The development of regret and relief about the outcomes of risky decisions

McCormack, Teresa; O'Connor, Eimear; Beck, Sarah; Feeney, Aidan

DOI:
10.1016/j.jecp.2016.02.008

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
McCormack, T, O'Connor, E, Beck, S & Feeney, A 2016, 'The development of regret and relief about the outcomes of risky decisions', *Journal of Experimental Child Psychology*, vol. 148, pp. 1–19. https://doi.org/10.1016/j.jecp.2016.02.008

Journal of Experimental Child Psychology 148 (2016) 1–19

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

The development of regret and relief about the outcomes of risky decisions

Teresa McCormack [a,*], Eimear O'Connor [a], Sarah Beck [b], Aidan Feeney [a]

[a] School of Psychology, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK
[b] School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:
Received 23 October 2015
Revised 7 February 2016
Available online 17 April 2016

Keywords:
Regret
Counterfactual thinking
Relief
Decision making
Risk
Development

ABSTRACT

Although a number of studies have examined the developmental emergence of counterfactual emotions of regret and relief, none of these has used tasks that resemble those used with adolescents and adults, which typically involve risky decision making. We examined the development of the counterfactual emotions of regret and relief in two experiments using a task in which children chose between one of two gambles that varied in risk. In regret trials they always received the best prize from that gamble but were then shown that they would have obtained a better prize had they chosen the alternative gamble, whereas in relief trials the other prize was worse. We compared two methods of measuring regret and relief based on children's reported emotion on discovering the outcome of the alternative gamble: one in which children judged whether they now felt the same, happier, or sadder on seeing the other prize and one in which children made emotion ratings on a 7-point scale after the other prize was revealed. On both of these methods, we found that 6- and 7-year-olds' and 8- and 9-year-olds' emotions varied appropriately depending on whether the alternative outcome was better or worse than the prize they had actually obtained, although the former method was more sensitive. Our findings indicate that by at least 6 or 7 years children experience the same sorts of counterfactual emotions as adults in risky decision-making tasks, and they also

* Corresponding author.
E-mail address: t.mccormack@qub.ac.uk (T. McCormack).

http://dx.doi.org/10.1016/j.jecp.2016.02.008

suggest that such emotions are best measured by asking children to make comparative emotion judgments.

Introduction

There has been a recent surge of research interest in the development of counterfactual thinking (for reviews, see Beck & Riggs, 2014; Rafetseder & Perner, 2014). Some of this research has focused on the development of emotions thought to require counterfactual thinking abilities, specifically regret and relief (Burns, Riggs, & Beck, 2012; McCormack & Feeney, 2015; O'Connor, McCormack, & Feeney, 2012; O'Connor, McCormack, & Feeney, 2014; Rafetseder & Perner, 2012; van Duijvenvoorde, Huizenga, & Jansen, 2014; Weisberg & Beck, 2010; Weisberg & Beck, 2012). Researchers studying early and middle childhood have primarily focused on attempting to pinpoint the age at which children first experience these emotions (O'Connor et al., 2012; Rafetseder & Perner, 2012; van Duijvenvoorde et al., 2014; Weisberg & Beck, 2010, 2012). Although this research has proved fruitful, there is still considerable disagreement over when regret can first be observed developmentally (Rafetseder & Perner, 2012; Weisberg & Beck, 2012). It is possible to identify three methodological issues that make it difficult both to resolve this disagreement and to integrate developmental findings with the larger body of research on regret conducted with adults. The current study directly addresses these issues.

All of these developmental studies have used a simple paradigm requiring children to choose between two boxes to win a prize of stickers or candies. Children see the prize from their chosen option and then rate how they feel about that prize on an emotion rating scale ranging from very happy to very sad (the exact nature of this scale differs across studies). When giving these initial ratings, children have received only what is termed *partial* feedback; they have seen the prize resulting from their choice, but they have not yet seen the prize they would have won if they had made a different choice. Children are given *complete* feedback when they see both what they have won and what they would have won if they had made a different choice. Emotion ratings are subsequently made after complete feedback. If the prize from the unchosen option is better than that from the chosen option, children reporting that they now feel sadder are assumed to be experiencing regret about their choice; if it is worse, children reporting that they now feel happier are assumed to be experiencing relief.

Although all of the developmental studies employed this basic procedure, they differ from each other in the exact way that counterfactual emotions are assessed; moreover, the task they use also differs in important ways from the type of task typically used to examine counterfactual emotions in adolescents and adults (Burnett, Bault, Coricelli, & Blakemore, 2010; Camille et al., 2004; Coricelli et al., 2005). We focus on three methodological issues stemming from these differences: the nature of the choices children need to make, the extent to which their emotional responses can be based on a single comparison between the prize received and the best or worst prize available, and the way in which emotion ratings are used to measure regret/relief. Each of these issues is described in turn.

Choice and risky decision making

Experiments examining regret and relief in adolescents and adults have used paradigms in which the choice that participants make is more complex than simply choosing between two boxes. Such studies typically align with research in the broader decision-making literature insofar as the tasks involve choosing between alternatives that vary in associated risk. In most tasks, the choice is between two gambles, with gambles giving participants the opportunity to win or lose points (e.g., Bault, Coricelli, & Rustichini, 2008; Coricelli et al., 2005; Mellers, Schwartz, & Ritov, 1999). For example,

Burnett et al. (2010) asked participants to choose between pairs of gambles that differed in risk such as a choice between a gamble with a 50% chance of winning 50 points along with a 50% chance of losing 50 points (+50/−50) and a gamble with an 80% chance of winning 200 points but a 20% chance of losing 200 points (+200/−200). On a regret version of such a trial, if participants chose the +50/−50 option, they were then shown that they would have won 200 points if they had chosen the other gamble; on a relief version, if participants chose that option, they were then shown that they would have lost 200 points if they had chosen the alternative gamble.

There is a good reason why studies of counterfactual emotions in adults have used tasks in which participants need to choose between options varying in risk: Researchers interested in the processes underlying risky choice have been examining whether counterfactual emotions are an important element of decision making (Coricelli, Dolan, & Sirigu, 2007; Mellers et al., 1999). Such a hypothesis stems from an influential tradition of formal models of the role of regret in economic choice (Bell, 1982; Loomes & Sugden, 1982). Moreover, neuropsychological studies using these tasks have attempted to identify the brain systems that underpin regret (Sommer, Peters, Gläscher, & Büchel, 2009). Studies using such tasks with brain-damaged patients who do not experience regret have helped to establish that this emotion may indeed play a role in risky decision making (Camille et al., 2004; Coricelli et al., 2007; but see also Levens et al., 2014).

Although there are good methodological reasons for using simpler tasks with young children, the fact that the developmental studies on regret use a quite different task that does not involve risky choice means that it is difficult to integrate the findings of such studies with those from the research with adults. In particular, we do not know whether young children experience counterfactual emotions in the same sorts of decision-making situations as adolescents and adults or whether such emotions may play a role in explaining developmental changes in risky decision making. The current study used a task that closely resembled in structure the tasks used with adults insofar as it involved choosing between two gambles that varied in risk. However, the task was simplified. In the studies with older participants, not only does the magnitude of possible prizes vary across gambles, but the odds of winning prizes can vary as well (in the example given above, the odds in one gamble are 50/50 and those in the other are 80/20). By contrast, in our task only the magnitude of the possible prizes varied, with each gamble always involving a 50% chance of winning one of two prizes. For example, children needed to choose between a gamble in which there was a 50% chance of winning 10 points and a 50% chance of winning 7 points and a more risky gamble in which prizes of 16 points and 1 point were equiprobable. Although this simplified the task, it nevertheless allowed us to examine the developmental profile of counterfactual emotions in the context of risky choice.

