Child Development, March/April 2006, Volume 77, Number 2, Pages 413–426

Children's Thinking About Counterfactuals and Future Hypotheticals as Possibilities

Sarah R. Beck
University of Birmingham

Elizabeth J. Robinson
University of Warwick

Daniel J. Carroll and Ian A. Apperly
University of Birmingham

Two experiments explored whether children's correct answers to counterfactual and future hypothetical questions were based on an understanding of possibilities. Children played a game in which a toy mouse could run down either 1 of 2 slides. Children found it difficult to mark physically both possible outcomes, compared to reporting a single hypothetical future event, "What if next time he goes the other way . . ." (Experiment 1: 3–4-year-olds and 4–5-year-olds), or a single counterfactual event, "What if he had gone the other way . . .?" (Experiment 2: 3–4-year-olds and 5–6-year-olds). An open counterfactual question, "Could he have gone anywhere else?," which required thinking about the counterfactual as an alternative possibility, was also relatively difficult.

A powerful feature of adult thinking is the readiness to reflect on possibilities, to consider what we know about reality within a broader context of what might be the case or what might have been the case. There are two distinct circumstances under which adults think about possibilities. In the first, the information currently available about past, present, or future reality is limited. We do not know the actual outcome, but we can speculate about what might be or might have been. For example, suppose I have agreed to collect my friend's child from school, but when I arrive there I find there are two gates from which the children can emerge and I could wait at either one of them. In the second circumstance we know for sure what is the case, but we speculate about alternatives that could realistically have happened instead, or could realistically happen in the future. For example, suppose the child comes out of the gate where I am not standing and sadly assumes she has been forgotten. When I eventually find her I might think "I should have waited at the other gate" or "Next time I'll wait at the other gate." These counterfactual (the former) and future hypothetical or prefactual (the latter) thoughts allow humans to escape from the here and now and entertain alternatives to reality. In the experiments reported below, we examine the relationship between children's thinking about possibilities in these two different circumstances.

There is reason to expect future hypotheticals and counterfactuals to be related in development. It is implicit in thinking about what could realistically have been the case (in the example above, "I should have waited at the other gate"), that there was a point in the past when either the counterfactual event or the actual event could have occurred (I could have waited at either of the gates). Similarly, it is implicit in thinking about how to avoid the same problem next time (where should I wait tomorrow?), that there was an earlier point in time when I could have chosen to behave differently. This is what is "special" about counterfactuals: Counterfactual events are not merely events that did not happen, they *could have* replaced the actual. Thus, the understanding that at any point in time multiple possible events could occur in the relative future underpins both future hypothetical and counterfactual thoughts.

Research on the cognitive processes that underpin counterfactual thought is relatively rare, but two authors, who have explored the cognitive bases for counterfactual thought, make this analysis explicit. First, Byrne (1997, p. 108) emphasized that counterfactual and actual possibilities are closely related and that counterfactuals are "grounded in the factual reality from which they depart." She describes counterfactuals as "dual possibilities":

This research was supported by a grant from the Economic and Social Research Council, UK (RES-000-22-0305).

Correspondence concerning this article should be addressed to Sarah R. Beck, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. Electronic mail may be sent to s.r.beck@bham.ac.uk.

 0009-3920/2006/7702-0012

> 'if only he were handsome, I would marry him' conjures up two possibilities, an imaginary possibility in which the man is handsome and the speaker marries him, but equally, a reality in which the man is not and she does not. (p. 25)

Byrne (1997; see also 2005) is clear that thinking about a counterfactual requires one to hold in mind two possibilities. A similar description is given by Perner (2000, p. 394), who describes a counterfactual "not just as a possibility but as a point for point alternative to real events." Critically, here Perner (like Byrne) contrasts thinking of a single possible event and holding two possibilities in mind. It is the latter that both see as necessary for counterfactual thought. What of thinking about the future? By its very nature the future has yet to be determined. Thus, thinking about a possible event in the future necessarily involves its inherent uncertainty (see, e.g., Atance & O'Neill, 2005).

On the basis of this analysis, we might expect that children engage in counterfactual or future hypothetical thinking only if they can identify the set of possibilities afforded by limited information. The published data strongly suggest otherwise however. We review these findings in the following sections and offer reasons for the contradiction between our expectation and what seems to be the case.

Research on children's acceptance of undecidability in logical reasoning has used tasks similar to the example above of the child coming out of school. In these circumstances, the protagonist has insufficient information to decide between possibilities, only one of which will turn out to be true. Child participants are provided with limited visual or verbal information about a state of affairs, such as which dolls are in a house, or which set of materials a model was made out of, and judge whether or not they can tell what is the true state of affairs (e.g., Fay & Klahr, 1996; Piéraut-Le Bonniec, 1980). Children seem to accept the possibility of undecidability at around the age of 6–8 years, whereas younger children are likely to treat the information as affording more certainty than it actually does, by incorrectly reaching a firm decision about what is the case. Very similar results appear in the large literature on children's understanding of ambiguity in utterances. In a typical ambiguity procedure, the child hears a message which is intended to refer to one item in an array, but which in fact refers equally well to more than one. In other studies, children see a partially informative view of an object, insufficient to be sure of its identity. In these tasks, there are multiple possible interpretations of the input. Children's understanding of these possibilities and the strategies they can employ to deal with them have been thoroughly investigated. There is a substantial body of research on children's evaluations of the informativeness of utterances (e.g., Beal & Flavell, 1982; Flavell, Speer, Green, & August, 1981; Robinson & Apperly, 2001; Robinson & Whittaker, 1987), and smaller literatures on children's explanations for different interpretations of ambiguous input (Carpendale & Chandler, 1996; Chandler, Hallet, & Sokol, 2002), and on their nonverbal responses to ambiguous input (e.g., Beck & Robinson, 2001; Patterson, Cosgrove, & O'Brien, 1980; Plumert, 1996).

Taken together, this research suggests that it is not until the age of around 6–8 years that children achieve reflective understanding that utterances (or visual input) can be ambiguous. Before this age children also have difficulties with multiple possibilities in a related area of research when they are confronted by logical inconsistency (Ruffman, 1999). In this task, children were shown a picture of a boy, with the head hidden. They were told "Peter's wearing a hat" and also "Peter's not wearing a hat." Young children failed to recognize this conflict between two possibilities as problematic. We are interested in children's ability to reflect on multiple possibilities as this should underpin children's counterfactual and future hypothetical thinking. Younger children may hesitate when trying to interpret ambiguous input (Plumert, 1996), can correctly revise a wrong interpretation in the light of subsequent input (Beck & Robinson, 2001), and even 12-month-old infants discriminate between ambiguous and unambiguous situations in their nonverbal reactions. Baldwin (2000) reported that on hearing a novel word, infants were more likely to look toward their mothers when there were two novel objects in view than when there was only one. Success on these measures does not necessarily mean that one understands and can think about multiple possibilities.

