 2-year-olds genuinely refer to their own desires in their spontaneous, everyday speech supports this view. Nevertheless, understanding others' desires might be more demanding, because recognition of a situation that matches another's desired state requires an appreciation of this desired state.

Wellman (1990) showed that children by the age of three have the capacity to predict a protagonist's behaviour based on the protagonist's desire, even when they themselves hold a different desire. He asked participants whether they preferred swimming or playing with a dog. The children were then told that the protagonist had an opposite preference. Almost none of the 3-year-olds had a problem with this 'not-own desire task' and a majority accurately predicted the protagonist's behaviour based on the protagonist's desire. Similarly, Moore, Jarrold, Russell, and Lumb (1995) let participants choose between two stickers of equal attractiveness. Again, they were told that the protagonist preferred the sticker that was not chosen by participants and a clear majority of the 3-year-olds predicted the protagonist's choice based on the protagonist's preference. These results suggest that preschoolers indeed appreciate that other people's desires can be distinct from their own. However, the conclusion that young children acknowledge others' desires was undermined when Moore *et al.* (1995) added a slightly, but fundamentally, different set-up. This time, one sticker was obviously much more attractive than the other. Moore *et al.* made it plausible for the participants that the protagonist would choose the less attractive sticker, because it could be inferred from the story that the protagonist would fear the animal on the more attractive sticker. For example, the protagonist had been frightened by this kind of animal in a previous situation. Contrary to previous findings, two-thirds of the 3-year-olds predicted that the protagonist would choose the sticker they preferred themselves, the more attractive sticker. Four-year-olds produced slightly better results, and almost all 5-year-olds based their prediction of the protagonist's choice on the protagonist's desire.

These results suggest that children before their fifth year appreciate the fact that different people have different desires, as long as these different desires seem reasonable to the children. It seems as if young children fail to acknowledge others' desires in the case of items that the young child perceives as highly undesirable. Or, as Astington and Gopnik (1991) phrased it 'understanding that desires may be different for different people isn't the same as understanding that what is desirable may be different for different people' (p. 43). They argue that young children perceive desirability as an objective characteristic of the world. Perner (1991) also states that desirability is often perceived as objective, despite the common-sense expression that taste is not disputable.

Some experiments appear to offer counter-evidence for the viewpoint that children treat desires as objective. In Repacholi and Gopnik's (1997) experiment, children first saw the experimenter displaying disgust over a bowl of cookies and the experimenter then expressed happiness over a bowl of broccoli. When asked to hand her some, without defining which one, 18-month-olds accurately gave the experimenter some broccoli, whereas 14-month-olds gave her the cookies. The authors explain this difference between these two age groups by arguing that the 18-month-olds attributed a

mental state to the experimenter. Another experiment was performed by Flavell, Flavell, Green, and Moses (1990) in which 3-year-olds admitted that Ellie (a second experimenter) did not like a cookie that they liked themselves. Moreover, 3-year-olds said that Ellie, who ate the cookie in the presence of the participants, thought the cookie was *yucky*. However, it is ambiguous if children attributed desires different from their own in these two experiments. Probably, participants in both experiments responded simply to the experimenter's facial gestures of disgust when she tasted the cookie, without any attention to the relevant desire and thus, without any understanding that what is desirable can be different for different people. As both Lillard (1993) and Currie (1998) have argued, there is a difference between the ability to identify others' mental states and carrying out 'scripted routines'. In other words, although children might react appropriately to observable features such as facial expressions or actions, they may fail to appreciate the underlying mental representations. For example, Lillard (1993) presented preschool children with a protagonist who did not know what birds were and who had never seen a bird. Yet, when the protagonist acted like a bird, most preschoolers affirmed that he was pretending to be a bird.

Moore *et al.* (1995) discussed that young children have difficulties in attributing a mental state to others, 'so long as there is a conflict between their own mental state and that to be attributed' (p. 5). However, it can be argued that Wellman's 'not-own desire' task and Moore *et al.*'s 'equally attractive stickers' condition also created a conflict, although small, because children had expressed their preference for one item over another in both experiments. We suggest that it is not the presence or absence of a conflict, but the *intensity* or *saliency* of the conflict that is the issue here. Or, to formulate this more precisely, the diverging results between Moore *et al.*'s two conditions can be explained in terms of the *distance* that participants experience between the protagonist's desire and their own desire. In the equally attractive stickers condition, the distance between the more desirable and less desirable stickers – as perceived by the participant – was rather small. Consequently, the protagonist's choice for the less desirable sticker was less saliently in conflict with the protagonist's ideas about desirability than in the unattractive sticker condition, where the participant's choice for the unattractive sticker was much more saliently in conflict with the participant's preference. This leads to the assumption that an increasing distance will increasingly cause neglect of the protagonist's desire. Instead, children will increasingly predict the protagonist's behaviour based on what they perceive themselves as desirable.