Single reference point versus multiple reference points

Using a gambling task allowed us for the first time to examine the early development of counterfactual emotions in the context of risky choice. However, it also allowed us to address a further issue regarding the flexibility with which children make spontaneous counterfactual comparisons. One claim in the literature (see Rafetseder & Perner, 2012; Rafetseder & Perner, 2014) is that younger children's emotions in a box-choosing task may be underpinned by just the simple thought that "I do not have the best prize" (or conversely "I do not have the worst prize") and thus are best described as frustration rather than regret. This explanation implies that (a) younger children's emotions are not the result of counterfactual thinking and (ii) the comparisons children make between what they obtained and what else was available are inflexible. That is, if children's emotions are a result of such thoughts, we need only to assume that, unlike adults, children are making a single comparison between what they actually won and the other available prize.

Adults' emotional ratings on the gambling task reflect an ability to flexibly make different types of comparisons between the reward obtained and (at least) two other reference points. By *flexibly*, we mean that adults can shift their point of reference appropriately in response to the different types of information that they receive. Adults' emotional responses indicate that they make a comparison following partial feedback between what they did win on the gamble they selected and what they could have won on that particular gamble. Using the example of a choice between two gambles of +50/−50 and +200/−200, when participants choose the +50/−50 gamble, unsurprisingly they report a negative

emotion if they lose 50 points; negative emotions in such circumstances are interpreted as reflecting disappointment (Mellers et al., 1999). Regret and relief are assumed to be a result of a further comparison when complete feedback is provided (i.e., on being shown the result of the other unchosen gamble) between what they won and what they would have won had they chosen the other gamble. Using our example, a participant who lost 50 points might nevertheless report the positive emotion of relief on seeing that he or she would have lost 200 points had he or she chosen the other gamble.

By using our simplified gambling task, we were able to examine whether children can make similar types of comparisons. Consider a situation in which children choose between a gamble to win either 16 points or 1 point (16/1) or either 10 or 7 points (10/7). Imagine that a child chooses the safer gamble, 10/7, and wins 10 points. That child already knows that he or she has not won the best available prize of 16 points. Nevertheless, the child may report feeling relatively happy because he or she did receive the best prize from the chosen gamble. If the child does report feeling relatively happy, we can be confident that the child's response is not solely based on a single comparison between the actual outcome and the best (or worst) prize available. Now assume that complete feedback is provided, revealing that the alternative prize from the unchosen outcome is 16 points, and the child reports feeling sad. We can be confident under such circumstances that the child has flexibly used two different points of reference: what the child could have won in the chosen gamble and what the child would have won if he or she had selected the alternative gamble. Each of these different points of reference would yield different (and, crucially from the point of view of observing them, contrasting) emotional responses. Although we know that older children, like adults, can indeed use both of these points of reference (Habib et al., 2012), the design of previous studies with younger children has meant that it is not clear at what age such an ability is present.

Within-trial versus between-trial comparisons of emotion rating

The third issue that this study addresses concerns an important difference between existing studies in how emotion ratings are actually used and interpreted. In what we term *within-trial* procedures, participants make ratings following partial feedback, are subsequently given complete feedback, and then make a second rating. This is the method that has been used in the majority of developmental studies. By contrast, in what we term *between-trial* procedures, participants make only a single rating on any given trial either following partial feedback or following complete feedback (e.g., Burnett et al., 2010; Rafetseder & Perner, 2012). These ratings can then be compared across trials as necessary. The majority (although not all) studies with adults have used procedures in which a single emotional rating was collected following either partial or complete feedback (e.g., Burnett et al., 2010; Chua et al., 2009; Coricelli et al., 2005; Mellers et al., 1999; Zalla et al., 2014). As we will see, one possible explanation for the inconsistencies observed in developmental studies lies in whether within-trial or between-trial comparisons of emotion ratings have been used.

Studies examining when regret first emerges developmentally that have used within-trial procedures have all found evidence of regret or relief in children from at least around 6 or 7 years of age (Burns et al., 2012; McCormack & Feeney, 2015; O'Connor et al., 2012, 2014; van Duijvenvoorde et al., 2014; Weisberg & Beck, 2010, 2012); indeed, using such a procedure, Weisberg and Beck (2012) argued that regret may even be observable in children as young as 4 years. Rafetseder and Perner's (2012) study is the only one to examine the developmental emergence of regret using a between-trial comparison of emotion ratings. Notably, they found that in children under 9 years of age, there was no difference in ratings following partial versus complete feedback. On the basis of their findings, they concluded that regret was not apparent in children until around 9 years.

Interestingly, it seems to be the case that whether further age differences are reported in levels of counterfactual emotions later in development also depends on which method is used. In a study comparing young adolescents, mid-adolescents, and adults using the gamble choice paradigm and a between-trial method, Burnett et al. (2010) found that there was only limited evidence of developmental changes. There was no evidence of developmental increases in the intensity of negative emotion in regret trials, and although they did find increases in the intensity of positive emotions in relief trials between young and mid-adolescents, there were no further changes after this age. By contrast, using a similar paradigm but with a within-trial procedure, Habib et al. (2012) found

evidence of developmental increases in regret, with both 11-year-olds and an adolescent group showing less regret than an adult group. These studies differed in that Burnett et al. (2010) index of regret and relief was simply the intensity of reported emotion on complete feedback trials, whereas Habib et al. (2012) calculated the difference between emotion ratings involving partial and complete feedback taken within a single trial. Habib et al. speculated that the results of the two studies are not consistent because of this difference in the way that regret and relief were measured.

Looking across the findings from developmental studies, it is clear that whether it is better to use a within-trial or between-trial procedure is an important methodological issue. Researchers working with younger children disagree as to which is the most appropriate method (McCormack & Feeney, 2015; Rafetseder & Perner, 2012). Rafetseder and Perner (2012) argued that it is not appropriate to use a within-trial method because children may feel compelled to change their response if asked to give a second emotion rating in a single trial; children's reports of feeling sadder after complete feedback may simply reflect a tendency to give a different rating after double questioning. Rafetseder and Perner (2012) found some evidence to support this suggestion in a study where they directly compared a within-trial method with a between-trial method. Younger children were more likely to report feeling sadder after seeing the alternative outcome in the within-trial condition, where they made two emotion ratings, than in the between-trial condition, where they made only one emotion rating. However, O'Connor et al. (2012, 2014; see also McCormack & Feeney, 2015; van Duijvenvoorde et al., 2014) included an extra trial type to control for the effects of double questioning in which the prizes in both the chosen and unchosen boxes were identical. They found that very few children were likely to report feeling sadder in these control trials, suggesting that children's tendency to report feeling sadder in regret trials was not an artifact of double questioning. Nevertheless, given the burgeoning number of studies on the development of counterfactual emotions, it would be useful to reexamine this methodological issue directly. The current study aimed to do so by using within-trial and between-trial comparisons of emotion ratings in the same study to directly examine whether these methods do indeed yield different patterns of developmental findings.