Children seem to find it much easier, however, to reason about counterfactual or future possibilities (e.g., Harris, German, & Mills, 1996; Riggs, Peterson, Robinson, & Mitchell, 1998), even if we exclude pretense, which has been seen as revealing precocious counterfactual thinking (e.g., Custer, 1996; Harris & Kavanaugh, 1993; Leslie, 1987; Lillard, 1993) and reasoning about counterfactuals within a pretend context (Richards & Sanderson, 1999; see also Dias & Harris, 1988; Leevers & Harris, 2000). At around the age of 4 years, realistic counterfactuals appear in children's everyday speech (Kuczaj & Daly, 1979), and at around this age (e.g., Riggs et al., 1998) or even earlier (e.g., Harris et al., 1996) children can entertain realistic counterfactuals in experimental settings. In these tasks children hear a story, typically acted out with puppets or illustrated with pictures. They are then asked a question about a counterfactual antecedent and they must work out the consequences of this change. For example, in a story used by Riggs et al. (1998), a character Jenny makes a painting, which she leaves on the table in the garden while she goes into the house. While she is away the wind blows the painting up in to a tree. The counterfactual test question is ''What if the wind hadn't blown, where would the picture be?'' Riggs et al. found that 3-year-olds tended to give answers about the real world ''In the tree'' whereas by 4 years children were able to speculate about the counterfactual alternative ''On the table.'' Other authors have used similar tasks (e.g., German, 1999; Guajardo & Turley-Ames, 2004; Perner, Sprung, & Steinkogler, 2004) and none of these have questioned that typically developing 4-year-olds can entertain counterfactuals with ease.

There have been fewer studies of children's abilities to reason hypothetically about the future, using questions with the format ''What if X happens, how will the world be?'' Like counterfactual questions children need to put aside what they know to be true about the world here and now, and consider an alternative to it. Both future hypothetical and counterfactual tasks require the child to resist describing the world as it is. But where the counterfactual event is supposed to stand in for what is known to have happened in the past, the future hypothetical is not, and possibly because of this, future hypotheticals have consistently been shown to be easier for young children than matched counterfactuals. For example, Robinson and Beck (2000) showed 3–4-year-olds a toy road with a garage at each end. A toy car began in the middle of the road and drove along it to one of the garages. Most 3–4-year-olds answered correctly a question about a possible event in the future ''What if next time he drives the other way, where will he be?'' A matched counterfactual question ''What if he had driven the other way, where would he be?'' was significantly more difficult for these children, indicating that children did not just respond to the test question by pointing to ''the other'' garage. Other studies confirm that many 3-year-olds find future hypothetical questions consistently easy (Perner et al., 2004; Riggs et al., 1998).

Explanations of why children pass future hypotheticals before they can answer counterfactuals have tended to focus on the fact that the imagined alternative is not in conflict with reality. Authors have appealed to different theoretical accounts of cognition; for example, Robinson and Beck (2000) draw on the mental models literature (Johnson-Laird, 1983), and Frye (2000) makes use of cognitive complexity and control theory (e.g., Zelazo, Frye, & Rapus, 1996). Putting this debate aside, we emphasize that the consensus of researchers based on the literature is that children's thinking is adult-like by 3 years of age for future hypotheticals and 4 years of age for counterfactuals. Byrne's (1997, 2005) analysis that counterfactual thinking requires thinking about multiple possibilities seems to be assumed implicitly by developmental psychologists. Similarly, those authors who have used future hypothetical tasks talk about them as asking the child to consider an uncertain future, one that is possible or might happen (Perner, 2000; Riggs et al., 1998; Robinson & Beck, 2000).

It seems clear then that children much younger than 6–8 years of age can answer questions about realistic alternatives to reality, even though the literatures on understanding of undecidability and ambiguity suggest that these young children would be unable at an earlier time point to identify two possibilities, one of which eventually happens and the other of which could have happened or could happen in the future. How might we reconcile this empirical evidence with our analysis of the relationship between the two kinds of thinking about possibilities? First we consider whether methodological differences could explain this discrepancy.

One possibility is that as very different tasks have been used there could be uninteresting methodological explanations for differences in difficulty. In our experiments reported below, we created matched tasks to avoid such problems of comparison. Furthermore, the tasks used to assess children's handling of alternative possibilities might be more demanding than necessary because of the researchers's focus on explicit evaluation of the informativeness of the input. For example, in the research on understanding of ambiguity, children are often expected to report *why* alternative interpretations are possible, and not just identify what the alternatives are.

One much simpler procedure that indirectly assessed children's ability only to specify alternative possibilities was used by Ruffman, Garnham, Import, and Connolly (2001). These authors' focus of attention was on implicit understanding of false belief, but they used an explicit control task to assess children's certainty, the results of which suggest that even 3-year-olds can identify possibilities when currently only limited information is available. In this control task, children were shown a container which might hold either a red square or a green ball, and two slides, one of which could accommodate only red squares and the other of which could accommodate only green balls. Children were asked to distribute 10 counters beside the slides to indicate where the shape would come. The correct response, indicating awareness of uncertainty about the outcome, was to place counters beside both slides. On other trials when children knew for sure that the container held only red squares (or green balls), the correct response was to place all the counters beside one slide. In the first experiment reported, 21 out of 28 (75%) children aged 3 and 4 years distributed their counters between the two slides when the outcome was uncertain in that they did not know which shape was in the container, whereas the great majority (90% on one trial and 93% on another trial) placed all their counters beside the correct slide when they knew what the shape was. In the second experiment, when the container held five red squares and five green balls, 55% of children distributed counters between the two possible slides, compared with 90% who placed all their counters by the correct slide when the container held 10 shapes of a single color. Although there is room for improvement, children's performance here is vastly superior to that reported on typical (and more complex) ambiguity tasks. These results suggest that children aged 3 and 4 years can identify two possible outcomes when they have only limited information about which one is going to happen. Following Ruffman, we used a very simple task in our studies, in which children had only to identify the two possibilities (not explain their occurrence). Thus, any difference between our new task and standard future hypothetical or counterfactual questions could not be due to these procedural differences.

In the experiments reported here, children were asked to acknowledge both possible outcomes of an event as yet unknown, and also to acknowledge a future hypothetical (Experiment 1) or a counterfactual (Experiment 2) alternative to the one eventual outcome. There are two likely patterns of results. In these well-matched tasks, we may find that children who consider future hypotheticals and counterfactuals can also identify the possibilities that existed at an earlier point in time. This would show that very young children can handle multiple possibilities under some circumstances. Alternatively, children who can answer the future hypothetical and counterfactual questions might find even our improved undetermined trials difficult. This would raise questions regarding children's thinking when they answer the standard questions, and would have implications for other findings in the counterfactual literature.

