
The Mysterious Emotional Life of Little Red Riding Hood

Paul L. Harris^a · Marc de Rosnay^b · Samuel Ronfard^a

^aDepartment of Psychology, Harvard University, Cambridge, Mass., USA;

^bSchool of Psychology, The University of Sydney, Sydney, N.S.W., Australia

Abstract

It is well-established that children realize around the age of 4 years that someone might approach a situation with a false belief about what it holds in store. Despite this insight, 4- and even 5-year-olds have difficulty in working out what the person will feel about the upcoming situation. They claim for example that even if Little Red Riding Hood mistakenly thinks there is only her grandmother waiting for her inside the cottage and that she knows nothing about the wolf, she will still feel afraid. We review a variety of experiments in which children display this gap between the grasp of a protagonist's emotion as compared to a protagonist's belief. We also describe new findings showing that children's attribution of emotion to a story character is a dynamic process. As the story character approaches an unexpected denouement, children's tendency to misattribute emotion intensifies. By implication, children's emotion attributions are not a fixed or static function of their current theory-of-mind. They fluctuate as the story unfolds.

Copyright © 2014 S. Karger AG, Basel

When Little Red Riding Hood knocks on the door of her grandmother's cottage, does she feel afraid? For the adult who reads the classic story to a young child, the answer is straightforward. No, at that moment in the story, she doesn't feel afraid because she does not yet know what has happened to her grandmother. In fact, even when she goes into the cottage, she still does not fully understand the danger she is in. She is puzzled by what she sees – her grandmother looks strangely different. But she does not yet feel the intense fear that would be all too appropriate if she realized that a hungry wolf is in her grandmother's bed. This tension between the actual situation facing Little Red Riding Hood and her emotional reaction to that situation is obvious to the adult reader. But a solid body of findings shows that it is not obvious to young children. The dramatic irony of the story is lost upon them. They effectively jump the gun by attributing fear to Little Red Riding Hood even before she could know that there is something to be afraid of.

In this chapter, we discuss this fascinating and revealing lacuna in young children's understand-

ing of emotion. We first describe some of the initial findings pointing to children's difficulty in understanding belief-based emotions. Second, we discuss an initially plausible explanation of that difficulty – the hypothesis that children fail to understand Little Red Riding Hood's mistaken beliefs and instead attribute her emotion to the situation that they know she faces. We will show that this goes part of the way toward an explanation but ultimately fails. Third, we review various studies that drive home the fact that there is a considerable lag between children's understanding of a protagonist's mistaken beliefs and their grasp of the emotions that flow from such beliefs. We also describe recent studies highlighting both the tenacity of children's misunderstanding as well as its lability. Finally, we propose an explanation for the overall pattern of findings.

A Nasty Surprise

Our first exploration of this topic was guided by the basic idea that at a certain point in development, children come to think about emotion like cognitive therapists. They come to acknowledge that emotional reactions depend not on the objective features of the situation that a person faces but on the person's appraisal of that situation. Even young preschoolers can make good use of a script-like analysis of what situations elicit particular emotions – for example, they attribute sadness to someone who has just lost a toy or happiness to someone who is about to get a treat [Borke, 1972]. Nevertheless, at some point they presumably appreciate that it is not the objective situation that causes emotion but a person's appraisal of that situation – be it mistaken or accurate – that is the critical determinant of what the person feels.

To examine this idea, we gave 4- and 6-year-olds a set of nasty surprise stories [Harris, Johnson, Hutton, Andrews & Cooke, 1989]. Children were first introduced to Mickey the Monkey, a toy

monkey who was 'always playing tricks on the other animals'. Next, four other animals were presented and the desire of each animal for a particular food or drink was explained. For example, children were shown a toy bear and told: 'Bertie the Bear is very hungry and his favorite snack is Smarties (M&Ms)'. Bertie was then made to leave the scene to go for a walk. Next, the experimenter demonstrated how Mickey played a trick on Bertie. A Smarties box was introduced, Mickey was made to empty out all the Smarties and replace them with stones. The narrative continued: 'Mickey put the Smarties box with stones inside it on Bertie's table. Bertie came home and saw this Smarties box on his table.' Children were then asked three questions about Bertie's emotional reactions: (a) How does he feel when his Mom gives him a box of Smarties? (b) How does he feel when he first looks at the box on the table before he looks inside? (c) How will he feel when he has a look inside and finds that there are stones inside instead of Smarties? Comparable stories were told regarding three other victims of Mickey's mischief.

Across all four stories, both age groups were excellent at answering the first and third questions. They realized that the animals would normally feel happy at receiving a treat and sad at discovering the nasty surprise inside the container. As expected, the second question about the animal's initial reaction on seeing the gift from Mickey was much more challenging. With a few exceptions, the children answered according to one of two patterns. They either systematically claimed that the animal would be sad or they systematically claimed that the animal would be happy. The first pattern was common among 4-year-olds (55%) but infrequent among 6-year-olds (25%) whereas the second pattern was infrequent among 4-year-olds (35%) but frequent among 6-year-olds (70%).