The current study

The current study used our simplified version of the gambling task with children aged 6 or 7 years and an older group of 8- and 9-year-olds. Our youngest age group was selected because the majority of researchers have concluded that most 6- and 7-year-olds are able to experience regret, as measured by the box-choosing task. However, it is not clear whether this age group will show evidence of counterfactual emotions in the context of risky decision making. We included an older group because we were also interested in whether there would be developmental increases across this age range in the likelihood that participants experienced counterfactual emotions. Van Duijvenvoorde et al. (2014) demonstrated that the likelihood that children experienced counterfactual emotions even in a simple box-choosing task increased over this period of childhood, and this study allowed us to examine whether this was also the case in a risky decision-making task.

Participants selected between gambles varying in risk (see Table 1), with a fixed 50% chance of each outcome. Children received many fewer trials than in studies with adolescents and adults, but we constructed these trials carefully to allow an examination of whether children were able to use two points of reference in their emotion judgments. There were two general types of trials: regret trials and relief trials. In regret trials, children selected between two gambles such as 10/7 and 16/1, where 10/7 is the safer option compared with the riskier 16/1 option. In Experiment 1, if children chose the safer option in regret trials, they always won the best prize from that option (10 points). However, they were also always shown that they could have won a better prize (16 points) if they had chosen the riskier option. In relief trials, children also needed to make a choice between a safer option and a riskier option such as between 9/6 and 14/1. If they chose the safer gamble, they always won the worse prize (6 points) but were shown that they would have won fewer points if they had chosen the risky gamble (1 point). Experiment 2 used a similar set of trials but, as will be described, also included additional regret and relief trials in which the choice of gambles and magnitude of the received prize were matched between these trial types. In Experiment 2, we also varied whether children initially won the best or worst prize from their chosen gamble.

Table 1
Number of points available in each box for each trial number along with the outcomes of the safe and risky gambles when the safe option was chosen.

Trial number	Safe gamble	Risky gamble	Outcomes when safe choice made		Type
			Actual	Alternative	
1	10 or 7	16 or 1	10	16	Regret
2	11 or 7	17 or 1	11	17	Regret
3	9 or 6	14 or 1	6	1	Relief
4	8 or 4	11 or 1	4	1	Relief
1a	10 or 7	16 or 1	10	1	Relief
2a	11 or 7	17 or 1	11	1	Relief
3a	9 or 6	14 or 1	6	14	Regret
4a	8 or 4	11 or 1	4	11	Regret

Note: Experiment 1 used Trials 1 to 4 only, whereas Experiment 2 used all 8 trial numbers.

In Experiment 1, in addition to varying whether trials were relief or regret trials, we also varied whether participants made two emotion ratings per trial (one after partial feedback and a second one after complete feedback) or just one emotion rating (only after complete feedback). This allowed us to directly compare regret and relief as measured either using the within-trial method or the between-trial method. Of interest was whether the developmental patterns differed depending on the measure used and whether either measure was more sensitive.

Experiment 1

Method

Participants

Participants were 37 6- and 7-year-olds (18 girls; M = 84.4 months, range = 77–91) and 31 8- and 9-year-olds (13 girls; M = 110.8 months, range = 103–118). Children were recruited from schools local to the university of the first author, were predominately from lower- to middle-class backgrounds, and were of Caucasian origin.

Apparatus

The task used 12 cardboard boxes (16 × 26 × 15 cm). The 4 boxes used in the practice trials were painted blue and red, whereas the 8 boxes used in the experimental trials were painted black and white. The boxes were vertically divided into two colored sections, and each section contained a card representing the number of points won in two ways: as a printed symbolic number and in concrete form as a picture of a stack of discs, with the number of discs corresponding to the number of points won (see Figs. 1 and 2). Each box had a hinged lid that could be opened to reveal just one side of the box. Along the division on the outside of each box were two yellow stars indicating the number of points in the two sections displayed as symbolic numbers. There was no indication of which colored side contained which amount of points. Separately, there were also concrete representations of the possible points to be won in the form of cylindrical stacks of metal discs, with each disc representing 1 point. These stacks were placed to the side of each box (see Figs. 1 and 2). Two dice were used; the practice die was painted half blue and half red, and the experimental die was painted half black and half white. During the task, a horizontal number line from 1 to 100 with a small arrow was used to indicate the number of points accumulated during the game.

A 7-point affective response scale was used, with cartoon faces varying in emotional expression from very happy on the left-hand side to very sad on the right-hand side of the scale. Children indicated their emotional responses using this scale and a three-pronged arrow. The three-pronged arrow had leftward-, rightward-, and upward-pointing prongs indicating feeling happier, sadder, and the same, respectively. Seven small pictures representing different emotion-provoking scenarios

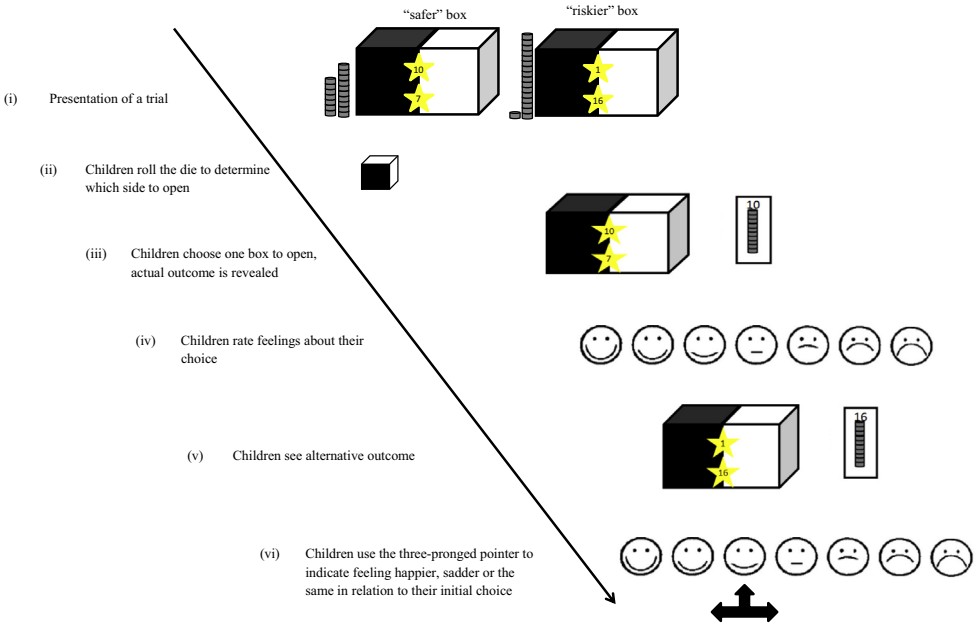

Fig. 1. Experimental design of a Two Rating trial. (i) A set of boxes was presented. (ii) A die was rolled to determine which side of the chosen box would be opened. (iii) Children saw the actual outcome. (iv) Children rated their feelings about their choice on the 7-point scale. (v) The unchosen box was opened, and the alternative prize was revealed. (vi) Children used the three-pronged pointer to indicate whether they now felt happier, sadder or the same in relation to their initial choice.