Experiment 1

Our prediction is that children who can answer questions about alternative possibilities (future hypotheticals or counterfactuals) will acknowledge that at a given point in time multiple possibilities may occur. As mentioned above, children aged 3–4 years find it easier to acknowledge a future hypothetical of the form "What if next time such and such happens . . .?" than a matched counterfactual of the form "What if such and such had happened . . .?" (Perner et al., 2004; Riggs et al., 1998; Robinson & Beck, 2000). It therefore makes sense to test our prediction first for future hypotheticals, because if it is supported for future hypotheticals, there is no need to test it again for counterfactuals.

Method

Participants. The final sample consisted of 69 younger children (28 boys), mean age 3 years 7 months (range 3 years 2 months to 4 years 2 months) from nursery classes, and 65 older children (38 boys), mean age 5 years 1 month (range 4 years 7 months to 5 years 7 months) from the first year of formal schooling. The children attended schools serving working and middle-class populations in Birmingham, UK. These schools serve diverse ethnic populations, with children from Caucasian, Asian, African, and Caribbean backgrounds. All participants' English was deemed competent by their teacher for comprehension of the procedure. We confirmed this with check questions during the procedure. All children in the classroom were included in our sample. Three additional children from the 3- and 4-year-old sample and 1 from the 4- and 5-year-old sample were excluded because of failure to complete the tasks.

Materials. We used the apparatus shown in Figure 1. This comprised a board approximately 1.5 m^2 with two slides made from semicircular tubes mounted on it. The left-hand slide was painted blue and ran vertically down the board. The right-hand slide was painted red and split halfway down the board to give two possible exits. At the junction were two gates that could be removed to open their respective sides. Beyond the gates, one end was painted with spots and the other with stripes. We used two sets of cards (approximately 10 cm × 3 cm). Cards in one set were red or blue on one side and plain white on the other. Those in the second set were patterned spotted or striped on one side and plain white on the other. We also used a small toy mouse and three cotton wool mats.

An important feature of the game was that the path of the mouse on each trial was determined by the

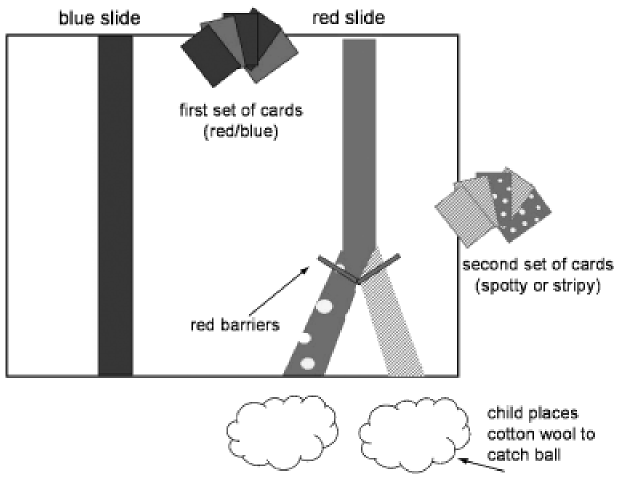

Figure 1. Diagram of apparatus.

child's choice of card which he or she selected from a face-down set. This was intended to make clear that neither experimenter nor child controlled the mouse's exit point, and to make it as easy as possible for children to realize when there was more than one possible exit on a particular trial.

Procedure. Children were tested in a quiet area away from the other children in the class. The game was explained to the child: The mouse would come down one of the slides and the child's job was to put out cotton wool mats to ensure the mouse did not hurt himself on the hard floor. This was a crucial feature of the game, designed to minimize any inclination the children had to guess, as making a guess risked the mouse hurting itself on the floor. Some authors have suggested that young children fail undecidability tasks because they prefer to guess an interpretation (e.g., Braine & Rumain, 1983; Speer, 1984). Others show that children's difficulties go much deeper than mistakenly thinking it is appropriate to guess (e.g., they have difficulty identifying the message as inadequate; see, e.g., Robinson & Apperly, 2001; Singer & Flavell, 1981). Although a preference to guess does not explain children's reported difficulties with multiple possibilities in the undecidability literature, we sought to minimize any chance children might deem guessing appropriate in our new task.

The game began with four *warm-up trials*, during which the child became familiar with the possible paths the mouse could take, the role of the gates at the junction of the red slide, the role of the two sets of cards, and her role in placing cotton wool. On the first warm-up trial, the experimenter let the mouse come down the blue slide after showing the child how to put out a mat in the right place to ensure that the mouse landed safely. On the second warm-up trial, the experimenter let the mouse run down the red slide as far as the gates. The experimenter made explicit that the child did not know whether the mouse would come down the stripy or the spotty side, and directed him or her to put out two mats, at the end of the stripy and the spotty slides. The experimenter then opened one of the gates and the child watched the mouse come down and land safely.

The third and fourth warm-up trials introduced the child to the cards that on some experimental trials would determine which way the mouse would go. The child was shown the red and blue cards, and the experimenter explained that if a red card was picked the mouse would go down the red slide, and if a blue card was picked the mouse would go down the blue slide. The experimenter then invited the child to pick a card from a choice of two, presented face down. Whichever color the child chose on the third trial, on the fourth trial the experimenter used sleight of hand to ensure that the child picked the other color. When a blue card was chosen the child put out one mat under the blue slide and then the mouse was released. When a red card was chosen the mouse was allowed to go down the red slide to the gates. The experimenter then showed the child the set of spotty and stripy cards and explained that they determined whether the spotty or stripy gate was opened. The child chose one of two face-down cards, the experimenter explained that "you've picked the spotty [or stripy] card, so he'll come down the spotty [stripy] side" and the experimenter ensured that the child placed one mat under the appropriate end of the red slide. The gate was removed and the mouse landed safely on the mat.

Hence by the end of the warm-up, the child had seen the mouse come down both the red and blue slides, had chosen from both sets of cards, and had put out one and two cotton wool mats when appropriate. The warm-up finished with a reminder of the importance that cotton wool was in place to catch the mouse, and that sometimes more than one mat would be needed.

It could have been that children routinely placed two or even three mats on trials when the mouse's exit was completely determined (e.g., when a red and then a stripy card had been chosen), but pilot work showed that children did not do this, and so there was no need to instruct them to place only the minimum number of mats. Pilot work also showed that when two mice were released at the same time, one down the blue slide and the other down the red slide with a single gate open, children performed near ceiling: 18 out of twenty-one 3–4-year-old children placed mats correctly beneath both exits. Given this good performance, we did not include two mouse trials in the main experiment.

After the warm-up trials there followed two future hypothetical trials, two undetermined trials, and two filler trials. On *future hypothetical trials* the experimenter offered the child a choice of two face-down cards, both of which were red although the child believed one was red and one blue. After child and experimenter had seen that the chosen card was red, the experimenter let the mouse go down the red slide to the gates. The child then had a genuine choice from a face-down pair of a spotty and a stripy card. After looking at the chosen card, the child was asked which gate would be opened and was asked "Can you put out cotton wool to make sure he lands safely?" The experimenter removed the gate indicated by the spotty or stripy card, and the mouse landed on the cotton wool. The experimenter then asked the future hypothetical test question "What if next time he goes the other way, where will he be?" This question was similar to future hypothetical questions used in previous research (Robinson & Beck, 2000).