Children show a response pattern that fits this account when asked to predict others' emotions. Young children predict a protagonist's emotion based on the protagonist's desire (Harris, Johnson, Hutton, Andrews, & Cooke, 1989). They do so, even when protagonists' preferences for one item over another (milk or coke) are alternated. Note that the distance between the desirability of the two items is probably small. Yuill, Perner, Pearson, Peerbhoy, and van den Ende (1996) presented participants with so-called 'wicked' (immoral) desires. For example, the protagonist has the desire to hit another child on the head with a ball. In contrast to the findings of Harris *et al.* (1989), their results suggest that 3-year-olds ignore the protagonist's desire in their emotion prediction, whereas 5-year-olds predicted the protagonist's emotion based on the protagonist's immoral desire. These results suggest that there are specific changes between the ages of 3 and 5. However, it is unclear whether immoral desires can be

equated with desires for milk or coke which were used by Harris *et al.* (1989). It might be, for example, that young children reacted differently to the immoral actions in Yuill *et al.*'s experiment, because of previous socialization experience.

Thus, little is actually known about how emotion predictions are influenced by a(n increasing) distance between a child's ideas about what is desirable and the desires of a protagonist. Our expectation is that with an increasing distance between the protagonist's desire and children's perception of what is desirable, young children will be less inclined to acknowledge the protagonist's desire as the basis of their prediction of the protagonist's emotion. Instead, they will increasingly base their emotion prediction on what they think is desirable. As a consequence of this suggestion, desirability in our experiment will be presented as *relative* (item X is presented as 'more desirable' than item Y), whereas the previous noted studies used *absolute differences* ('desirable' vs. 'undesirable'). Since there is a lack of research as to whether young children understand these relative statements, it is possible that very young children translate these statements in absolute terms. Yet, it is difficult to predict the possible consequences of this translation. 'X is desirable, so Y has to undesirable' and 'X as well as Y are desirable' are both plausible outcomes. This issue will be further taken up in our discussion.

EXPERIMENT 1

(Uncommon desires)

Introduction

Note that children's perceptions of desirability supposedly correspond to their own desires and preferences and so an experiment was conducted in which a smaller or larger distance between the participants' own preference for two items was designed. The preference of the protagonist and the participant for one of these items was either consistent or in conflict. We therefore created four conditions in which the distance between the protagonists' and participants' preferences could be ranked. The largest distance emerged in the most saliently *conflicting* condition, whereas the smallest distance emerged in the most saliently *consistent* condition. Participants were asked for the protagonist's emotional reaction when the protagonist's desire was fulfilled or remained unfulfilled. We tested 3-, 4- and 5-year-olds. Correct emotion predictions were expected to increase with age and, more importantly, when the distance between the protagonist's desire and participants' desire was reduced.

Method

Participants

Participants in this study were 38 3-year-olds (mean age 3;6 years; range 3;0 to 3;11 years); 52 4-year-olds (mean age 4;7 years; range 4;0 to 4;11 years) and 46 5-year-olds (mean age 5;8 years; range 5;0 to 6;4 years). Half of the participants were male, the other half female in each age group. The children were tested in day-care centres and primary schools in the suburbs around Amsterdam, the Netherlands.

Procedure

Children were tested individually in a separate, quiet room in one session of approximately 20 minutes. This session consisted of a 'snack condition' and a 'toy condition'. In both conditions participants were

asked to rank seven items from most to least liked. The items were seven different snacks (piece of chocolate, chips, piece of apple, piece of cucumber, piece of cheese, a potato and rice) and seven different toys (doll, tea-set, Lego, ball, crayons, car and airplane). Half of the participants started with the snack condition, the other half with the toy condition. Following the ranking procedure, the first, third, fifth and seventh ranked item were used in four stories.

The ranking procedure started with a short introduction, after which the experimenter showed two snacks (toys) and asked participants which item she or he preferred to eat (play with). For example 'What do you like to eat most: a piece of apple or a piece of cheese?' The experimenter continued with randomly presenting several possible combinations of two items. She presented different combinations until the complete rank-order of the seven items was clear. The reliability of the obtained order was controlled for by presenting a few extra possible combinations. Children did not experience difficulties with this procedure and there were no inconsistencies. This scale construction was followed by a short conversation about feelings and a warm-up task to familiarize participants with the terms (happy and unhappy) and the pictures (a happy and a sad face) that would be presented later in the session:

> Do you also sometimes feel different from one day to another? Yes? How do you feel today? And how come that you feel . . .? Hm, I can imagine that. Okay, now I'll read you some stories in which a girl or a boy feels happy. But sometimes the girl or the boy doesn't feel happy. Look, here are two pictures that show you how a girl [boy] feels.
>
> Here, the girl feels happy [androgynous picture with happy face], and here the girl [boy] feels . . .? [Child is shown an androgynous picture with sad face]
>
> [Picture of a girl in a coat] Very well, now here comes the first story. This is Daphne. Daphne was given a new coat. Daphne feels very happy. Can you show me the picture that shows how Daphne feels? Very good.
>
> [Picture of a boy at the table with an empty plate in front of him] Now I have another story. This is Jasper. It is Jasper's birthday today and he wants a piece of cake. But look, his plate is still empty. Jasper feels very unhappy. Can you show me the picture that shows how Jasper feels? Okay.