Scrutiny of children's justifications reinforced this age change. Children who claimed that the animal would be sad typically explained that at-

tribution by referring to the actual situation (e.g., 'there's no Smarties' or 'cos of the trick') whereas children who claimed that the animal would be happy often referred to the animal's desire for the treat that was apparently on offer (e.g., 'he likes Smarties') or ignorance of what was actually inside the container (e.g., 'cos she didn't see what's inside'). These results offered support for our initial hunch about the course of development. Younger children typically worked out what the animal was feeling by focusing on reality – the nasty trick that had been played by the mischievous monkey whereas older children were more likely to acknowledge the victim's naïve appraisal of reality.

A very similar pattern emerged in a second experiment in which we gave 4-, 5- and 6-year-olds either four nasty surprise stories – similar to those just described – or four pleasant surprise stories in which a now benign but nevertheless tricky Mickey again doctored the contents of a familiar container but introduced something desirable rather than undesirable. To take a concrete example, one of the tricked animals was Ellie the elephant. Children learned that Ellie liked only one kind of drink. Half the children were told that she liked only coke and the remaining children were told that she liked only milk. During her daily stroll in the jungle, Mickey tricked her by pouring the coke out of a coke can, replacing it with milk, and leaving the coke can on the table to await her return.

Children were asked how Ellie would feel on first spotting the coke can, and how she would feel after taking a sip. We reasoned that the second question was relatively straightforward: Children needed only to work out whether the actual contents – which she would discover on taking a sip – corresponded to Ellie's preferred drink – be it coke or milk. More specifically, as long as children realized that desires play a key role in determining emotional reactions, they would be expected to answer this question correctly – to appreciate that Ellie would get a nasty surprise if she

preferred coke but a pleasant surprise if she preferred milk.

We expected the first question, by contrast, to prove more challenging. To answer correctly, children had to keep in mind Ellie's preferred drink – a preference that did not necessarily match their own. In addition they had to keep in mind whether or not the expected contents – coke, since it was a coke can – matched that preference. Effectively, children needed to coordinate their knowledge of Ellie's desire for a particular drink with an understanding of her mistaken presumption about the contents of the can.

Our expectations were borne out. Children in all three age groups performed very well on the second question. With few exceptions they answered correctly for all four stories. By implication, they had no difficulty in remembering the animals' individual preferences (e.g., for coke not milk – or vice versa) and in realizing whether a particular animal would end up having a nasty or pleasant surprise. Answers to the first question, by contrast, produced a marked age change. More than half of the 4-year-olds ignored the animals' mistaken belief and replied in terms of reality (i.e., what they knew to be inside the container). Only 19% managed to give correct replies on three or more stories (out of 4 trials). Among the 5- and 6-year-olds this percentage climbed to 44 and 75%, respectively. Using a much more liberal criterion – answering correctly on at least one trial – a similar change with age was observed. Thus, 38% of 4-year-olds, 63% of 5-year-olds and 88% of 6-year-olds were correct at least once.

The pattern of justification was very similar to that observed in the first experiment. Children tended to focus on either the reality of the situation or on the animal's mistaken appraisal of that reality with the latter pattern gradually displacing the former in the course of development. Taken together, these findings lent strong support to our expectation that children would increasingly realize that a person's emotion depends less on their actual circumstances than on their appraisal of

those circumstances – even if the appraisal is ill-founded – and indeed will soon prove to be so. Setting these findings in the context of the story of Little Red Riding Hood, we would expect 4-year-olds to miss the protracted tension in the story – to assume prematurely that Little Red Riding Hood is afraid of the wolf.

On reflection, however, we realized that this straightforward developmental account overlooked something important. We shared our findings with Josef Perner who readily saw parallels with the classic false belief task that he and Heinz Wimmer had introduced some years earlier [Wimmer & Perner, 1983]. Still, he also pointed out that the parallel was less than exact. At the time, and indeed ever since [Wellman, Cross & Watson, 2001], the evidence shows that many 4-year-olds and most 5-year-olds do well on standard false belief tasks. By contrast, the success rate for our nasty and pleasant surprise tasks was a good deal poorer. For some mysterious reason, children found it hard to appreciate the impact of beliefs on emotion.