Undetermined trials began in the same way as future hypothetical trials in that the mouse went down the red slide as far as the gates, but this time the child was asked to place cotton wool *before* she chose a spotty or stripy card. Hence at the time the cotton wool was placed it was undetermined which side the mouse would go down, and to ensure the mouse did not hurt himself it was necessary to cover both possibilities. In order to give the children the best chance of showing that they knew that the mouse could go either way, those who put out only one mat were prompted "Could it go anywhere else?" and thereby given the opportunity to add a second mat. This also indicated to children that they should not make a guess about the outcome. Children then chose a spotty or stripy card from a face-down pair, were given the opportunity to adjust placement of the cotton wool if they had put out only one mat in the wrong place, the gate was removed, and the mouse went down the slide to land safely on the cotton wool.

On *filler trials* the mouse came down the blue slide. The child chose from a set of face-down cards which were in fact both blue, but she believed one was red and one blue. The child put out cotton wool, and the mouse came down the blue slide.

The order of trials was counterbalanced: Children had either two future hypothetical, then one filler, then two undetermined trials, or they had two undetermined, then one filler, then two future hypothetical. The younger (3- and 4-year-old) children had an additional filler trial at the end of the warm-up to check that they understood the use of the red and blue cards and the cotton wool placement. This check was unnecessary for the older (4- and 5-year-old) children.

Throughout the game, further checks were made on children's understanding of the connection between choice of card and the slide the mouse came down. At the start of every experimental or filler trial, once the child had picked the card, the experimenter asked where the mouse would start. He also asked the younger (3- and 4-year-old) children which gate would be removed on the future hypothetical trials, once the card had been picked.

Results and Discussion

Performance was very good on all the check questions: 8 children (of 134) made one mistake and 6 made two mistakes. Children appeared to have a good understanding of how the game worked.

First, we examined whether children would be more likely to mark two potential future locations when the outcome was uncertain (undetermined trials), than when it was known (blue filler trials). We had included a prompt on the undetermined trials to give children the best possible chance of revealing competence, but as this was not included on the filler trials, for this comparison we looked at children's spontaneous performance on the undetermined trials, that is, before the prompt. Children rarely put out multiple mats when the mouse was coming down the blue slide (the younger children did this on 2 [of 138] trials. The older children who had only 1 trial of this type never put out multiple mats). In contrast, when the outcome was undetermined, 3- and 4-year-old children spontaneously put out two mats on 21% of trials and 4- and 5-year-old children did this on 37% of trials (29 out of 138 trials and 48 out of 130 trials respectively). Although children were more likely to put out one mat than two on the undetermined trials, we replicated Ruffman et al.'s (2001) finding that children discriminated between undetermined and blue trials, using our variation on their task.

Our analysis predicted that children should succeed on the future hypothetical trials only if they could acknowledge multiple possible outcomes (on the undetermined trials). However, in line with previous studies, performance on future hypothetical trials was excellent with children making mistakes on only 6% of trials. Given the relatively poor performance on undetermined trials described above it is clear that some children who were unable to pass the undetermined trials were passing the future hypothetical questions. We examined our data further using ANOVA.

Data Coding. On *future hypothetical* trials, after the mouse had come out at either the spotty or the stripy side of the red slide, children were asked ''What if next time he goes the other way, where will he be?'' On each of the 2 trials children scored one point if they pointed to the other exit of the red slide, giving a maximum score of 2. Some children pointed to the exit of the blue slide. This occurred on 32 of 270 trials, and we also counted these responses as correct. Incorrect responses were pointing to the exit where the mouse currently was (9 trials), pointing to the shelf where the mouse started the game (4 trials) and incomprehensible replies (e.g., ''flying,'' 6 trials).

On *undetermined trials*, children put out cotton wool while the mouse waited at the barrier. Children scored 1 if they put a mat at both exits of the red slide, either spontaneously or after the prompt ''Could he go anywhere else?'' As mentioned above, this generous coding gave children the best chance possible to show their understanding. On undetermined trials, adding a second mat after the prompt occurred on 11 of 138 trials with younger (3- and 4-year-old) children, and on 31 of 130 trials with older (4- and 5-year-old) children. The prompt increased the number of children passing undetermined trials, but we noted that hearing the prompt on the first undetermined trial did not inflate spontaneous performance on the second trial. No more children ($N = 7$) put out two mats spontaneously on the second trial having put out only one of the first (and thus heard the prompt), than passed the first trial and failed the second ($N = 8$). On 4 trials, children placed a mat correctly under each exit of the red slide, and a third mat was placed in between them. On 1 trial, a child placed one mat under one exit of the red slide, and one under the junction of the red slide. Both these responses were scored as correct. Incorrect responses were putting a single mat even after the prompt (younger children 98 trials and older children 39 trials), putting multiple mats in a pile beneath one red exit (8 trials) or putting a mat beneath the blue slide (4 trials). We coded placing a mat beneath the blue slide as wrong here because it indicated that the child had not understood the physical possibilities: that when a red card was chosen, the mouse had to go down the red slide. In contrast, blue was coded as correct for future hypothetical trials, as the question could be ambiguous, in that ''other way'' could mean other slide. Furthermore, pointing to the blue slide is not the typical realist error made on future hypothetical trials. Despite this we ran our analysis coding the blue responses on future hypotheticals as incorrect. The difference between trials remained for the younger group, supporting our finding that children can answer future hypothetical trials without being able to acknowledge multiple possibilities.

Comparison of trial types. We ran a repeated measures ANOVA with trial type (future hypothetical, undetermined) as a within-subject variable, and age (younger, older) and order (future hypothetical or undetermined trials first) as between-subject variables. The dependent variable was the number of trials correct for each trial type. There was a significant main effect of trial type, $F(1, 130) = 160.28$, $p < .001$, $\eta^2 = .55$. Performance on future hypothetical trials (mean = 1.85, maximum 2) was better than undetermined trials ($M = 0.90$). There was also a relatively small main effect of age, $F(1, 130) = 6.71$, $p = .011$, $\eta^2 = .05$, with older children ($M = 1.49$) outperforming younger children ($M = 1.26$). There was a significant interaction between age and trial type, $F(1, 130) = 29.13$, $p < .001$, $\eta^2 = .18$. All other effects were nonsignificant (highest $F = 0.83$, lowest $p = .36$). The mean scores for each trial type by age group are shown in Table 1.