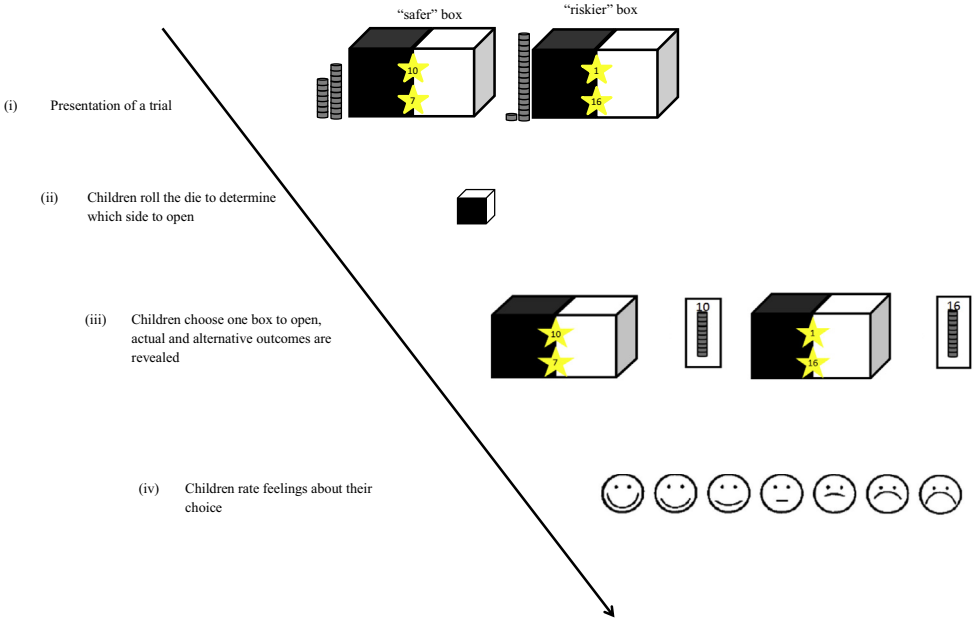

Fig. 2. Experimental design of a One Rating trial. (i) A set of boxes was presented. (ii) A die was rolled to determine which side of the chosen box would be opened. (iii) Children saw both the actual and alternative outcomes. (iv) Children rated their feelings about their choice on the 7-point scale.

(e.g., a lost dog poster, a trophy) and two puppets with toys (a camera, a mobile phone, and a watch) were used in the pretraining session where children learned how to use the emotion scale. Prizes were small "goodie" bags containing age-appropriate toys.

Procedure

Children were invited to play a game and were told that if they won enough points by the end of the game, they would receive a goodie bag. The 7-point scale was first introduced along with the seven cartoon pictures. The cartoon faces were described from left to right as feeling "really really happy" through to "really really sad." Children were then asked to indicate which face represented one of the aforementioned feelings. Next, children were shown the seven cartoon pictures and were told that different things can make us feel happy or sad. The pictures and the scenarios they represented were discussed before children were asked to rate how they would feel in each scenario (e.g., "Imagine you won a trophy; how would you feel?"). Each picture represented a different feeling on the scale. Answers were discussed and corrected if necessary. This technique was used so that children were comfortable with the idea of degrees of happiness or sadness and thus willing to use the full range of the scale.

Children were then introduced to the two puppets. One puppet received a small toy and was described as feeling "really happy," and the other puppet received two toys and was described as feeling "really really happy," with the arrow pointing to the appropriate face on the scale. Each puppet received an additional toy, and children were asked to indicate, using the three-pronged arrow, whether that puppet now felt happier than before (leftward prong), sadder than before (rightward prong), or the same as before (upward prong). This demonstration was repeated but with both puppets losing their toys. Children who gave incorrect answers were corrected, and the trials were repeated. Full details of this training procedure are provided in O'Connor et al. (2012, Experiment 2).

Next, a pair of blue and red practice boxes was introduced. Each of the two boxes was split into two sides, one red and one blue, and the boxes were opened to show children the two sections. The experimenter showed the two cards depicting the prizes for each box. In the first practice trial, the two cards for one of the boxes depicted 9 and 4 points, respectively, and the two cards for the other box depicted 8 and 5 points, respectively. The experimenter then placed one card in each colored section of each box, keeping the number on each card hidden. The experimenter pointed to the two stars on the outside of each box that depicted its two prizes (e.g., 9 and 4 points) along with the concrete representations of these in the form of the cylindrical stacks of metal discs that were placed in the front of each box. The experimenter made it clear that there was only one card per section and that children could not be sure which card was in which section. Children were then asked a series of comprehension questions to ensure that they understood how the points were allocated. Children who gave incorrect answers were corrected, and the process was repeated until they showed a full understanding of how the points were distributed.

Children were told that to play the game they first needed to roll the colored die because this would determine whether they would be using the blue or red sections of the boxes.[1] After the die was rolled, children were asked to choose just one box for a prize of points. The experimenter opened the predetermined colored section (blue or red) of the chosen box and revealed the printed card (e.g., 9 points). The corresponding section of the unchosen box was then opened, and the card was revealed (e.g., 8 points). Children completed one more practice trial with the two remaining blue and red boxes; in this second

[1] In the practice and test trials, the die was rolled in advance of children's choice of box and determined in advance which side of the box children would get. Children themselves chose which box they wanted (i.e., which gamble). The roll of the die essentially determined what the outcome of each gamble would be. We decided that these outcomes should be fixed in advance of children's choice of box because we wanted to ensure that children could be confident that they would have obtained the specific counterfactual prize available in the other box had they chosen differently. If the die was not rolled until after children had made their choice, there could have been some ambiguity over what the nature of the counterfactual prize would have been had they chosen differently because that prize would not have actually been decided until after they had made their choice. It was essential that children could be sure that had they chosen a different gamble, they would have obtained a prize of a particular magnitude, and if the die had not been thrown until after they had made their choice, they might not have been confident that this was the case.

trial, one box contained 7 and 3 points and the other box contained 6 and 4 points. The presentation order of the practice trials was counterbalanced.

Children then moved on to experimental trials that involved four sets of black and white boxes. Each experimental trial included a safer box and a riskier box; see Table 1 (Trials 1–4) for a breakdown of the number of points displayed on the outside of each box. Unknown to children, trial outcomes were manipulated in advance of the game by planting appropriate cards in the sections of the boxes; actual outcomes are depicted in Table 1 for each trial type, assuming that children made the safer choice. Children who chose the riskier box always won just 1 point because it was hoped that this would deter children from choosing the risky boxes. This decision was made because the critical trials for analysis were trials in which children chose the safer box.[2]

The experimental trials were presented in blocks of four, with each trial number (Trials 1–4 in Table 1) occurring once per block. For each child, the order of the trials remained constant across blocks but was varied between children. We did this by assigning children randomly to one of 12 different possible trial orders, with 6 of these orders having a regret trial as the first one in the block and 6 having a relief trial as the first one in the block. Because we were only interested in children's performance on trials in which children made the safer choice (see also Burnett et al., 2010; Habib et al., 2012), the number of blocks and trials that children received varied depending on the number of safer choices they made; we detail below how this was managed.