The puzzling nature of this lag did not fully sink in until it was dramatically highlighted in later experiments. But before describing those experiments it is worth underlining one initially appealing explanation of the lag. It might be argued that standard false belief tasks are relatively simple because they call only for a diagnosis of the protagonist's mistaken belief. By contrast, surprise tasks, whether nasty or pleasant, call for a diagnosis not just of the protagonist's mistaken belief but also a further inference about the protagonist's belief-based emotion. Perhaps it is this extra inferential step that creates the lag between the pattern of development for the standard false belief task and the pattern of development for the belief-based emotion tasks described above. However, this account ignores a crucial point. In most belief tasks [including Wimmer & Perner, 1983], children are not asked directly what the protagonist believes. Rather they are invited to work out what the protagonist will do – for ex-

ample, where he or she will search – in the box or the cupboard. Alternatively, they are invited to work out whether a protagonist will say that there are pencils or M&Ms in a container [Gopnik & Astington, 1988]. In other words, the standard false belief tasks also call for an inferential step, whether it is from belief to action or from belief to utterance. So, the added difficulty of the emotion tasks cannot be due only to the addition of an inferential step. Of course, it might be the case that inferential steps are not equal in difficulty: The link between belief and emotion might be harder for children to understand than the link between belief and action or between belief and utterance. Still, that line of explanation – which we will return to later – implies that the problem has to do with the *nature* of the inference and not with making the inference per se.

What Does Little Red Riding Believe?

If there is indeed a developmental lag between children's diagnosis of a belief and their diagnosis of an ensuing belief-based emotion, we can predict that individual children will pass through a paradoxical phase in which they correctly diagnose the belief but incorrectly diagnose the emotion. Compelling evidence of this paradox was reported by Bradmetz and Schneider [1999]. Using the story of Little Red Riding Hood as a vehicle, they questioned children aged 3–6 years about her thoughts and feelings. For example, they asked, 'When Little Red Riding Hood goes into grandmother's house, does she think the wolf is in the bed or does she think the grandmother is in the bed?' and 'When Little Red Riding Hood goes into grandmother's house does she feel afraid? Why?'

Almost half of the 3- and 4-year-olds answered both questions incorrectly, albeit coherently – they said that Little Red Riding Hood thought the wolf was in the bed and felt afraid. Conversely,

nearly half of the 5- and 6-year-olds answered both questions correctly – they said that Little Red Riding Hood thought her grandmother was in the bed and did not feel afraid. The remaining children displayed the anticipated paradox – they diagnosed Little Red Riding Hood’s belief correctly but not her emotion. They said that she thought it was her grandmother in the bed but then went on to claim that she felt afraid.

A similar pattern emerged in two follow-up studies by Bradmetz and Schneider [1999]. One was based on a story with a similar structure to Little Red Riding Hood involving a wolf that tricked a little goat. The other was closely modeled on a standard false belief task – Maxi, the main character, put his chocolate in a container and left the scene. In Maxi’s absence, his brother ate most of the chocolate but put a small remaining piece in a different container. On Maxi’s arrival at the door of his house, children were asked: ‘Maxi is in front of the door, where will he look for his chocolate?’ and ‘When Maxi is in front of the door, is he happy?’.

In both studies, a *décilage* between replies to the two test questions was observed: Although children often replied incorrectly to both questions – or replied correctly to both questions – approximately half the children answered the belief question correctly and the emotion question incorrectly. The reverse pattern was never observed. By implication, a correct reply to the belief question is a necessary but not a sufficient condition for a correct reply to the emotion question.

Another experiment involving 6- to 8-year-olds further underlined children’s persistent difficulty in appreciating the contradiction between their replies to the two questions. When children correctly said where Maxi would search, the interviewer posed the emotion question and then – depending on how children replied – offered a counter-suggestion. For example, children who incorrectly claimed that Maxi felt unhappy were asked whether he knew that his brother had eaten the chocolate. Children who correctly claimed

that Maxi felt happy were reminded that his brother had eaten the chocolate. Then, in each case, the emotion question was repeated. These counter-suggestions had an impact but only in one direction. Among the children who had initially answered the emotion question correctly almost half changed their mind – reverting to an incorrect answer. By contrast, all the children who had initially answered the emotion question incorrectly – by claiming that Maxi was unhappy – clung to this incorrect answer.

Taken together, this series of experiments provides strong evidence for a paradoxical pattern of responding between the ages of 3 and 8 years. Children may appreciate the protagonist’s mistaken belief but they do not necessarily work out the implications of that belief for what the protagonist feels – and even when they have done so, a counter-suggestion can easily unsettle them. To explain the gap between children’s understanding of belief and their understanding of emotion, Bradmetz and Schneider [1999] proposed that emotional cues are vivid – their message has a high priority. Presumably, what these authors mean by this claim is that the affective implications of the situation facing the story character – whether or not he or she is aware of the situation – are strong and compelling for young children as they listen to the story. For example, as they listen to the story of Little Red Riding Hood they are likely to regard the wolf inside the grandmother’s cottage as a strong and compelling cue for fear. That association between the wolf and the emotion of fear might be sufficiently rapid and vivid that children then attribute fear to Little Red Riding Hood even though they can work out that she was unaware of the presence of the wolf.

Indirect support for this line of interpretation can be gathered by thinking back to the influential findings of Zajonc [1980] on affective processing. He argued that the affective valence of a stimulus can be processed rapidly and efficiently. Indeed, in some cases, such processing can occur

before the stimulus has even been recognized. For example, when presented with geometric shapes for brief intervals, adults did no better than chance at saying which stimuli they recognized as old versus new. Despite this absence of recognition, subjects expressed a preference for stimuli that they had seen before over those they had not.