We ran post hoc t tests to explore the interaction, making a Bonferroni correction for four tests so that the criterion for significance was $p = .0125$. The difference between trial types was significant for both age groups: younger $t(68) = -13.205$, $p < .001$; older, $t(64) = -5.05$, $p < .001$. Performance on the undetermined trials improved with age, $t(132) = -4.27$, $p < .001$. Performance on the future hypothetical trials was very good for both age groups, although performance by the younger group was slightly better. This difference approached but failed to meet significance having made the conservative Bonferroni correction, $t(87.71) = 2.49$,

Table 1

Mean Scores on Undetermined and Future Hypothetical Trials by Age Group in Experiment 1

Age group	Trial	Mean scores maximum = 2 (95% confidence intervals)
3- and 4-year-olds	Undetermined	0.6 (0.4 – 0.8)
	Future hypothetical	1.9 (1.9 – 2.0)
4- and 5-year-olds	Undetermined	1.2 (1.0 – 1.4)
	Future hypothetical	1.8 (1.7 – 1.9)

$p = .014$. A decline in performance is clearly surprising, but we note that only a small minority of children in each age group made any mistakes on future hypothetical questions (4/69 younger and 11/65 older children).

The important result, which occurs despite the generous scoring on the undetermined trials, is that both age groups found it easier to answer future hypothetical questions than they did to acknowledge multiple possibilities on the undetermined trials. Children answered future hypothetical questions correctly, apparently without thinking about the specified event as one of a set of possibilities.

Experiment 2

The results of Experiment 1 were clearly inconsistent with our prediction that children would answer the future hypothetical questions correctly only if they could specify both possible outcomes before one of them had happened. We already know that counterfactuals are more difficult than future hypotheticals. The accounts previously offered focus on difficulty imagining replacing current reality with, or relating it to, a counterfactual alternative, which is unnecessary for a future hypothetical (e.g., Perner et al., 2004; Riggs et al., 1998; Robinson & Beck, 2000). An alternative or additional factor might be that counterfactuals, but not future hypotheticals, require children to think of the replacement as a possibility that could have happened at an earlier point in time. If this is true, then success on a counterfactual question should imply success identifying multiple possibilities and should be of similar difficulty to the undetermined trials.

However, there may be an alternative explanation of the mismatch in performance on counterfactual and undecidability tasks. The particular tasks that have typically been used to assess young children's thinking about counterfactuals and future hypotheticals may allow them to answer correctly without necessarily thinking in terms of alternative possibilities. In the example from Riggs et al. (1998) given above, children were told what counterfactual situation to imagine (''What if the wind hadn't blown . . .?'') and asked about a feature of that situation (''Where would the picture be?''). Perhaps children could imagine how things would be and answer the question without thinking back to an earlier time when the wind might or might not have blown.

One notable departure from this type of question is the antecedent counterfactual task devised by Guajardo and Turley-Ames (2004). In this task, children heard a short story and were asked to generate a counterfactual antecedent rather than a counterfactual consequent. For example, the child was told a story about walking across the floor in muddy boots. The standard consequent tasks ask ''What if you had taken off your boots, would the floor be clean or dirty?'' The new antecedent question was ''What could you have done so that the kitchen floor would not have gotten dirty?'' This question allows children to generate multiple possible antecedents: You could have taken off your boots, you could have wiped them on the mat, you could have avoided the mud outside. Guajardo and Turley-Ames found that some 3-year-olds could generate appropriate counterfactual antecedents, with substantial improvement between 4 and 5 years. It is difficult to make direct comparison between ''success'' on consequent and antecedent tasks, because for the former there is only one correct answer, whereas for the latter children were scored for each alternative given. However, the overall conclusion is that children who could answer consequent questions could also generate appropriate counterfactual antecedents.

Does Guajardo and Turley-Ames' antecedent question require the child to think about multiple possibilities? One can think of the antecedents generated as different events that could have led to the given consequent. We are interested in whether children understand that at any one point in the past multiple events could have occurred. Thus, although Guajardo and Turley-Ames' task cleverly asks whether children understand that there were multiple points in the causal chain that could have been altered (e.g., removing boots, wiping them on mat), we are asking whether children understand the tension between multiple possibilities at the same point (one possibility that becomes the actual, one the counterfactual), that is, removing one's boots at the door, or not removing them. An open counterfactual question such as ''What else could have happened?'' might require the thinking about possibilities which seems to underlie adult-like counterfactual thinking. Like the antecedent counterfactual, the question uses the same speculative ''could'' phrasing, but asks the child to think about the point when multiple possibilities could have happened. Note also that many alternatives can be generated to answer an antecedent question, whereas the correct answer to an open counterfactual is to indicate the alternative possibility at the key point in time. This permits direct comparison with standard counterfactual and undetermined trials. In Experiment 2, we included an open question of this kind.

In Experiment 1, even the older children, aged 4–5 years, performed quite poorly on the undetermined trials. In Experiment 2, therefore, our older age group was a year older, 5–6 years, to allow us to check whether performance approached ceiling levels at this age. As in Experiment 1, our younger age group was 3–4 years.

Method

Participants. We tested 64 (36 males) younger children, mean age 3 years 11 months (range 3 years 6 months to 4 years 5 months) from two nursery classes, and 60 (28 males) older children, mean age 6 years 1 month (range 5 years 7 months to 6 years 6 months) from two classes of children in their second year of formal schooling. All attended schools serving working and middle-class populations in Birmingham, UK. As in Experiment 1, children were from Caucasian, Asian, African, and Caribbean backgrounds, and all participants' English was deemed competent by their teacher for comprehension of the procedure. All children in the classroom were included in our sample.

Materials. We used the same apparatus and materials as in Experiment 1. A cotton wool ball replaced the mouse for standard counterfactual and open counterfactual trials, because on these trials there was no mat placement and we did not want children to witness a mouse landing unprotected on the hard floor.

Procedure. The warm-up trials were the same as in Experiment 1. There followed *filler* trials and *undetermined* trials the same as those in Experiment 1. In addition, there were two new trial types: standard counterfactual and open counterfactual. On these trials, the experimenter removed the gates so the ball ran freely down the red slide and went down the spotty or stripy side at chance. The experimenter explained to the child that there was no need to put out mats to catch the ball. On *standard counterfactual trials*, the child chose from a face-down pair of cards to determine whether the ball would come down the red or the blue slide (unbeknown to the child both cards were red) and then the cotton wool ball was allowed to run freely down the red slide. Once it had landed on the floor beneath the stripy or the spotty exit, the experimenter asked the standard counterfactual test question, similar to that used in previous research (e.g., Riggs et al., 1998), "What if it had gone the other way, where would it be?" On open counterfactual trials, the procedure was exactly the same as on standard counterfactual trials, except that the test question was "Could it have gone anywhere else?"

Undetermined, standard counterfactual, and open counterfactual trials were presented in pairs, with a blue filler trial between each pair. Pairs were presented in six fully counterbalanced orders (standard [S], undetermined [U], open [O]; SOU; OUS, OSU; USO; UOS).

As in Experiment 1, children's understanding of the game was checked at various points: Before the mouse or ball entered the red or blue slide, the experimenter checked understanding of the role of the red and blue cards, and on filler trials the child's correct placement of cotton wool beneath the blue exit was checked.