Two different methods were used to measure children's feelings about their choices during the game. The *within-trial* method involved trials in which children made two emotional ratings, which had the following structure depicted stage by stage in Fig. 1. In Stage (i), children were shown the pair of black and white boxes with the possible points depicted on them and the corresponding representations of the number of points in stacked discs. In Stage (ii), the black and white die was thrown, determining whether children would get the black half or the white half of the boxes. In Stage (iii), children were asked to select one box from the pair, and depending on the die throw the experimenter opened either the black or white section of the chosen box and children were shown the number of points that they had obtained (in the example in Fig. 1, the child chooses the safer box, and the appropriate side is opened to reveal a card depicting 10 points; the stack of 10 discs is then placed beside the box, and the stack of 7 discs is removed to make it clear that the child has won 10 points). In Stage (iv), children were asked to rate how they felt about the outcome using the 7-point scale (emotion rating after partial feedback). In Stage (v), once this rating had been made, the alternative outcome was then revealed from the appropriate section of the unchosen box (in the example in Fig. 1, the child finds out that he or she would have won 16 points had he or she chosen the other box). In Stage (vi), children were asked to indicate, using the three-pronged pointer, whether they now felt happier, sadder, or the same (updated emotion rating after complete feedback). Thus, for these trials, children made two emotion ratings: one following partial feedback and one following complete feedback; these are labeled Two Rating trials.

For the between-trial method, the measure of interest was whether children's emotion ratings differed depending on whether they had received complete or partial feedback. This measure necessarily involved a comparison of emotion ratings from two separate trials of the same type; a rating after partial feedback was compared with a rating given on a separate but otherwise identical trial in which only complete feedback was provided. Ratings following partial feedback were taken from the Two Rating trials [i.e., at Stage (iv) listed above and as depicted in Fig. 1]. These were compared with ratings taken from a separate block of otherwise identical trials that had the same Stages (i) to (iii) as above but involved children making an emotion rating only once they had complete feedback. Fig. 2 shows a sample One Rating trial. As can be seen from the figure, these differed from Two Rating

[2] If children chose the riskier box, by default they must always have received either the best or worst prize available (e.g., if the choices were 10/7 and 16/1 and children chose the 16/1 box, they must always have received either the best possible prize of 16 or the worst possible prize of 1). This meant that they could learn nothing new about whether the prize in the unchosen box was better or worse when it was opened (i.e., even before the corresponding side of the other box was opened, they would know that it must contain either a worse prize than 16 or a better prize than 1). Because of this, the distinction between partial feedback and complete feedback does not exist for the riskier choices. Thus, we did not analyze emotional responses when children made riskier choices (this was also the case in the two previous developmental studies using a gambling task with adolescents; see Burnett et al., 2010; Habib et al., 2012).

trials in that children saw the contents of both the chosen and unchosen boxes before making a single emotion rating.

Children first completed two blocks of 4 trials, with each block consisting either of 4 One Rating trials or 4 Two Rating trials. Each block involved the 4 trials depicted in Table 1. Once these blocks had been administered, because only data from trials where the safer option had been chosen could be included in the analysis, we administered a single repetition of trials in which children had made the riskier choice. Testing proceeded after the administration of the first two blocks of trials as follows. If children had chosen the safer option in both the One Rating and Two Rating blocks for any specific trial number (from Trials 1–4 in Table 1), they were no longer given any more trials of that number because this provided the full data set needed for an assessment of performance on that trial number using both the within-trial and between-trial methods. This meant that if children had chosen only safer options on all 8 trials in the first two blocks (which was rare), testing was terminated. However, if children had made a riskier choice for a specific trial number in either the One Rating or Two Rating block, that specific trial number was repeated once as necessary in a subsequent block in order to obtain the data needed for that trial. For example, with regard to Trial 1 (10/7 vs. 16/1), if children chose the safer gamble (10/7) when given both One Rating and Two Rating versions of that trial, they did not repeat Trial 1. However, if they chose the riskier option (16/1) in the One Rating version of the trial, the One Rating version was repeated in a subsequent One Rating block (the same was true if they chose the riskier option in the Two Rating version of that trial). Thus, repetitions of the One Rating and Two Rating blocks included only trials for which children had not chosen the safer option. If children failed to choose the safer option again on a given trial number, we did not administer it a further time, and we set an upper limit of four blocks of trials per child to prevent fatigue with the task. This meant that children received variable quantities of trials, with a minimum of 8 and a maximum of 16 trials ($M = 12.06$, $SD = 1.97$).

Children were shown a running total of the number of points they had won. The experimenter used the 1 to 100 number line to do this, recording the number of points children won during the game by updating the position of the arrow on each trial. All children received a goodie bag at the end of the game regardless of the actual number of points won.

Results

The younger children completed on average 1 more trial than the older children ($M = 12.51$, $SD = 2.01$ vs. $M = 11.52$, $SD = 1.81$); this was due to the younger children making significantly more risky choices, $t(66) = 2.22$, $p < .05$. All subsequent analyses reported here are only of trials in which children chose the safer box.

We began by comparing initial emotional ratings following partial feedback only from blocks of Two Rating trials to examine whether children's responses to their actual prize varied depending on whether they had received the best possible prize from their chosen box. Ascribing a score of 1 = *really really sad* to 7 = *really really happy,* the average first emotion rating on the 7-point scale given in regret trials was 6.47 ($SD = 0.92$) and in relief trials was 5.42 ($SD = 1.08$); children were initially significantly happier after seeing the actual outcome in the regret trial (where they received the best outcome from their chosen gamble) than they were in the relief trial (where they received the worst outcome from their chosen gamble), $t(57) = 6.51$, $p < .001$.

Within-trial analysis of counterfactual emotions

This analysis focused on the second emotional ratings given after complete feedback in Two Rating trials using the three-pronged arrow. These ratings are always relative to the first ratings following only partial feedback and are categorical (i.e., happier than, sadder than, or the same as before the unchosen box was opened, depending on which prong of the arrow was chosen). Not all children chose the safer option for at least one Two Rating regret trial and one Two Rating relief trial, meaning that there were a small number of children who could not be included in the analyses; the analysis includes only the 31 6- and 7-year-olds and 27 8- and 9-year-olds who chose the safer option on at least one regret trial and one relief trial in a Two Rating block. The proportions of categorical responses (happier, sadder, or the same) given by each age group in regret and relief trials are shown in Fig. 3. A

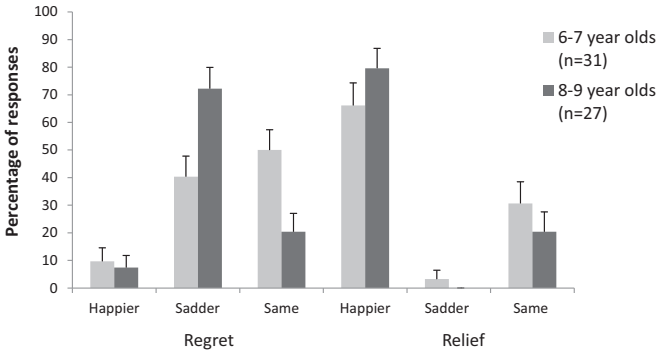

Fig. 3. Proportions of times children reported feeling happier, sadder, or the same on regret and relief trials in Experiment 1 as a function of age using within-trial measurement.

two-way analysis of variance (ANOVA) with a between-participants factor of age and a within-participants factor of trial type (regret or relief) was conducted on the proportion of sadder responses. There was a significant effect of trial type, $F(1, 56) = 86.61$, $p < .001$, $\eta_p^2 = .61$, with more sadder responses in the regret trials, and a significant interaction between age and trial type, $F(1, 56) = 8.94$, $p < .01$, $\eta_p^2 = .14$. Further analyses (making a Bonferroni correction for 4 tests, $\alpha = .0125$) showed that the effect of trial type was significant for each of the two age groups: 6- and 7-year-olds, $t(30) = 4.28$, $p < .01$; 8- and 9-year-olds, $t(26) = 9.37$, $p < .01$. However, the proportion of sadder responses in regret trials increased with age, $t(56) = -2.96$, $p < .01$. A further two-way ANOVA with a between-participants factor of age and a within-participants factor of trial type was conducted on the proportion of happier responses. The main effect of trial type was significant, $F(1, 56) = 98.84$, $p < .001$, $\eta_p^2 = .64$, with more happier responses in relief trials than in regret trials, but there were no other significant effects.