On the other hand, when we take into account the full range of findings, an explanation for the belief-emotion gap in terms of the rapid and compelling nature of affective reactions begins to seem less plausible. Consider the experiment described earlier in which children listened to stories about animals with fairly idiosyncratic, and indeed narrow, preferences for a particular food or drink [Harris et al., 1989, experiment 2]. For example, Harry the Horse was described as liking only one kind of snack – either peanuts (for half the children) or chewing gum (for the other half). In either case, Mickey the monkey tricked Harry by replacing the contents of a peanuts packet with chewing gum. Can we maintain that chewing gum will immediately have negative implications for the children who have been told that Harry likes peanuts but positive implications for the children who have been told that he likes chewing gum? To the extent that children recognized that chewing gum has different implications for Harry depending on his desires, it would seem wrong to insist that it carries an automatic and inevitable affective message. We need to recognize that even younger children fine-tune the affective message depending on the particular, idiosyncratic desires of the protagonist.

To summarize the pattern of results so far, then, young children are oddly incoherent in their reasoning about mental states. On the one hand, they understand by 4 or 5 years of age that a protagonist may misconstrue reality – for example, by mistakenly thinking that there are Smarties in a Smarties box, that grandmother is in her bed, or that chocolate is in the same place as before. On the other hand, children of this age are prone to attribute emotion to that same pro-

tagonist on the basis of what they themselves know to be really the case, ignoring what the protagonist thinks to be the case. We have examined two initially plausible explanations and found them wanting. The proposal that emotion attributions are difficult because they call for some extra inferential step appears to be inadequate – standard tests of belief attributions also call for some extra inferential step – linking the belief to an action or utterance. Similarly, the proposal that emotion attributions might be disrupted by affectively charged cues that are rapidly processed sounds plausible for lurking wolves but less plausible for an unexpected packet of chewing gum that would disappoint one animal but please another. In the next sections, we discuss whether children show this lag on other belief-emotion problems as well as describe relations between performance and individual differences in language ability and input. Finally, we seek to offer a more adequate explanation of the pattern of results.

Beliefs, Emotion, and Language

Attachment theorists have suggested that children's understanding of emotion is likely to depend on the extent to which they can express, think about, and discuss their feelings in the context of a secure attachment relationship [Cassidy, 1994]. In an exploration of such individual differences, De Rosnay and Harris [2002] asked whether attachment security was linked to performance on different variants of the nasty surprise task and this offered an opportunity to ask a simple but important question. Do individual children perform in a consistent fashion across different variants of the task? In an initial examination of this question, children ranging from 3 to 6 years were given three different tasks. The dog-rabbit task was very similar to the nasty surprise tasks devised by Harris and his colleagues [Harris et al., 1989]. Children were introduced to Roger Rabbit

and told that he likes Smarties and dislikes peanuts. Gromit the dog then played a trick on Roger by removing Smarties from a Smarties box, replacing them with peanuts, and leaving the box to await Roger's return.

A second task was quite different in both content and format. Children saw a short video clip showing a toddler left alone in a waiting room while her mother went to another room for an appointment. At a certain point, there was a knock at the door but to the toddler's disappointment, it was a stranger who came in rather than her mother. Children were asked how the toddler felt on first hearing the knock at the door. The rationale for this choice of content was that the video played on attachment-related themes, notably separation from the mother and a hoped-for reunion. Accordingly, it might be expected that secure children would be especially good at understanding the toddler's mistaken hope that his mother was returning.

The third task was similar to the second but more emotionally charged. Children saw a video clip showing a toddler left alone in a waiting room by his mother but in this case the toddler reacted to her departure with considerable distress. A knock at the door again proved misleading – and upsetting – because it was a stranger who entered rather than the mother.

Despite the variation among the three tasks with respect to the identity of the main protagonist, the nature of the protagonist's mistaken expectation, the degree of expressed emotion, and the format of the presentation, children proved to be quite stable in their performance on the key test questions in which they were asked to predict and explain how the protagonist felt (when seeing the Smarties box or when hearing the knock at the door). Thus, when performance on any two of the three tasks was compared, children tended to either pass both or fail both tasks. Indeed, performance on the Dog-Rabbit task and the more emotionally charged maternal separation task was highly concordant, with 44 of the

total sample of 51 children (86%) either failing both or passing both. Note that these results cast further doubt on the explanation in terms of emotional vividness discussed earlier. That is to say, the prospect of disappointment would seem to be more compelling for the maternal separation video depicting a distressed toddler hearing a stranger's knock at the door compared to the prospect of disappointment in the case of Roger Rabbit. After all, he is simply getting peanuts (a snack that children might actually like) instead of getting Smarties. Yet, children responded similarly in each case.