Results and Discussion

Performance on the check questions was good: 1 child made two errors and 13 children made one error over the eight check questions. As in Experiment 1, children seemed to understand the workings of the game.

As in Experiment 1, children rarely put out two mats when the outcome was known (blue trials). In each age group, this happened on only 2 trials (of 128 trials for the younger group and 120 trials for the older group). When it was appropriate to put out two mats (undetermined trials), 3- and 4-year-old children did so spontaneously on 31% of trials (40 of 128) and 5- and 6-year-old children on 68% of trials (81 of 120). In line with Ruffman et al. (2001), children could discriminate between trials where it was appropriate to put out one and two mats.

Despite this discrimination between undetermined and blue trials, it is clear that, contrary to our prediction, children's performance on the undetermined trials was much poorer than on the standard counterfactual trials: 3- and 4–year-old children gave correct answers to 87% of these questions and all but one 5- and 6-year-old child answered both standard counterfactuals correctly (98% of trials correct). However, the performance on open counterfactual trials was not so good: 3- and 4-year-old children answered correctly on 66% of trials, and 5- and 6-year-old children answered correctly on 83% of trials. In what follows, we investigated the relative difficulty of the three types of trials more closely using ANOVA.

Data coding: On the *undetermined trials*, scoring was the same as in Experiment 1. Children who placed a second mat appropriately on prompting were coded as correct. As in Experiment 1, we checked whether hearing the prompt on the first trial influenced performance on the second. Twelve children failed the first trial but passed the second, whereas 16 passed the first and failed the second. On

13 trials children put out a third mat under the junction between the two red exits, in addition to a mat beneath each red exit, and we coded these responses as correct. We also scored as correct a child who put one mat under the stripy slide and one under the junction. On 10 trials children included a mat under the blue slide. This response was coded as incorrect, as it indicated confusion over the rules of the game.

On *standard counterfactual trials* and on *open counterfactual trials*, scoring was similar to that used for the future hypotheticals in Experiment 1: Children gained a score of 1 if they indicated the other red or the blue exit.

Table 2
Mean Scores on Undetermined, Standard Counterfactual, and Open Counterfactual Trials by Age Group in Experiment 2

Age group	Trial type	Mean scores maximum = 2 (95% confidence intervals)
3- and 4-year-olds	Undetermined	1.0 (0.9 – 1.2)
	Standard counterfactual	1.7 (1.6 – 1.8)
	Open counterfactual	1.3 (1.1 – 1.5)
5- and 6-year-olds	Undetermined	1.7 (1.5 – 1.9)
	Standard counterfactual	2.00 (1.9 – 2.0)
	Open counterfactual	1.7 (1.5 – 1.9)

Comparison Between Trial Types

We ran a repeated measures ANOVA with trial type as a within-subject factor (undetermined, standard counterfactual and open counterfactual) and age (younger, older) and order (six counterbalanced trial orders) as between-subject variables. The dependent variable was the number of trials correct for each trial type.

There was a significant effect of trial, $F(2, 224) = 18.29$, $p < .001$, $\eta^2 = .14$ (Means: undetermined = 1.4, standard counterfactual = 1.9, open counterfactual = 1.5), and a main effect of age, $F(1, 112) = 29.07$, $p < .001$, $\eta^2 = .21$, older children ($M = 1.8$) outperformed younger children ($M = 1.4$). There was a significant interaction between age and trial type, $F(2, 224) = 3.55$, $p = .03$, $\eta^2 = .03$. All other effects were nonsignificant (highest $F = 1.45$, lowest $p = .160$). Means are shown in Table 2. We conducted post hoc t tests, making a Bonferroni correction for nine tests (significance level now $p = .006$) to explore this interaction. There were improvements with age on undetermined, $t(116.87) = -4.94$, $p < .001$, and standard counterfactual trials, $t(94.45) = -3.22$, $p = .002$, but for open counterfactuals the difference only approached significance, having made the strict Bonferroni correction, $t(121.51) = -2.40$, $p = .018$. For the younger (3- and 4-year-old) children standard counterfactuals were easier than undetermined trials, $t(63) = -5.80$, $p < .001$, and easier than open counterfactual trials, $t(63) = 4.06$, $p < .001$. There was no significant difference between undetermined and open counterfactuals, $t(63) = -2.01$, $p = .049$. The same pattern was true for the older children. Standard counterfactuals were easier than undetermined trials, $t(59) = -2.91$, $p = .005$, and open counterfactuals, $t(59) = 3.10$, $p = .003$. There was no difference between undetermined and open counterfactuals, $t(59) = 0.38$, $p = .70$. Thus our main conclusion is that standard counterfactual trials were easier for these children to answer than either open counterfactuals or undetermined trials.

One problem with comparing performance on the open counterfactual and standard counterfactual questions is that the former might be considered more linguistically demanding. Open-ended questions that do not have a single correct answer could be more difficult for children to answer. However, note that Guajardo and Turley-Ames' (2004) antecedent question mentioned earlier is also open ended, and there are many possible antecedents that could lead to the given counterfactual consequent. These authors consider this potential limitation, but are rightly reassured by the very low frequency of irrelevant responses. This gives us confidence that the question is a reasonable one to use with children at a similar age to those in the previous study. However, it remains an interesting avenue for future research to explore the role of linguistic ability in various tests of counterfactual thinking.

Comparing Tables 1 and 2, the children in the younger groups appear to have performed better on undetermined trials in Experiment 2 than in Experiment 1. The procedures in the two studies were exactly the same. We attribute the difference in performance to the use of a slightly older sample in Experiment 2, that may also have been influenced by other sample differences. As mentioned above, the 4 – 5-year-old children in Experiment 1 performed quite poorly on the undetermined trials. In Experiment 2, the 5 – 6-year-olds performed very well (mean score 1.7 out of 2). These older children performed quite well even without prompts to put out a second mat: The mean score for putting out two mats spontaneously prior to the prompt was 1.3 (95% confidence intervals 1.1 – 1.6).

In line with Experiment 1 younger children found undetermined questions relatively hard. In contrast they performed very well on the standard counterfactual questions. This went against the original proposal that thinking about counterfactual alternatives requires the ability to identify the multiple future possibilities that can ensue at any point in time. However, we also speculated that the standard counterfactual question may not demand such recognition. We introduced a new open counterfactual that required children to think about the multiple possibilities. Indeed, this question proved to be significantly more difficult than the standard counterfactual and of similar difficulty to the undetermined trials.

We conclude that children can answer the standard counterfactual questions by thinking about the counterfactual event in isolation, without recognizing its temporal relationship to the actual event. Although children can do this by around the age of 4 years, the ability to entertain multiple possibilities develops somewhat later. The relatively late success on the open counterfactual trials indicates that further developments in children's counterfactual thinking occur after the preschool years.