These analyses demonstrate that children's emotional responses varied appropriately between regret and relief trials and that there was an increase with age in the likelihood that children reported feeling sadder in regret trials. We also examined whether children were more likely than chance to report feeling sadder in regret trials and happier in relief trials. Because there were three possible choices (sadder, happier, or the same), we compared whether proportions of sadder responses in regret trials and happier responses in relief trials differed significantly from 0.33 using a one-sample *t*-test. These analyses showed that the older group gave significantly more sadder responses in regret trials and happier responses in relief trials than would be expected by chance, $t(26) = 5.09$, $p < .001$ and $t(26) = 6.49$, $p < .001$, respectively; however, although the 6- and 7-year-olds gave more happier responses in relief trials than would be expected by chance, $t(30) = 4.06$, $p < .001$, the proportion of sadder responses in regret trials did not differ significantly from chance, $t(30) = 0.98$, $p = .34$. Note, however, that this group did not respond randomly in regret trials (see Fig. 3); the younger children produced very few happier responses (unlike in relief trials), with most of their responses being either sadder or the same. Taken together, these findings suggest unambiguously that the older children experienced both regret and relief. However, whereas the younger group experienced relief, the younger children did not consistently experience regret.

Between-trial analysis of counterfactual emotions

In this analysis, children's initial emotional ratings following partial feedback for a specific trial type [taken from Stage (iii) in Two Rating trials] were compared with their ratings from an otherwise identical trial type in which they made a rating only following complete feedback (taken from One Rating trials). To generate categorical data analogous to those used in the within-trial analysis, we examined the proportion of trials in which the latter ratings were the same as, sadder than, or happier than the former ratings (see Fig. 4). To make the required comparisons for analysis, it was necessary to have data from two trials of the same type in which children made the safer choice (one where

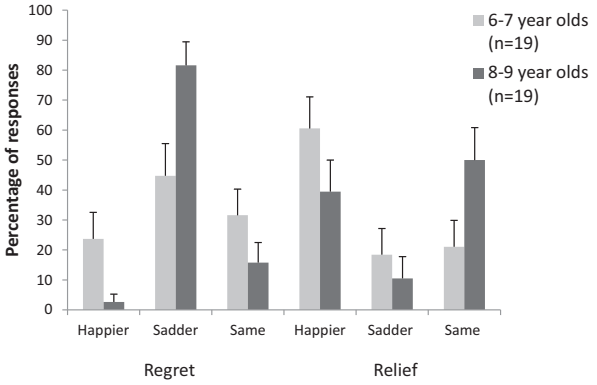

Fig. 4. Proportions of times children reported feeling happier, sadder, or the same on regret and relief trials in Experiment 1 as a function of age using between-trial measurement.

children made an emotional rating following partial feedback and one where they made a rating only after full feedback) for at least one regret and one relief trial. There were 19 children from each age group who made a sufficient number of safer choices to generate these data. A two-way ANOVA with a between-participants factor of age and a within-participants factor of trial type was conducted on the proportion of sadder responses. The main effect of trial type was significant, $F(1, 36) = 36.67$, $p < .001$, $\eta_p^2 = .51$, as was the interaction between trial type and age group, $F(1, 36) = 7.74$, $p < .01$, $\eta_p^2 = .18$. Post hoc *t*-tests (making a Bonferroni correction for 4 tests, $\alpha = .0125$) showed that the effect of trial type was marginally significant for the 6- and 7-year-olds, $t(18) = 2.04$, $p = .06$, and significant for the 8- and 9-year-olds, $t(18) = 7.39$, $p < .01$. The effect of age was significant for the proportion of regret trials in which children made a sadder rating, $t(36) = -2.77$, $p < .01$, with older children being more likely to report feeling sadder on these trials. A two-way ANOVA with a between-participants factor of age and a within-participants factor of trial type was conducted on the proportion of happier responses. The main effect of trial type was significant, $F(1, 36) = 17.95$, $p < .001$, $\eta_p^2 = .33$, as was the main effect of age, $F(1, 36) = 5.69$, $p < .03$, $\eta_p^2 = .14$, with younger children being more likely to report feeling happier than older children across both trial types. However, the interaction between age and trial type was not significant, $F < 1$. In summary, although these analyses are based on an entirely different set of emotional ratings, the findings are very similar to those reported for the within-trial analysis in suggesting that both age groups experienced regret and relief and an increased likelihood of children experiencing regret with age.

Additional analyses

To check whether the within-trial and between-trial measures differed from each other in terms of sensitivity, for each age group we used paired-sample *t*-tests to compare the proportion of sadder responses in regret trials yielded by each measure as well as the proportion of happier responses in relief trials yielded by each measure. These analyses showed that none of these proportions differed from each other (all *p*s > .30) except for the difference between the proportions of happier responses in relief trials for the older children, for whom the within-trial measure yielded a significantly higher proportion of happier responses, $t(18) = 2.36$, $p < .05$.

The experimental procedure involved children completing different numbers of trials, meaning that children had quite variable exposure to different sets of outcomes, which could have affected their responses as they moved through the task. Therefore, we conducted a final set of analyses in which we focused just on the very first Two Rating regret trial and Two Rating relief trial in which children made a safer choice. In the first regret trial on which children chose the safer box, 19 of 33 (57.6%) 6- and 7-year-olds and 23 of 31 (74.2%) 8- and 9-year-olds reported feeling sadder following complete feedback. In both of these groups, the probability of children reporting that they now felt

sadder was significantly greater than chance on a binomial test, both *p*s < .01, assuming that one third of responses will be sadder by chance. In the first relief trial where children chose the safer box, 20 of 32 (62.5%) 6- and 7-year-olds and 23 of 27 (85.2%) 8- and 9-year-olds reported feeling happier following complete feedback, in both cases a proportion greater than chance, binomial test, both *p*s < .01.

Discussion

Children's initial emotion ratings after partial feedback demonstrated that they felt sadder if they had received the worse prize from their chosen box than if they had received the better prize. However, children of both ages were able to flexibly update this initial emotional response once they had seen the prize in the unchosen box. Analysis of the within-trial data indicated that children were then likely to subsequently report feeling sadder if the prize in the unchosen box was better (regret trials) but happier if the prize in the unchosen box was worse (relief trials). However, although the younger children did not respond randomly in regret trials (they very rarely gave happier responses, unlike in relief trials), they did not report feeling sadder significantly more often than chance (many of the children reported feeling the same). Thus, although this group did vary their emotional responses across trial types, they did not all consistently experience regret. The likelihood that children experienced regret increased developmentally, with older children more reliably reporting feeling sadder in regret trials.