Nevertheless there were considerable individual differences among the children in their overall performance – for example, some children as young as 3 years 11 months were able to identify the protagonist's 'mistaken' emotion and to explain it appropriately for one of the three tasks whereas some children as old as 5 years 10 months failed to answer these two questions for all three tasks. A regression analysis highlighted the important role of language ability (as measured by the BPVS, a measure of receptive vocabulary) and also – consistent with the earlier findings of Fonagy, Redfern and Charman [1997] – the role of attachment security, as measured by the Separation Anxiety Test [Klagsbrun & Bowlby, 1976], in which children are invited to discuss the feeling of a protagonist who is depicted as experiencing separation from one or both parents. Neither chronological age nor gender made a significant independent contribution to emotion understanding. By contrast, verbal mental age independently explained a considerable portion of the variance in emotion understanding (27.2%), and attachment security accounted for a smaller but significant portion (8.4%).

A follow-up analysis of these same children cast more light on their difficulties with the task and also on the facilitating role of language ability [de Rosnay, Pons, Harris & Morrell, 2004]. In line with the findings of Bradmetz and Schneider [1999], there was clear evidence that an under-

standing of false belief may be necessary for an understanding of belief-based emotion but is not sufficient. On each task, very few children answered the emotion question correctly and the belief question incorrectly. By contrast, a considerable proportion of the children gave the wrong answer to the emotion question – failing to realize that the protagonist would feel happy – but correctly attributed a mistaken belief to the protagonists. Thus, when asked about what Roger Rabbit thought was in the box or who the toddlers thought was knocking at the door, children managed to answer correctly. Furthermore, children’s language ability proved to be a predictor both of correct replies to the false belief question and also of correct replies to the emotion question. Indeed, the relation between language and correct replies to the emotion question held up even when the contribution of language to performance on the false belief question was taken into account. Making this same point differently, it appears that language serves as a stepping-stone at two successive points. It first contributes to children’s understanding of false belief and then it makes an additional, further contribution to their understanding of belief-based emotion [Harris, de Rosnay & Pons, 2005]. These findings are, of course, consistent with the more general point that has emerged repeatedly, namely that children’s insight into a protagonist’s false belief about a situation is no guarantee that they will proceed to a correct diagnosis of the emotion triggered by such a false belief.

To consolidate this pattern of findings, De Rosnay et al. [2004] conducted a further study in which they measured not just children’s understanding of beliefs and belief-based emotions (using the Dog-Rabbit task and the emotionally charged maternal separation task) but also the nature of the language environment that children were exposed to. The experimenter invited mothers to describe their children (aged between 4.5 and 6 years) with an open-ended prompt (‘Can you describe [child’s name] for me?’). The fre-

quency with which mothers talked about their children’s mental life was assessed. As in the earlier studies, there was a sizeable group of children who correctly diagnosed the protagonist’s mistaken belief but failed to attribute the emotion that would be appropriate given that mistaken belief. For example, in the maternal separation task, 33% displayed this pattern whereas only 9% showed the reverse pattern (i.e., correctly attributed the emotion but not the mistaken belief). As usual, those children who attributed the wrong emotion cited the emotion that would be likely given the actual situation facing the protagonist. However, the gap between emotion understanding and belief understanding was less marked among children whose mothers were more ‘mind-minded’ in that they often talked often about their child’s mental life when describing them [Meins, Fernyhough, Russell & Clark-Carter, 1998]. More specifically, mothers whose descriptions of their child included proportionately more references to the child’s thoughts and feelings rather than to the child’s behavioral or physical attributes had children who answered correctly to both the belief and emotion questions.

Summarizing, children are fairly stable in their grasp of belief-based emotions. They show the same pattern of correct – or incorrect – attribution across a variety of tasks. Correct attribution of emotion is associated with language in two ways. Children with greater language ability and children with mothers who use more mental-state language make more correct attributions. Moreover, the contribution of language cannot be explained by its well-known association with an understanding of false belief [Astington & Baird, 2005]. The contribution to an understanding of emotion emerges even when prior allowance is made for a contribution to the understanding of false beliefs. This pattern of findings reinforces the conclusion set out earlier; the understanding of belief-based emotion calls for some insight or ability that goes beyond an understanding of false beliefs.

Understanding Our Own Emotions

The studies so far have focused on children's ability to figure out what someone else feels – whether it is Maxi, Little Red Riding Hood, Roger Rabbit or a toddler in a video. But we can also ask how far children are able to report their own emotions accurately, especially when those emotions are based on a misconstrual. Consider the following variant on another classic false belief task. Children are shown a familiar box of Smarties and asked to indicate how they feel about eating what's inside. Next, they are shown the contents – which turn out to be inedible beads rather than candy. The beads are then poured back into the box and children are asked what they initially thought was inside the box and also how they felt about eating it.