General Discussion and Conclusions

Do our findings support our original proposal that entertaining counterfactual or future hypothetical thoughts implies recognition of the multiple possibilities that can occur at any one point in time? It is clear that they do not. Even the youngest children performed very well on the standard tasks asking them to speculate about alternative events in the past or future. However, these 3- and 4-year-olds found it particularly difficult to pass our undetermined trials. They did not consistently put out two cotton wool mats to catch the mouse which could go either way, even though their verbal responses to the questions implied that they understood he could take/have taken an alternative route. Younger children's poor performance on the open counterfactual trials is in keeping with their difficulty with multiple possibilities. Put another way, we have no evidence to support the proposal that when 3- and 4-year-olds answer counterfactual and future hypothetical questions they are treating them as possibilities.

Although there is some disagreement in the literature about the age at which children begin to think counterfactually, there is agreement that children can do this by around the age of 4 years. However, some studies have shown that children can have difficulty with counterfactuals in some circumstances. The authors of these studies do not doubt that children are able to think counterfactually, but the findings raise questions about the circumstances under which they are able to this. In the light of our findings, we will now re-examine these studies and offer a new explanation.

First, German and Nichols (2003) have presented data showing that a young sample (3-year-olds) could answer counterfactual questions about short but not long causal chains. In their study, children were presented with a chain of events. For example, Mrs. Rosy is in the garden and calls her husband to come out of the house. When Mrs. Rosy opens the kitchen door, the dog escapes from the kitchen, runs round the garden, and tramples on a flower, which makes Mrs. Rosy sad. Children found it easy to answer a question "What if the dog hadn't squashed the flower, would Mrs. Rosy be happy or sad?" whereas a longer chain of events "What if Mrs. Rosy hadn't called her husband, would Mrs. Rosy be happy or sad?" caused more problems.

The authors interpret this result as showing children's competence entertaining counterfactual thoughts, which is masked by the greater information processing demands of considering a longer causal chain. Critically, this means that children engage (or try to engage) in the same counterfactual thinking when asked both long and short counterfactual questions. But when there are more steps in the story (the long counterfactual), task irrelevant processing demands hamper their ability to complete the process. Presumably the child then falls back on a strategy of giving a realist answer. We propose an alternative explanation, based on the difference found between standard and open counterfactuals. We suggest that the long causal chain requires one to "think back in time" to a point when multiple possibilities could have happened. Children have difficulty with this because it involves thinking about both the actual and counterfactual events as possibilities: for example, Mrs. Rosy could have called or could not have called her husband. The same process of thinking back could solve the short causal chain counterfactual. However, children could in principle give the right answer to short counterfactual chains by imagining only an alternative world and not relating it to the current world. That is, one can reason simply "dog does not squash flowers, therefore Mrs. Rosy happy." In contrast, thinking "Mrs. Rosy does not call her husband" does not lead to an obvious "Mrs. Rosy happy" or "Mrs. Rosy sad" conclusion. The child *must* think back in time and consider the consequences of this change.

Although German and Nichols (2003, p. 515) suggest that difficulty with long chains comes not from "counterfactual reasoning per se but rather . . . the complexity of the inferences required," we argue that the difficulty comes very much from the counterfactual thinking. The long causal chain demands that the child relate the counterfactual and actual, in other words he or she engages in the kind of thinking about multiple possibilities authors like Byrne believe is the essence of counterfactual thinking. Although the sample in German and Nichols' study is younger than ours, we think our analysis of counterfactual possibilities and events offers a new perspective on their results.

Our current data cannot differentiate between our proposal and German and Nichols' information processing account. However, our proposal does offer an alternative explanation of their findings. Importantly, in our study both standard and open questions asked about the same counterfactual possibility, and hence the same length causal chain. To account for our data the information processing account would need to be broadened, such that open questions necessarily posed increased processing demands over standard questions, as well as the increased demands of considering a longer causal chain. Thus, the child has competence in counterfactual thinking, which is sometimes masked by increased information processing demands. Although we accept that open counterfactuals may make different information processing demands, we argue that the open counterfactual requires a qualitatively different type of thinking to the standard question. Direct comparisons between the types of task, perhaps using open counterfactual questions with long and short causal chains, should be used in future research to help resolve these contrasting accounts.

Second, recent research on young children's understanding of emotions has found some surprising limits on performance when emotions involve counterfactual thinking. Understanding of regret (among other complex emotions) requires counterfactual thought. The actual outcome is viewed as negative in comparison with an alternative that could have happened. For example, consider a game in which you are given a choice between two boxes and you receive the prize held in the chosen box. Adult evaluations of the prize received are influenced by the contents of the box not chosen. Thus, opening your box to find a sweet would be pleasing. Yet, on discovering that the other box contains 20 sweets you will likely feel disappointed and regret your choice, thinking "If only I'd picked the other box."

In a recent study, Guttentag and Ferrell (2004) presented children and adults with narratives in which two characters made choices that lead to relatively negative consequences. Although the characters experienced the same outcome, their decision making differed in such a way that leads adults to attribute greater regret to one character. To use the example above (which was indeed one of the narratives in Guttentag and Ferrell's study), if the first character chose one box but then at the last minute changed to the other and the second character chose the same (other) box without changing, the former is judged more regretful when both discover that the unchosen box contains a better prize. Five- and 6-year-olds in this study did not make the same attributions about regret as older children and adults. In one study they failed to differentiate between the two characters in the stories. In a second study, their judgments about who would feel worse were influenced by the alternative event but in a surprising way. Characters who should have felt regret were judged to feel better than those who should have been relieved (the opposite judgment to that made by adults).

It seems that 5- and 6-year-olds' judgments while sometimes reflecting the counterfactual event, did not contrast it with the actual event. The authors (p. 770) suggest that although 5-year-olds can identify counterfactual events, "what was missing from their judgment strategy was a comparison of the feelings that the character would experience in the two situations and an understanding that a positive alternative would intensify the negative feelings produced by the outcome that actually occurred." However, they acknowledge that they do not yet have an explanation for what would change between 5 and 7 years. We suggest that our results may offer one possible explanation for the kind of thinking that is needed to understand these emotions.

In line with Guttentag and Ferrell's analysis (and the findings in the literature), 5-year-old children in our study found it very easy to identify a counterfactual event in response to our standard counterfactual question "What if X had happened?" Children did not have any difficulty setting aside the current reality and working out the consequences of an alternative event. However, when we used a question (the open counterfactual) that required that they think back in time to a point where either event could have happened, and to view the two alternatives as possibilities, children encountered more problems. Our suggestion is that understanding regret involves the same type of understanding as our open question. In both, it is critical that one appreciates that there was a point in time when either the actual or counterfactual event could have happened.

We might consider whether one could experience regret without this understanding. Even though one could understand that choosing the other box would make you feel happier, without the acknowledgment that there was a point in the past when you made this decision (i.e., when both outcomes were possible) one would not be able to make the comparison between what was chosen and what could have been. Without the comparison or contrast between actual and possible, there is no regret. There may be alternative explanations for why complex emotions are difficult for 5-year-olds. Future research should explore the relationship between our open counterfactual questions and the emotion narratives. But at the very least our findings are in keeping with Guttentag and Ferrell's results.