The between-trial comparisons of emotional ratings given in separate trials under partial or complete feedback also revealed that children's emotional responses in both regret and relief trials was sensitive to whether or not children had received full feedback. However, for this measurement the proportion of sadder responses in regret trials was only marginally significantly different from that in relief trials for the younger group, and the likelihood that children would report feeling sadder in regret trials also increased with age. Thus, the pattern of findings is generally consistent across both ways of measuring regret and relief. We note, however, that the within-trial method may be somewhat more sensitive, with the proportion of older children reporting feeling happier in relief trials being larger under this measure than under the between-trial measure. The within-trial method also has the advantage of requiring administration of half the number of trials as the between-trial method, which is very advantageous when testing young children.

Experiment 2

The findings of Experiment 1 were primarily based on a comparison between the proportions of sadder/happier responses on regret trials and those on relief trials. However, these trials differed not just in terms of whether the prize in the unchosen box was better or worse than the one in the chosen box; the choices that children needed to make on regret trials were not completely identical to those that they needed to make on relief trials (cf. Trials 1 and 2 with Trials 3 and 4). Moreover, the actual prize that children received from their chosen gamble on regret trials (10 or 11 points) was always better than the actual prize that children received on relief trials (6 or 4 points). This latter aspect of the design was deliberate to provide us with a test of whether children could update their emotion appropriately on seeing the prize in the unchosen box. However, it might be argued that these differences between regret and relief trials mean that it is difficult to straightforwardly compare patterns of responses across these two trial types.

In our second experiment, we used a similar task as in Experiment 1 but made three key changes. First, we added an additional 4 trials to ensure that regret and relief trials were matched both in terms of the choices children needed to make in each trial type and in terms of the size of the actual prize obtained. These are listed in Table 1 as Trials 1a to 4a. Second, because the results of Experiment 1 had indicated that there was no benefit in also including a between-trial measurement of regret/relief, we used only Two Rating trials, yielding only a within-trial measurement of regret/relief. Third, all children received the same number of trials to ensure that they were all administered an identical task.

Method

Participants

Participants were 38 6- and 7-year-olds (19 girls; M = 79.6 months, range = 72–94) and 36 8- and 9-year-olds (17 girls; M = 107.8 months, range = 96–119 months). All children were recruited from the same population as in Experiment 1.

Apparatus

The stimuli used were the same as those used in Experiment 1.

Procedure

The procedure was very similar to that of Experiment 1, with the addition of Trials 1a to 4a in Table 1. The experimental trials were presented in blocks of 4 Two-Rating trials, with each block consisting of 2 regret trials and 2 relief trials. Block 1 consisted of Trials 1, 3, 2a, and 4a, whereas Block 2 consisted of Trials 1a, 3a, 2, and 4. All children received the same number of trials in this task (the two blocks of 4 trials shown in Table 1). Children were randomly assigned to one of 24 possible trial orders, with half receiving a regret trial first and half receiving a relief trial first.

Results and discussion

The 6- and 7-year-olds tended to choose the riskier box (M = 4.82, SD = 1.27) significantly more often than the 8- and 9-year-olds (M = 3.31, SD = 1.33), $t(72) = 5.00$, $p < .001$. As in Experiment 1, only data from trials in which children chose the safer box were analyzed. We first examined whether children's initial emotion ratings following partial feedback differed depending on whether they had received the best prize (Trials 1, 1a, 2, and 2a) or the worst prize (Trials 3, 3a, 4, and 4a) in their chosen box. The average first emotion rating given in trials where children won the best possible prize from their chosen box was 6.9 (SD = 0.27) compared with 5.7 (SD = 1.01) in trials where children won the worst possible prize; children were initially significantly happier in the former trials, $t(68) = 9.13$, $p < .001$.

Subsequent analyses focused on the proportion of times children gave sadder, happier, or the same responses once they had been given complete feedback. To be included in these analyses, children needed to make the safer choice in at least one regret trial and one relief trial; here, 29 of the 6- and 7-year-olds and 34 of the 8- and 9-year-olds were included. Fig. 5 shows the proportion of sadder, happier, and same responses for each trial type. A two-way ANOVA with a between-participants factor of age and a within-participants factor of trial type was conducted on the proportion of times children reported feeling sadder after complete feedback. There was a main effect of trial type, with more sadder responses given in regret trials than in relief trials, $F(1, 61) = 231.02$, $p < .001$, $\eta_p^2 = .79$. No other effects were significant. Because the results of Experiment 1 indicated that the proportion of sadder responses given by older children was significantly greater than that given by younger children, planned comparisons examined this age effect. However, although the proportion of sadder responses in regret trials was larger for the older group, the age effect was not significant, $t(61) = -1.12$, $p = .27$. A further two-way ANOVA with a between-participants factor of age and a within-participants factor of trial type was conducted on the proportion of times children reported feeling happier after complete feedback. There was a main effect of trial type, with more happier responses being given in relief trials than in regret trials, $F(1, 61) = 398.06$, $p < .001$, $\eta_p^2 = .87$, and no other effects were significant.

Although all children completed the same number of trials in this experiment, the number of trials used in the analyses to yield the data depicted in Fig. 5 differed between children because children differed in terms of the number of times they chose the riskier box (M = 4.08, SD = 1.50, range = 0–8). This might be seen as problematic because children with a tendency to make riskier choices will produce fewer data points but also potentially may be less likely to experience regret or relief. To check this, we examined initially whether there were correlations between the number of times children made a riskier choice and the proportion of sadder responses in regret trials and between the number of times children made a riskier choice and the proportion of happier responses in relief

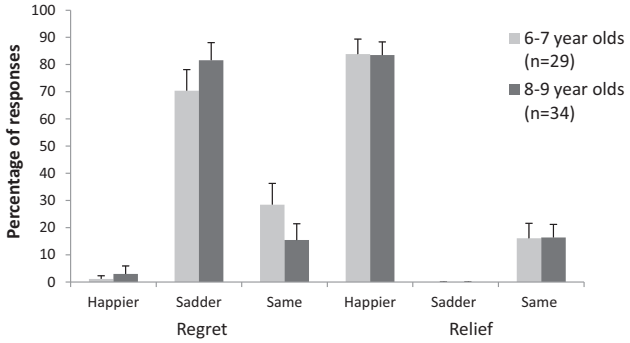

Fig. 5. Proportions of times children reported feeling happier, sadder, or the same on regret and relief trials in Experiment 2 as a function of age.

trials. Neither of these correlations approached significance (both *ps* > .17). We also reran the ANOVAs described above including number of riskier choices as a covariate and found a qualitatively identical pattern of findings, with only the main effect of trial type being significant in both analyses. These analyses indicate that whether or not children experienced regret or relief was not confounded by the number of riskier choices they made.

We conducted a final set of analyses on the very first regret trial and relief trial in which children chose the safe option. In their first completed regret trial, 22 of 35 (62.9%) 6- and 7-year-olds and 29 of 35 (82.9%) 8- and 9-year-olds reported feeling sadder. In both of these groups, the probability of children reporting that they now felt sadder was significantly greater than chance on a binomial test, both *ps* < .01. In their first completed relief trial, 25 of 31 (80.6%) 6- and 7-year-olds and 32 of 35 (91.4%) 8- and 9-year-olds reported feeling happier, in both cases a higher proportion than would be expected by chance on a binomial test, both *ps* < .01. These findings are very similar to those yielded by the analysis of the full set of trials.