Depending on how we identify our emotions, two different outcomes seem feasible. Suppose that particular emotions such as happiness or sadness are natural kinds, each associated with a distinctive inner feeling, a particular pattern of physiological arousal, and a distinctive mode of expression via the face, voice and posture. Classic theories of emotions, rooted in Darwin's evolutionary approach, have long advocated such distinctive emotional states [Ekman, 1999; Ekman, Campos, Davidson & de Waal, 2003]. On this view, when first shown the Smarties box, children would be likely to identify their emotional state – for example, their emotion at the prospect of eating Smarties – based on its distinctive inner quality. Subsequently, when they are asked to think back to how they felt before learning the actual contents of the Smarties box, they should be able to remember that specific emotional state – just as they might remember other feeling states, such as a feeling of thirst or an itch in the neck.

However, suppose instead that there is an inextricable link between appraisal processes and emotional experience. More specifically, suppose that a given emotional experience is constructed on the basis of an appraisal of the situation that

one faces [Lindquist & Barrett, 2008]. In that case, faced with a Smarties box and the prospect of eating the candy inside it, children might appraise the situation positively and judge themselves to be happy – not so much because of some identifiable inner glow of happiness but because they know that such a feeling is appropriate to the pleasant prospect of eating candy. On this view, it might be difficult for children to reconstruct their feelings about the prospect of Smarties, once that prospect is no longer in the offing. An accurate reconstruction of how they felt would call for: (a) an accurate retrieval of how they first viewed the Smarties box, and (b) an inference as to what they would have felt given that mistaken appraisal. But rather than engage in such a two-step accurate reconstructive process, children might instead take their current appraisal as a guide. Because they now view the box negatively – after all it contains only inedible beads – they may claim to have felt negatively even when first shown the box.

To compare these alternative accounts, Bender and his colleagues [Bender, Pons, Harris & de Rosnay, 2011] interviewed 5.5-year-olds and 7-year-olds. When presented with a nasty surprise along the lines just described, a considerable proportion of children especially in the older group displayed the by now familiar pattern. Although they correctly stated that they had initially thought there were Smarties in the box, they incorrectly stated that they had felt sad about eating the contents. Thus, even with respect to their own emotions, children were prone to report an emotion that was appropriate to what lay in store for them but inappropriate given their ignorance at the time. By implication, children did not report on their past emotion by delving into their memory and retrieving the record of some inner glow when they were first offered the Smarties box. Their report of their past emotion was contaminated by knowledge of what they discovered later – even if this knowledge did not contaminate their report of their past belief.

Overview and Interpretation

Summing up the pattern of findings, across a variety of different procedures and laboratories, young children around the age of 4 and 5 years can often do quite well in appreciating that someone who lacks the relevant perceptual access to a given situation may misconstrue that situation. Thus, children of this age realize that someone might think that a situation is rewarding or safe when, in fact, it is not. Despite this facility in understanding how a person – including themselves – might not know about what is in store for them, children frequently misattribute emotions to that person. They attribute emotions that make sense only in light of the impending situation. It is as if children look too far ahead, ignoring the fact that the person does not yet know what lies in the future. The likelihood of these ‘premature’ attributions varies from child to child. Two different language measures have been shown to account for some of this variation. Children with superior language ability and children whose mothers make more references to mental states are less prone to making such misplaced attributions of emotion.

As discussed earlier, it is tempting to explain these errors in terms of the extra inferential step required when imputing a belief-based emotion – as opposed to a belief *per se*. Yet there is no indication that children invariably have difficulty with such additional inferencing: They readily work out what someone will say or do on the basis of a false belief. Similarly, although it is plausible that the vivid emotional implications of an upcoming surprise – such as the presence of a wolf – are rapidly processed and highly salient, it is less plausible that the affective implications of the machinations carried out by Mickey the Monkey or Roger Rabbit (e.g., the replacement of Smarties by peanuts) are so highly salient, especially when the valence of those implications depends on the particular preferences of the victim (e.g., for Smarties over peanuts or vice versa).

However, it is possible to construct a more complex account that builds on these ideas while taking care of the aforementioned objections. We may suppose that children ordinarily appraise an imminent situation in light of their own ongoing goals and preferences. However, it is also plausible that when thinking about another person – whether it is a friend, a child in a video, or a story protagonist – children readily adopt a different appraisal strategy. They evaluate the situation that the person is facing in light that person’s ongoing goals and preferences – not their own. Indeed, Repacholi and Gopnik [1997] showed that this flexibility emerges at an early age: having seen an adult express a preference for broccoli over crackers, toddlers who responded to the adult’s request for food typically offered broccoli – even though the adult had not identified which particular food they were requesting.

So, when asked to say how a person feels, it is plausible that young preschoolers are able to attribute emotions based on a rapid appraisal of whether the situation that the person is facing matches that person’s desires or not. Making the same point differently, an inhibitory process will be needed for that desire-based attribution process to be overridden. Knowing that the person is, in fact, unaware of the actual nature of the situation can, in principle, trigger such an inhibitory process but – as we have seen repeatedly – it may fail to do so. On this account, from 4 years of age and upward, children make progress in attributing emotion not in their ability to analyze what the person thinks, but in their ability to use the outcome of that analysis to override a powerful, desire-based analysis of the person’s emotion.