Putting aside the literature on counterfactual thinking, what of our other experimental trials: future hypotheticals and undetermined? Our findings go against the most obvious interpretation of what children are doing when they answer future hypothetical questions, that is, speculating about an uncertain future in which multiple alternative possibilities could occur. Their relatively poor performance on undetermined trials supports the alternative proposal that children can consider a single future event without necessarily construing it as one *possibility* among others. Interestingly, recent work by Byrne and Egan (2004) suggests that although adults can think about the future as containing multiple possibilities, they tend to speculate about just one possibility when understanding prefactual "if X were to happen in the future then ..." and future indicative conditionals "if X happens in the future then ..." Do adults rely on the same type of thinking as our youngest children unless they are explicitly obliged to consider multiple possibilities? Perhaps even when this competence is gained, it is not always put to use.

Performance on the undetermined trials shows a marked improvement between the ages of 3 and 4 years and 5 and 6 years, with good performance among these older children. The 5- and 6-year-olds's success on these trials was perhaps surprising given that they were modeled on tasks in the ambiguity literature, which children typically find difficult until the age of about 7 years. However, in contrast with many of those tasks, children did not have to make a metacognitive evaluation of their knowledge or the quality of the information: They did not have to say *why* there were two possible interpretations, nor *how* to identify the correct one, but merely had to acknowledge their existence. Even so, and even though children were given prompts to put out a second mat, children younger than 5 years still found the task difficult.

To conclude, we suggest that there are real limitations on 3- and 4-year-olds' thinking about counterfactuals and future hypotheticals. They can pass standard questions in these domains by putting aside what they know about reality and speculating about single events. However, in keeping with the undecidability and ambiguity literature, acknowledging multiple possibilities poses serious problems for them. By around 5 or 6 years children can acknowledge multiple possibilities, as shown by our behavioral undetermined trials, and performance on our open counterfactual questions has improved. Only by this age can children's thinking about future and counterfactual possibilities have the mature quality of speculation about genuinely alternative worlds.

References

Atance, C. M., & O'Neill, D. K. (2005). Preschoolers' talk about future situations. *First Language*, *25*, 5–18.

Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. *Current Directions in Psychological Science*, *9*, 40–45.

Beal, C. R., & Flavell, J. H. (1982). Effect of increasing the salience of message ambiguities on kindergarteners' evaluations of communicative success and message adequacy. *Developmental Psychology*, *18*, 43–48.

Beck, S. R., & Robinson, E. J. (2001). Children's ability to make tentative interpretations of ambiguous messages. *Journal of Experimental Child Psychology*, *79*, 95–114.

Braine, M. D. S., & Rumain, B. (1983). Logical reasoning. In J. H. Flavell & E. M. Markman (Eds.), *Cognitive development* (Vol. III, 4th ed., pp. 266–340). New York: John Wiley.

Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. *The Psychology of Learning and Motivation*, *37*, 105–154.

Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT Press.

Byrne, R. M. J., & Egan, S. M. (2004). Counterfactual and prefactual conditionals. *Canadian Journal of Experimental Psychology*, *58*, 113–120.

Carpendale, J., & Chandler, M. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*, 1686–1706.

Chandler, M., Hallet, D., & Sokol, B. W. (2002). Competing claims about competing knowledge claims. In B. Hofer & P. Pintrich (Eds.), *Personal epistemology*. Mahwah, NJ: Lawrence Erlbaum.

Custer, W. L. (1996). A comparison of young children's understanding of contradictory representations in pretense, memory and belief. *Child Development*, *67*, 678–688.

Dias, M. G., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British Journal of Developmental Psychology*, *6*, 207–221.

Fay, A. L., & Klahr, D. (1996). Knowing about guessing and guessing about knowing: Preschoolers' understanding of indeterminacy. *Child Development*, *67*, 689–716.

Flavell, J. H., Speer, J. R., Green, F. L., & August, D. L. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development*, *46(5)*.

Frye, D. (2000). Theory of mind, domain specificity, and reasoning. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind*. Hove, UK: Psychology Press.

German, T. P. (1999). Children's causal reasoning: Counterfactual thinking occurs for ''negative'' outcomes only. *Developmental Science*, *2*, 442–447.

German, T. P., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science*, *6*, 514–523.

Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, *19*, 53–80.

Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: Age differences in judgments of regret and relief. *Developmental Psychology*, *40*, 764–775.

Harris, P. L., German, T. P., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*, 233–259.

Harris, P. L., & Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, *58(1)*.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.

Kuczaj, S. A., & Daly, M. J. (1979). The development of hypothetical reference in the speech of young children. *Journal of Child Language*, *6*, 563–579.

Leevers, H. J., & Harris, P. L. (2000). Counterfactual syllogistic reasoning in normal 4-year-olds, children with learning disabilities, and children with autism. *Journal of Experimental Child Psychology*, *76*, 64–87.

Leslie, A. M. (1987). Pretence and representation: The origins of ''Theory of mind''. *Psychological Review*, *94*, 412–426.

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

Patterson, C. J., Cosgrove, J. M., & O'Brien, R. G. (1980). Non-verbal indicants of comprehension and noncomprehension in children. *Developmental Psychology*, *16*, 38–48.

Perner, J. (2000). About+belief+counterfactual. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind*. Hove, UK: Psychology Press.

Perner, J., Sprung, M., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development*, *19*, 179–201.

Piéraut-Le Bonniec, G. (1980). *The development of modal reasoning: Genesis of necessity and possibility notions*. New York: Academic Press.

Plumert, J. M. (1996). Young children's ability to detect ambiguity in descriptions of location. *Cognitive Development*, *11*, 375–396.

Richards, C. A., & Sanderson, J. A. (1999). The role of imagination in facilitating deductive reasoning in 2-, 3- and 4-year olds. *Cognition*, *72*, B1–B9.

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, *13*, 73–90.

Robinson, E. J., & Apperly, I. A. (2001). Children's difficulties with partial representations in ambiguous messages and referentially opaque contexts. *Cognitive Development*, *16*, 595–615.

Robinson, E. J., & Beck, S. R. (2000). What is difficult about counterfactual reasoning? In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind*. Hove, UK: Psychology Press.

Robinson, E. J., & Whittaker, S. J. (1987). Children's conceptions of relations between messages, meanings and reality. *British Journal of Developmental Psychology*, *5*, 81–90.

Ruffman, T. (1999). Children's understanding of logical consistency. *Child Development*, *70*, 872–886.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, *80*, 201–224.

Singer, J. B., & Flavell, J. H. (1981). Development of knowledge about communication—children's evaluations of explicitly ambiguous messages. *Child Development*, *52*, 1211–1215.

Speer, J. R. (1984). Two practical strategies young children use to interpret vague instructions. *Child Development*, *55*, 1811–1819.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*, 37–63.