All participants responded accurately to this warm-up task. Following this task the children were presented with four snack stories and four toy stories. Half of the participants started with the snack condition and the other half with the toy condition. Every story contained a combination of two items (a combination of the 1st and 7th ranked items; or the 3rd and 5th ranked items) that the protagonist could receive. The preference of the protagonist for one item over the other was expressed. This preference was consistent with the preference of the participants in one version, and in conflict in the other version. The conflict between preferences of participants and the protagonist was either less salient (5th vs. 3rd ranked item) or more salient (7th vs. 1st ranked item). Four story types with a decreasing distance were thus created: most saliently in conflict; less saliently in conflict; less saliently consistent and most saliently consistent. The order of the four story types was randomized in each condition. The protagonist either received the desired or the less desired item and the emotional outcome was therefore happy or unhappy. This was counterbalanced over story types and varied over participants using a Latin square design. Note that *undesired items* in this experiment are only undesired relative to the favoured items, and are not undesired in an absolute sense. In each condition two of the protagonists were male and two were female. An example of one of the stories:

> [Picture 1] This is Kees. Kees sees a potato on the table [7th ranked item] and a piece of chocolate [1st ranked item]. Kees wants to eat the potato, because he likes potatoes much better than chocolate. What does Kees want? (1)
>
> [Picture 2] Father comes into the room and gives Kees a potato [a piece of chocolate]. What do you think? Is Kees happy or is Kees unhappy? (2) What did Kees want? (3)

All children answered the desire questions (1) correctly. The last question (3) was a check on the memory for the protagonist's desire. Children were not corrected if they failed a memory question.

Results

Children performed better on the memory questions as age increased (79%, 86% and 96% correct for 3-, 4- and 5-year-olds respectively). A 3 (age) × 4 (distance) × 2 (scale: snack or toy) analysis of variance with repeated measures on the last two factors confirmed a significant main effect for age ($F(2,133) = 16.81$, $p < .001$), but also that the performances on the memory tasks were independent of story type; there was no significant main effect for distance.

Table 1 shows the mean score on the emotion prediction question, per age group, for four story types (distance): stories in which the protagonist's desire was most saliently in conflict [1] or less saliently in conflict [2] with the preferences of participants, and stories in which the protagonist's desire was less saliently consistent [3] or most saliently consistent [4] with the preferences of participants. Every story type was represented by two stories (in one story the protagonist received the desired item, in another she or he did not). Thus, collapsed over the snack and the toy scales, participants obtained a maximum score of 2 when they accurately responded to both emotion prediction questions, and a minimum score of zero when they gave wrong answers to both questions. Note that our use of the terms *correct* and *incorrect* emotion predictions refer to emotion predictions respectively based or not based on the protagonist's desire. These terms are not meant to connote value judgments. Stories with a happy or an unhappy outcome were answered equally well and independently of story type.

It had been assumed that children's pattern of responses would not differ for the snack and the toy scale. However, the results in Table 1 show that although correct scores increased with age for both the snack and toy scale, the four story types evoked a different pattern across the two scales. All respondents were included in the analysis below, but note that these results did not differ when children were excluded who did not respond correctly to all eight memory questions. A 3 (age) × 4 (distance) × 2 (scale snack or toy) analysis of variance with repeated measures on the last two factors confirmed a main effect for age ($F(2,133) = 31.58$, $p < .001$) and an interaction of scale × distance ($F(3,399) = 4.13$, $p = .007$). The results for the two scales were, therefore, examined separately. Note that, due to ceiling effects for the 5-year-olds, the assumption of homogeneity of variance was not met in conditions where they scored $> .90$. After exclusion of the 5-year-olds, the test for homogeneity showed that the only condition in which the 4-year-olds scored $> .90$ was still significant (Cochran's $C(44,2) = .76$, $p < .001$).

For the snack scale, most saliently conflicting stories produced the lowest scores, the two less salient story types were in-between, and most saliently consistent stories produced the highest scores. However, it can be seen that this trend is due to the response pattern of the 4-year-olds. Three- year-olds responded quite randomly (between 50% and 66% correct), while 5-year-olds correctly predicted the protagonist's emotion for all story types (91% to 94% correct). A *post hoc* 3 (age) × 4 (distance) analysis of variance for the snack scores with repeated measures on distance confirmed a main effect for age ($F(2,133) = 34.57$, $p < .001$), distance ($F(3,399) = 3.37$, $p = .019$) and an

Table 1. Mean numbers of correct emotions for the snack and toy item pairs predicted by the three age groups of children in Expt 1