This hypothesis implies that children will vary in the emotion that they attribute depending on whether the situation that the protagonist is approaching is a long way off – or imminent. When the situation is imminent, its positive and negative features – as appraised through the lens of the protagonist’s desires – will be highly salient. Hence, the inhibitory process may not be effective

enough to suppress an attribution of emotion based on its perceived valence for the protagonist. On the other hand, when the upcoming situation is more distant, the inhibitory process is more likely to succeed because the positive or negative features of the upcoming situation will be less salient.

We have recently tested this idea in a series of studies [Ronfard & Harris, 2013]. Children were reminded of the story of Little Red Riding Hood and then invited to consider her feelings as she made her way from her home toward her grandmother's cottage. More specifically, a figurine depicting Little Red Riding Hood was placed at four successive locations (at her own house, part way toward Grandmother's house, still closer to Grandmother's house and outside Grandmother's front door). Children were asked to say how Little Red Riding Hood felt at each successive location. In two experiments, a marked effect of distance was observed. Children were more prone to claim that little Red Riding Hood was afraid the closer she got to her destination. In a third experiment, children were asked about a protagonist who set off toward his own house, where unbeknownst to him, all his friends were waiting to play with him. A similar distance effect was observed but this time, it was misattributions of joy that became more frequent as the protagonist approached his destination.

An important implication of these findings is that children's attribution of emotion is not a fixed and direct function of their level of theory-of-mind understanding. In all three of the studies just described, individual children gave different answers depending on the distance of the protagonist from the upcoming surprise contrary to what we would expect if children's replies were entirely constrained by their theory-of-mind. The greater the immediacy of that surprise, the more children erred in their attributions. An important question for future research concerns the status of these fluctuating attributions. Should we insist that children 'really' do understand belief-based

emotions as shown by their more accurate performance when the surprise is further away? On this view, the growing salience of the surprise masks what children really know about belief-based emotion. Alternatively, should we acknowledge that children do not actually understand belief-based emotions given their misattribution of emotion when a surprise is imminent?

The best way to resolve this dilemma is to think more carefully about how children make developmental progress between 4 and 6 years. According to one possibility – the inhibition account – progress simply calls for greater inhibitory control as indexed by standard tasks such as the day-night task [Carlson & Moses, 2001]. By implication, 4- and 5-year-olds do understand belief-based emotions but lack sufficient inhibitory control to set aside interference from the emotional implications of an upcoming surprise – and that shortfall is increasingly obvious the closer and more immediate the surprise. A different possibility is that development is primarily conceptual – it involves an increasingly firm grip on the idea that emotions flow from appraisal processes so that situations lying beyond a protagonist's awareness are not allowed to infect the emotion attribution process. According to this hypothesis, 4- and 5-year-olds might not genuinely understand belief-based emotions even when the protagonist is at far distances. The greater likelihood of correct attributions at more distant locations occurs primarily because the emotional implications of the upcoming surprise are less salient to children; therefore they cause less interference in making attributions to the protagonist.

To assess these two alternatives, it is helpful to think about the exact sequence and timing of the mental steps involved in attributing emotion. We may speculate that there are two different processes involved: a rapid, quasi-empathic process that specifies the emotional valence of an upcoming outcome in relation to the desires of a particular protagonist; and a slower, more conceptual

and effortful process that first specifies which outcome the protagonist anticipates (especially in cases where the protagonist expects something different from the actual outcome) and then proceeds to specify the emotional valence of that anticipated outcome for the protagonist.

This model makes the following predictions: (a) children who mistakenly attribute fear to Little Red Riding Hood will do so quickly because they rely only on the faster process described above; (b) such mistaken attributions will be increasingly fast as the protagonist gets closer because the surprise will be increasingly salient to young children as they think about the movement and destination of the protagonist and notably his or her increasing proximity to the upcoming surprise; (c) the correct attribution of happiness to Little Red Riding Hood will be slower than incorrect attributions of fear because correct attributions of happiness call for the slower, reflective process described above, and, finally (d) the speed of such correct attributions will likely be slower at closer distances given the possibility of increased interference from the actual, upcoming situation in working out what exactly the protagonist anticipates.

How might we conceptualize developmental progress between 4 and 6 years of age in light of this model? One possibility is that both the faster, empathic process and the slower more conceptual process operate throughout that developmental period but that children become increasingly adroit at inhibiting and setting aside the output of the fast process and relying instead on the output of the slower process. Note that this corresponds to the assumptions of the inhibition account described above. If this account is correct, we can plausibly expect children to take more time to reply as they get older and shift from incorrect to correct replies. However, a second possibility is that the slower, more reflective process is only gradually put in place and becomes increasingly efficient and less effortful as children get older – effectively winning the race with the simpler empathic process to supply an answer to

the attribution question. This would be consistent with the second, more conceptually based account of development proposed above. On this hypothesis, we can plausibly expect that once children start to provide correct attributions (e.g., of no fear to Little Red Riding Hood), they will take increasingly less time to do so as compared to younger children because the slower, reflective process improves in efficiency.