	Distance between protagonists' and participants' preferences			
	Most saliently conflicting	Less saliently conflicting	Less saliently consistent	Most saliently consistent
Snack scale				
3-year-olds	.55	.58	.50	.66
(N = 38)	(.50)	(.50)	(.51)	(.48)
4-year-olds	.56	.81	.81	.92
(N = 52)	(.50)	(.40)	(.40)	(.47)
5-year-olds	.93	.94	.94	.91
(N = 46)	(.25)	(.25)	(.25)	(.29)
Total	.68	.79	.77	.85
(N = 136)	(.47)	(.41)	(.43)	(.36)
Toy scale				
3-year-olds	.71	.55	.68	.69
(N = 38)	(.46)	(.50)	(.47)	(.47)
4-year-olds	.81	.64	.87	.75
(N = 52)	(.40)	(.49)	(.35)	(.44)
5-year-olds	.91	.85	.96	.87
(N = 46)	(.29)	(.36)	(.21)	(.34)
Total	.82	.68	.85	.77
(N = 136)	(.39)	(.47)	(.36)	(.42)

Note. Range of responses is 0 to 1; standard deviation in parenthesis.

interaction of age × distance ($F(6,399) = 2.60$, $p = .017$). *Post hoc* analyses for polynomial trends per age group confirmed the expected linear trend for the 4-year-olds ($p < .001$). This shows that the increase in correct responses goes up in a linear fashion for the four story types that contain distance: from most saliently in conflict to most saliently consistent.

The participants' mean scores on the toy scale revealed an unexpected pattern: most saliently conflicting and less saliently consistent story-types produced higher scores than the two remaining story types. A *post hoc* 3 (age) × 4 (distance) analysis of variance for the toy scores with repeated measures on distance confirmed a significant main effect for age ($F(2,133) = 14.13$, $p < .001$), and distance ($F(3,399 = 3.99$, $p = .008$), but without a linear trend. Clearly, the two scales call for a different kind of explanation.

A possible explanation for the unexpected results for the toy scale is that the toy items appealed to possible sex-stereotyped beliefs about desirability. Some participants may have held sex-stereotyped beliefs about toy preference, for example, boys would liked the airplane and car most and the doll and tea-set least. Thus, a female protagonist who preferred a car or an airplane would have violated their sex-stereotyped beliefs about desirability. Therefore, children may have based their emotion predictions on their

normative beliefs rather than on their own preferences or the protagonist's preferences, resulting in inaccurate emotion predictions. In other words, regardless of the distance between the protagonist's and the participant's preference, stories could have been conflicting or consistent, depending on whether the protagonist's preference violated sex-stereotyped beliefs about desirability: a conflicting story could thus become consistent and vice versa.

We re-coded the stories accordingly, and note that it was no longer possible to include distance, because we could only see if there was a conflicting or a consistent situation between the protagonist's preference and participants' sex-stereotyped beliefs. *Post hoc* analysis confirmed that children predicted the protagonist's emotion more accurately when the protagonist had a traditional desire, for example a female protagonist wanted a doll, than when the protagonist had a non-traditional desire. A 3 (age) × 2 (desire type: i.e. traditional vs. non-traditional) analysis of variance with repeated measures on the last factor confirmed a significant main effect only for desire type ($F(1,68) = 5.97$, $p = .017$). The interaction effect was not significant.

Discussion

The results show that if stories contained snack items, 4-year-olds' accuracy improved when the distance between their own preference and the protagonist's preference was reduced. However, whereas 5-year-olds predicted the protagonist's emotion based on the protagonist's preferences under all circumstances – as is consistent with Yuill *et al.* (1996) – 3-year-olds seemed to react quite randomly. The fact that the 3-year-olds performed poorly also in the easiest condition, the condition in which their preference and the preference of the protagonist was most saliently consistent, suggests that they did not understand the task clearly. Especially the phrasing 'Kees likes X much better than Y' might have been too difficult for these children. The relative statement 'Kees likes X much better than Y' is probably interpreted in absolute terms and translated as 'Kees likes X' and 'Kees likes Y', which could explain their random reactions.

Unexpectedly, however, if the stories contained toy items, children's responses showed a different pattern. Participants predicted the protagonist's emotion more accurately when the protagonist preferred a toy that participants had ranked as least liked, than when the protagonist desired a toy that was ranked by participants as most liked. One possible explanation for this difference in results is that the distance between the ranked items of the two scales is not compatible. Thus, participants may have strongly disliked certain snack items, but liked all seven toy items. Yet, if a smaller distance between the ranked toy items accounted for the difference between the two scales, we would expect smaller differences between the different story types for the toy scale, rather than a different pattern.

Another reasonable explanation for the data is that we had falsely assumed that children's perception of desirability would correspond to their own desires. And indeed, *post hoc* analyses provided credibility to the idea that children's judgments on desirability were, at least partly, guided by sex-stereotyped beliefs about the desires in question. In the next experiment we aim to elaborate further on this line of argument.

EXPERIMENT 2

(Sex-stereotyped desires)

Introduction

To provide independent support for the hypothesis that sex-stereotyped beliefs about desirability influence the prediction of the protagonist's emotion, a second experiment was carried out. Children were presented with stories in which the protagonists' desires were either consistent or in conflict with the assumed sex-stereotyped belief of the participants (the assumption that 4- and 5-year-olds hold sex-stereotyped beliefs about toy preferences was first tested in a preliminary study). In the main study, the protagonist's sex was counterbalanced across story types and the protagonist's sex was explicitly emphasized to the participant. Out of the seven toy items that were used in the first experiment, four toys were expected to create the most salient distance in this respect: two stereotypical male toys (car or airplane) and two stereotypical female toys (doll or tea-set). We did not vary distance in this study, because we were primarily interested in the possible bias in children's emotion predictions resulting from sex-stereotyped beliefs. Preschool children, aged 4 and 5, were tested and correct emotion predictions were expected to increase with age. Moreover, participants were expected to predict the protagonist's emotion more accurately when the protagonist had a traditional desire (when a male protagonist preferred a car to a doll to play with) than when the protagonist had a non-traditional desire (when a female protagonist preferred a car to a doll).

Method

Participants

Participating in this study were 29 4-year-olds (mean age 4;7 years; range 4;3 to 4;11 years) and 29 5-year-olds (mean age 5;5 years, range 5;1 to 5;11 years). An additional 16 4-year-olds (mean age 4;7 years, range 4;1 to 4;11 years) and 16 5-year-olds (mean age 5;8 years, range 5;1 to 6;1 years) participated in the preliminary study described below. Half of the participants both in this study and in the preliminary study, were male and half were female and they came from middle class families. Participants were drawn from primary schools in the suburbs around Amsterdam.

Preliminary study

In order to study 4- and 5-year-olds' explicit beliefs about sex-stereotyped preferences for toys, children were asked to judge how much boys or girls liked to play with four toys: a car, an airplane, a doll and a tea-set. Participants were taken individually from their classroom to a separate room and tested for approximately 5 minutes. Children were presented with a horizontal bar and a ring, which could be moved from the extreme left side of the bar to the extreme right side. The reverse side of the bar contained a scale, which was only visible to the experimenter. The experimenter said:

> [Name child], do you like to play outside? Well then, on this side of the bar you see a happy face and on the other side you see a sad face. And this ring always starts in the middle. So, if you like to play outside, which face should you move the ring to? That's right, to the happy face. Then we move the ring back to the middle of the bar. And do you like to fall off your bike and hurt yourself? So, which side should you move the ring to?

Children were also instructed that if they [dis]liked an item a little bit, they should move the ring *a little bit* in the appropriate direction. This procedure was practised a few times. Then a picture of a toy (doll, tea-set, car and airplane) was shown and half the children were asked: 'Do boys like to play with [item]?' until all four items were scored. The items were presented in random order. Participants were then asked: 'Do girls like to play with [item]?' for the same toys. This order was reversed for the other half of the participants. When participants responded with: 'I like to play with [item]' the experimenter asked: 'But other boys [girls], do they like to play with [item] or don't they like to play with [item]?' When participants responded with: 'Some boys [girls] like to play with [item]' the experimenter asked: 'But most boys [girls]? Do most boys [girls] like to play with [item] or don't they like to play with [item]?'

   Children obtained a maximum score of 10 when the ring was moved to the extreme right side (happy face) of the bar and a minimum score of − 10 was obtained when the child had moved the ring to the extreme left side (unhappy face). Table 2 gives the mean score for traditional and non-traditional toys for the 4- and 5-year-olds. It can be seen that both age groups judged children to be more happy with sex-stereotyped toys (boys like a car and an airplane; girls like a doll and a tea-set) than with non-traditional toys. This pattern was more strongly evident for 5-year-olds than for 4-year-olds. A 2 (age) × 2 (participants' sex) × 4 (toys) × 2 (traditional vs. non-traditional preference) analysis of variance with repeated measures on the last two factors confirmed a main effect for traditional preference ($F(1,28) = 155.74$, $p < .001$) and an interaction for age × traditional preference, $F(1,28) = 13.52$, $p < .001$).

Table 2. How much do boys/girls like to play with . . .?

	Non-traditional toys	Traditional toys
4-year-olds	− 1.78	7.18
(N = 16)	(5.57)	(3.99)
5-year-olds	− 7.34	8.91
(N = 16)	(5.08)	(1.84)

Note. Range of responses is − 10 to 10; standard deviation in parenthesis.