In future research, we anticipate testing these various predictions by studying not just the accuracy or inaccuracy of children's attributions of emotion but also the speed with which they make those attributions. We also envisage examining the contribution of two other factors, one that we have already considered and one that we have touched on only in passing. If our analysis is correct, it should be feasible to alter the salience of the upcoming surprise for individual children. More specifically, depending on whether we ask children to attribute an emotion to the protagonist when he or she is at some distance from the upcoming surprise or alternatively, when it is imminent, the salience of that surprise should vary. In addition, if our emphasis on inhibitory control is correct, it is plausible that measures of individual differences in that ability will predict the ease or difficulty with which children inhibit their inclination to make attributions to the protagonist in light of the more rapidly, empathic process. We speculate that the pattern of performance elicited by variation in these two factors – relative distance from the surprise and inhibitory control may ultimately look quite similar. For example, children with superior inhibitory control would look similar to children tested when the protagonist is at some distance from the surprise and children with inferior inhibitory control would look similar to children tested when the protagonist is close to the surprise.

In conclusion, we are impressed by a nice paradox. Even if – as we have seen – children stumble in their attribution of belief-based emotions, they do make considerable progress in the space of 2 or 3 years. Developmental psychologists studying that

progress, and seeking to understand it, are quite slow by comparison. After a quarter of a century of puzzlement there are still unanswered questions. Still, we like to think we are moving forward – even if we are not exactly sure what lies in store.

Acknowledgement

The authors gratefully acknowledge helpful feedback from Kristin Hansen Lagattuta and Henry Wellman.

References

- Astington, J.W., & Baird, J.A. (Eds.). (2005). *Why language matters for theory of mind*. New York: Oxford University Press.
- Bender, P.K., Pons, F., Harris, P.L., & de Rosnay, M. (2011). Do young children misunderstand their own emotions? *European Journal of Developmental Psychology, 8*, 331–348.
- Borke, H. (1972). Interpersonal perceptions of young children: Egocentrism or empathy. *Developmental Psychology, 7*, 104–106.
- Bradmetz, J., & Schneider, R. (1999). Is Little Red Riding Hood afraid of her grandmother? Cognitive vs. emotional response to a false belief. *British Journal of Developmental Psychology, 17*, 501–514.
- Carlson, S.M., & Moses, L.J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.
- Cassidy, J. (1994). Emotion regulation: Influences of attachment relationships. In N. Fox (Ed.), *The development of emotion regulation: Biological and behavioral considerations. Monographs of the Society for Research in Child Development, 59*, 228–249.
- De Rosnay, M., & Harris, P.L. (2002). Individual differences in children's understanding of emotion: The roles of attachment and language. *Attachment and Human Development, 4*, 39–54.
- De Rosnay, M., Pons, F., Harris, P.L., & Morrell, J.M.B. (2004). A lag between understanding false belief and emotion attribution in young children: Relationships with linguistic ability and mothers' mental state language. *British Journal of Developmental Psychology, 22*, 197–218.
- Ekman, P. (1999). Basic emotions. In T. Dalgliesh & T. Power (Eds.), *Handbook of cognition and emotion* (pp. 45–60). New York: Wiley.
- Ekman, P., Campos, J.J., Davidson, R.J., & deWaal, F. (2003). *Emotions inside out: 130 Years after Darwin's the expression of the emotions in man and animals*. New York: New York Academy of Science.
- Fonagy, P., Redfern, S., & Charman, T. (1997). The relationship between belief-desire reasoning and a projective measure of attachment security (SAT). *British Journal of Developmental Psychology, 15*, 51–61.
- Gopnik, A., & Astington, J.W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59*, 26–37.
- Harris, P.L., de Rosnay, M., & Pons, F. (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science, 14*, 69–73.
- 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.
- Klagsbrun, M., & Bowlby, J. (1976). Responses to separation from parents: A clinical test for young children. *British Journal of Projective Psychology & Personality Study, 21*, 7–27.
- Lindquist, K.A., & Barrett, L.F. (2008). Constructing emotion: The experience of fear as a constructive act. *Psychological Science, 19*, 893–903.
- Meins, E., Fernyhough, C., Russell, J., & Clark-Carter, D. (1998). Security of attachment as a predictor of symbolic and mentalising abilities: A longitudinal study. *Social Development, 7*, 1–24.
- Repacholi, B.M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12–21.
- Ronfard, S., & Harris, P.L. (2013). When will Little Red Riding Hood become scared? Children's attribution of mental states to a story character. *Developmental Psychology, 49*, 12–21.
- Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*, 655–684.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.
- Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.

Paul L. Harris
Harvard Graduate School of Education
503A Larsen Hall, Appian Way 14
Cambridge, MA 02138 (USA)
E-Mail paul_harris@gse.harvard.edu