Procedure

Children were tested individually by one experimenter. The same introduction and warm-up task as in Expt 1 were used, so as to familiarize participants with the terms and pictures (see the Method section of Expt 1 for details). All participants responded accurately to this warm-up task and were presented with eight stories. The protagonist was male in four stories and female in the other four. Every story contained a combination of two items: a stereotypical male toy (car or airplane) and a stereotypical female toy (doll or tea-set). The protagonist would, for example, either receive the car or the doll. The protagonist's preference for one item over another was expressed. In four stories, this preference was traditional (the protagonist's sex matched his or her preference for a sex-stereotyped toy) and non-traditional (the protagonist's sex was opposite to his or her desire for a sex-stereotyped toy) in the other four stories. Children were asked to predict the protagonist's emotion upon receipt of the preferred item of a pair (happy stories) or the non-preferred item (unhappy stories). An example of one of the stories is:

> [Picture one, representing two toys] This is a story about a boy, Simon. Simon wants to borrow a toy from his friend. His friend has a tea-set and a car. Simon wants to play with the car, because Simon likes cars much better than tea-sets. What does Simon want? (1)

[Picture two, representing the remaining toy] Simon's friend chooses the car and Simon has to play with the tea-set. What do you think? Does Simon feel happy or unhappy? (2) What did Simon want? (3)

The last question (3) was a check on the memory for the protagonist's desire. Finally, participants were presented with pictures of the four items that had been used in the stories (car, airplane, doll and tea-set) and they were asked: 'If you could play with only one of these toys, which toy would you choose to play with?' The selected item received rank-number 1 and its picture was removed. Then participants choose from the three remaining items and so forth, until the four items were ranked from 1 to 4. Based on this rank-order, participants were ascribed to a traditional preference group (male participants: car and airplane ranked first and second, female participants: doll and tea-set ranked first and second; N = 37) and the other participants were ascribed to a non-traditional preference group (N = 21). Thus, the non-traditional preference group also includes six female participants with a traditional male preference. No male participants revealed a traditional female preference. Traditional and non-traditional preferences were equally distributed among 4- and 5-year-olds. This categorization was used to see if children who do not themselves subscribe to those sex-stereotyped norms (children with a non-traditional preference) would also be biased in their emotion predictions by sex-stereotyped beliefs about desirability.

Results

The memory questions concerning the protagonist's desire were answered well (86% correct) and irrespective of age and emotional outcome (happy or unhappy), but traditional desires of the protagonist (91% correct) were memorized better than non-traditional desires (82% correct). Also in the work of Carter and Levy (1988) it was shown that preschool children's memories were distorted by gender-atypical information. All respondents were included in the analysis below, but note that these results did not differ when children were excluded who did not respond correctly to all eight memory questions.

Table 3 demonstrates the percentage of correct emotion predictions for stories in which the protagonist had a traditional and a non-traditional desire; and for stories with a happy and an unhappy outcome. The scores are collapsed over age, the gender of the participants and over children with a traditional or a non-traditional preference for toys, because these factors were not influential on children's responses. Every story type is represented by two stories (a male and a female protagonist). It can be seen that children gave more accurate emotion predictions when the protagonist had a traditional desire than when the protagonist had a non-traditional desire. Concurrently, children were

Table 3. Percentage of correct emotions predicted in Expt 2

	Protagonist's desire	
	Non-traditional	Traditional
Happy	86%	90%
	(.59)	(.52)
Unhappy	66%[a]	80%
	(.78)	(.73)

[a]This percentage is greater than that expected by change ($\chi^2(2)$ = 19.52, $p < .001$) standard deviation in parenthesis
Note. N = 58.

more accurate when it concerned a happy than an unhappy emotion prediction. It can also be seen in Table 3 that the influence of traditional vs. non-traditional desires on children's predictions was somewhat stronger in the case of unhappy emotion predictions than in the case of happy predictions.

A 2 (age) × 2 (participant's preference: traditional vs. non-traditional) × 2 (protagonist's desire: traditional vs. non-traditional) × 2 (valence: happy vs. unhappy emotion predictions) × 2 (protagonist's sex: male vs. female) analysis of variance with repeated measures on the last three factors, confirmed a main effect for protagonist's desire ($F(1,54) = 5.41$, $p = .024$) and valence ($F (1,54) = 7.74$, $p = .007$), and an interaction of protagonist's desire × valence ($F (1,54) = 4.32, p = .042$). *Post hoc t* tests confirmed that more accurate responses were given for traditional than for non-traditional stories in the unhappy condition ($T = 3.24$, (57) $p = .002$), but not in the happy condition, although the direction of the difference in the happy condition also matches our expectations.

Based on participants' individual preferences for toys, stories were also classified as in conflict or consistent with the protagonist's desire. For example, a story in which the protagonist preferred a doll over an airplane was classified as in conflict with participants' preference when participants had ranked an airplane as more desirable than a doll. This story was classified as consistent in the case of participants who had ranked a doll as more desirable than an airplane. With this classification, four stories were always in conflict with participants' preference, whereas the other four stories were consistent. Participants performed equally well on stories in which the protagonist's desire and the participants' desire were consistent (82% correct) as compared to stories in which the two were in conflict (78% correct). A 2 (age) × 2 (participant's preference) × 2 (conflicting or consistent stories) analysis of variance with repeated measures on the last factor confirmed this pattern, because no main effects or interactions were shown.

Discussion

First, preschoolers were shown to have explicit beliefs about desirability for toys, based on sex-stereotyped ideas. Regardless of their own sex, they think that girls will be happy to play with traditional female toys and unhappy with traditional male toys, whereas they suppose an opposite pattern for boys. These results are in line with the work of Eisenberg, Murray, and Hite (1982), who showed that 3- to 5-year-olds frequently used sex-stereotyped arguments to justify a protagonist's choice for a toy, but explained their own preferences in terms of the toy's functionality. Also consistent with other work (Damon, 1977), these ideas increase with age: 5-year-olds in our research were more outspoken in their judgments than 4-year-olds.

Secondly, the findings of this experiment supported the hypothesis that sex-stereotyped beliefs about desirability biases children's predictions of others' emotions; children were more accurate in their emotion predictions when the protagonist had a traditional desire, than when the protagonist had a non-traditional desire, irrespective of children's own preferences for one toy over the other. Note, however, that this influence of traditional vs. non-traditional desires was more strongly evident in the case of unhappy outcomes than when happy outcomes were concerned. This difference between happy and unhappy conditions is due to two reasons. First, when a male protagonist

wanted a stereotypical female toy, but received a stereotypical male toy, quite a few children (inaccurately) predicted that the boy was happy. It seems reasonable that these kinds of story types are most difficult for children, because the protagonist receives an item that, consistent with the participant's sex-stereotyped beliefs about desirability, matches the protagonist's sex. Secondly, children were in general less accurate in the case of unhappy outcomes, than in the case of happy outcomes. In other words, children tend to predict happy emotions more frequently (Rieffe, Meerum Terwogt, Koops, & Hagenaar, 2000). This 'happy bias' is also evident in other experiments (e.g. Harris *et al.*, 1989) and in our research resulted in a ceiling effect for the happy condition. This ceiling effect prohibited an obvious difference between the traditional and non-traditional conditions, although the results show a trend in the expected direction.

GENERAL DISCUSSION

Although preschool children, overall, accurately predicted the protagonist's emotion based on the protagonist's desire in both experiments, inaccuracies emerged as well. When the desires concerned snack items, it became clear that 3-year-olds had severe problems with predicting others' emotions, even when the desire of the protagonist and the participant corresponded. Note that the construction in which the desirability of one item is relative to the other 'Kees likes X better than Y' might have been too demanding for these young children, because it requires them to work with relative rather than with absolute judgments. The responses of the 5-year-olds, on the other hand, showed a ceiling effect. They were not at all hampered by the syntax, nor biased by the task conditions: 5-year-olds correctly predicted the protagonists' emotions under all circumstances. The results of the 4-year-olds showed a pattern we had expected: when the distance between their own preference and the protagonist's desire increased, their predictions of the protagonists' emotions were less accurate.

Considering for a moment the possibility that children not only recognize their own desires, but that they have developed knowledge about what most people – or in this case, most children – find desirable, then it is plausible that children's own desires largely coincided with these so-called *beliefs about desirability* in our experiment, because their own desires for snacks and those of most children were likely to be equivalent. For example, 90% of our participants ranked chocolate at the most preferred end of the scale. In other words, there is no obvious reason why children's ideas about desirability should differ from their own desires.

Moreover, all other experiments in which there was a distance between the protagonist's and participant's desire (Moore *et al.*, 1995; Wellman, 1990; Yuill *et al.*, 1996) used situations in which it is most likely that participants' desires and the desires of most children corresponded. In the case of Moore *et al.*'s experiment with a less attractive sticker and a much more attractive sticker: most children will choose the more attractive sticker. In the case of Wellman's experiment: both alternatives are likely to be attractive to most children. And in case of Yuill *et al.* we expect that young children consider moral issues as fixed and objective (Piaget, 1932). In other words, when trying to explain the bias shown by our 4-year-olds in their emotion predictions, it is impossible to distinguish between participants' own desires and their knowledge about

children's preferences in general – their beliefs about desirability – because they were largely identical.

An answer to this issue could be given by those rare children who do not like sweets. If their reasoning about others' emotions is biased by their own desires, they would not be troubled by a protagonist who prefers a potato over a piece of chocolate, indeed, we would expect them to find this condition simple. Yet, if they developed knowledge about children's preferences in general, and this knowledge would bias their responses, their response pattern should be equivalent to that of the 4-year-olds in our snack experiment.

In the case of toys, however, our findings demonstrated that preschool children were not guided or biased by their own preferences when predicting how someone will feel about a toy given to him or her. Instead, they were guided by beliefs about desirability which did not necessarily correspond with their own desires: their sex-stereotyped expectations (note that this conclusion is only valid for stories with an unhappy outcome, because stories with a happy outcome showed the expected tendency, but also had a ceiling effect). Irrespective of their own preferences, children found it more difficult to predict that a girl would be unhappy to receive a doll than to predict that she would be unhappy to receive an airplane. Biased emotion predictions even emerged among girls with no clear-cut preferences for dolls over airplanes or boys with no preference for airplanes over dolls, clearly demonstrating that normative information can be assimilated and used by children who do not themselves subscribe to those norms. This is consistent with the work of Eisenberg *et al.* (1982) in which preschoolers used sex-stereotyped arguments to justify choices of others but not themselves.

In conclusion, it became clear that, in line with authors like Perner (1991) or Astington and Gopnik (1991), the subjective character of desires is not always recognized or accepted by preschool children. However, this perception of desirability as an objective property is not always due to the generalization of one's own desires. Our results from the toy experiment showed that preschool children's knowledge of what is desirable to most people (given their group membership) biased their emotion predictions, whereas their own preferences played a subordinate role in this process. In the snack experiment, children's own preferences will have largely coincided with their beliefs about the preferences of most children. Thus, we cannot be sure what actually caused the bias of our 4-year-olds here: their own desires, their ideas about what most children find desirable, or a combination of both. Nevertheless, we can conclude from these results that desires are sometimes perceived as objective features, but that the foundation of this perception does not necessarily stem from children's own preferences. Instead, these beliefs about desirability can be based on social, normative knowledge as well. The perception of desirability as an objective property is probably not restricted to children or to gender related issues. Adults can also attribute their own ideas about desirability to others under circumstances that are perceived as drastically deviating from their own cultural norms (Meerum Terwogt & Rieffe, submitted).

Note that these results are relevant to the debate about the two major theoretical accounts of the development of the child's theory of mind: 'simulation theory' (e.g. Gordon, 1992) and 'theory theory' (e.g. Stich & Nichols, 1992). Simulation theory – which assumes that children start by projecting their own mental states onto others and then make imaginative adjustments according to their knowledge about the other –

would indeed predict that young children might attribute their own desires, as was shown by the 4-year-olds in our snack experiment. Also the reported distance effect seems to fit in with simulation theory, because this approach would predict that the larger the distance one has to bridge by imaginative skills, the more likely it becomes that such an attempt will fail. Unexpectedly, however, it turned out that children also apply rules, they use generalizations, based on their knowledge about the world, which is not necessarily in line with their own desires. Here, it is obvious that children are not – at least not always – simulating. Instead, they seem to derive their knowledge from an everyday, *theory-like* framework of concepts, rules and generalizations. Of course, simulation theory never intended to exclude the influence of such domain specific theoretical notions. Stone and Davies (1996, p. 133) explicitly stated that 'inductive based generalisations play a role in real life use of mental simulation'. However, our findings clearly demonstrate that, at an early age, the effects of such rule-like generalizations are far from marginal. Under some conditions, they even outweigh the influence of children's own desires at an age as young as 4 years.

Acknowledgements

References

Astington, J. W., & Gopnik, A. (1991). Developing understanding of desire and intention. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading*, (pp. 39–50). Oxford: Blackwell.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford: Oxford University Press.

Carter, D. B., & Levy, G. D. (1988). Cognitive aspects of early sex-role development: The influence of gender schemas on preschoolers' memories and preferences for sex-typed toys and activities. *Child Development, 59*, 782–792.

Currie, G. (1998). Pretence, pretending and metarepresentation. *Mind and Language, 13*, 35–55.

Damon, W. (1977). Measurement and social development. *Counseling Psychologist, 6*, 13–15.

Eisenberg, N., Murray E., & Hite T. (1982). Children's reasoning regarding sex-typed toy choices. *Child Development, 53*, 81–86.

Flavell, J. H., Flavell, E. R., Green, F. L., & Moses, L. J. (1990). Young children's understanding of fact beliefs versus value beliefs. *Child Development, 61*, 915–928.

Frijda, N. (1995). *The irrationality of emotions*. Paper presented at the Fourth Annual Meeting of the European Society for Philosophy and Psychology, Oxford, 29 August – 1 September.

Gordon, R. M. (1992) The simulation theory: Objections and misconceptions. *Mind and Language, 7*, 11–34.

Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition and Emotion, 3*, 379–400.

Lillard, A. (1993). Young children's conceptualization of pretense: Action or mental representational state? *Child Development, 64*, 372–386.

Meerum Terwogt, M., & Rieffe, C. (submitted). Stereotyped beliefs about desirability: A potential problem in establishing the child's theory of mind. *Cognition and Emotion*.

Moore, C., Jarrold, C., Russell, J., & Lumb, A. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development, 10*(4), 467–482.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Piaget, J. (1932). *The moral judgement of the child*. New York: Harcourt & Brace.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*(1), 12–21.

Rieffe, C., Meerum Terwogt, M., Koops, W., & Hagenaar, J. (2000). The desirability of beliefs: Preschooler's appreciation of fact beliefs and subsequent emotions. *Infant and Child Development*, 9, 147–160.

Searle, J. R. (1983). *Intentionality*. Cambridge: Cambridge University Press.

Stich, S., & Nichols, S. (1992). Folk psychology: Simulation or tacit theory? *Mind and Language*, 7, 35–71.

Stone, T., & Davies, M. (1996). The mental simulation debate: A progress report. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 119–140). New York: Cambridge University Press,

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Yuill N., Perner J., Pearson A., Peerbhoy D., & Ende, J. van den (1996). Children's changing understanding of wicked desires: From objective to participative and moral. *British Journal of Developmental Psychology, 14*, 457–475.