In summary, as in Experiment 1, children were initially sadder (following only partial feedback) if they had received the worst prize from their chosen box. This suggests that children felt disappointed if their hope to win the best prize from that box was not realized. In this experiment, we found that children of both ages showed regret on seeing that the prize they would have won in the unchosen box was better and showed relief if the alternative prize was worse. In Experiment 1, we had found somewhat ambiguous evidence that the younger group experienced regret; although the youngest children were more likely to report feeling sadder on regret trials than on relief trials, this difference was only marginally significant using the between-trial measurement, and using the within-trial measurement these children did not differ from chance in the proportion of times they reported feeling sadder in the former trials. The best interpretation of those findings was that only some of the younger children experienced regret. However, in Experiment 2 even the younger group reported feeling sadder on regret trials 70% of the time. Older children were more likely to feel regret than younger children (in 81% vs. 70% of regret trials), although unlike in Experiment 1 this difference did not reach significance.

Why might the younger group be more likely to show regret in Experiment 2 than in Experiment 1? This is a different group of children, and it is possible that individual differences in ability might explain the contrasting patterns of results (although there are mixed findings over whether children's ability, as assessed using standardized measures, predicts reported regret/relief; see Burns et al., 2012; McCormack & Feeney, 2015; O'Connor et al., 2014). There are also notable methodological differences between the experiments that may explain the differing findings. Experiment 2 included a different set of trials than Experiment 1, and in the new trials the magnitude of the difference between the obtained reward and the counterfactual alternative was larger than in the original set of trials (see Table 1; compare Trials 1 and 2 with Trials 3a and 4a). However, this does not seem to explain the difference. We compared the percentage of times the younger children reported feeling sadder

in Trials 1 and 2 with the percentage of times they reported feeling sadder in Trials 3a and 4a and found no difference; the younger children were sadder 70% of the time in the former trials compared with 72% in the latter trials. Our best guess is that performance was somewhat better in the second experiment in the younger group because the experimental procedure was less complex. Children made the same pair of emotion ratings on every trial, whereas in Experiment 1 they completed both One Rating and Two Rating trials. Moreover, children always completed the same 8 trials in Experiment 2, but the vast majority of children completed more than this number of trials in Experiment 1 because of the need to repeat some trials. We note that our analysis of the first completed regret trial in Experiment 1 did show that as a group the younger children reported feeling sadder more often than would be expected by chance, whereas this was not the case for the full trial set. Thus, we suspect that our younger children may have found the larger set of trials in Experiment 1 to be onerous.

General discussion

This study was the first to examine regret and relief in younger children using the type of risky decision-making task used to study these counterfactual emotions in adolescents and adults. Taken together, the findings of both experiments suggest that children as young as 6 or 7 years do experience both regret and relief in this sort of task, although the evidence for regret in this age group was stronger in the second experiment. These findings are consistent with developmental findings from other studies in which children make a simpler choice between two colored boxes. The majority of these previous studies have also found that regret and relief are present from at least 6 or 7 years of age (Burns et al., 2012; McCormack & Feeney, 2015; O'Connor et al., 2012, 2014; Weisberg & Beck, 2010, 2012). However, they are not consistent with the findings of Rafetseder and Perner's (2012) study, which failed to find evidence of regret in children until 9 years of age. We discuss our findings in relation to the three methodological issues raised in the Introduction.

Choice and risky decision making

This study is the first to use a version of the gambling task with children as young as 6 or 7 years to examine counterfactual emotions. Its findings indicate that, like adolescents and adults, children of this age do experience both regret and relief in circumstances that involve risky choice. The findings pave the way for using the gambling task with a broader age range of participants to examine whether there are developmental changes between childhood and adulthood in the likelihood of experiencing these emotions. One of the reasons why there has been intense interest in counterfactual emotions in the gambling task is that researchers have been trying to establish the roles of these emotions in decision making. Thus, it has been argued that regret and relief may have an impact on how people make risky decisions and that decision making will be different in the absence of such emotions (e.g., due to brain damage) (Camille et al., 2004; Coricelli et al., 2007). This claim is consistent with long-standing suggestions in both economics (Bell, 1982; Loomes & Sugden, 1982) and psychology (Zeelenberg, 1999; Zeelenberg, Beattie, van der Pligt, & de Vries, 1996; Zeelenberg & Pieters, 2007) that individuals attempt to minimize future regret when making decisions. Thus, one important reason for examining the development of regret within the context of risky decision making would be to explore whether developmental changes in regret are accompanied by or indeed explain changes in risky decision making.

Single reference point versus multiple reference points

Our findings suggest that at least by the time children are 6 or 7 years old, their emotions are sensitive to two different reference points. When given partial feedback, children were less happy if they had won the worst of two prizes in their chosen gamble than if they had won the best of two prizes. This can be interpreted as disappointment (Mellers et al., 1999; Zeelenberg, van Dijk, & Manstead, 2000) and suggests that children hoped to win the better prize. As we have discussed, children's emotions were also sensitive to the outcome of the unchosen gamble, as indicated by their

emotion ratings following complete feedback that strongly contrasted with their ratings following partial feedback. This suggests that any explanation of children's emotion ratings in this task just in terms of a simple comparison such as "I do [do not] have the best prize" is inadequate, whereas such an explanation may be sufficient to explain such ratings following complete feedback in previous studies with young children. Following a safer choice, children are already aware, even before partial feedback, that they do not have the best or worst prize; their subsequent emotion ratings following partial feedback seem to be sensitive to what they could have won given their choice and, following complete feedback, sensitive to what they would have won had they chosen differently.

Does this mean that we can rule out an explanation of children's emotions in terms of simple frustration rather than regret or relief (see Rafetseder & Perner, 2012; Rafetseder & Perner, 2014)? We are assuming that this question remains open even under circumstances in which participants have full responsibility for the choice that led to the outcome and full feedback has been given (see O'Connor, McCormack, Beck, & Feeney, 2015). Certainly, any explanation of children's emotions in terms of frustration needs to be more complex than assuming a simple comparison between the prize obtained and the best or worst prize available; it would need to assume that this frustration can mutate within a single trial following complete feedback (from frustration to happiness in relief trials or from happiness to frustration in regret trials). One might argue that all that needs to happen is for children to systematically change the reference point by which they make their judgment regarding whether they have the best (or worst) prize and that this need not involve thinking counterfactually about their choices. We note, however, that children's judgments following complete feedback do seem to be sensitive not just to whether or not a better (or worse) prize was available but also to whether children would have won that prize had they chosen differently. That is, they made the comparison not between their actual prize and the possible contents of the unchosen gamble but rather between their actual prize and the specific outcome they would have obtained had they chosen the other gamble. Although we cannot definitely rule out the possibility that children make these comparisons without thinking counterfactually, we note that neuropsychological evidence suggests that the regions of the brain that seem to be important for being sensitive to complete feedback in this sort of gambling task also seem to be important for counterfactual thinking (Barbey, Krueger, & Grafman, 2009; but see Van Hoeck et al., 2013).

Within-trial versus between-trial comparisons of emotion ratings

As described in the Introduction, Rafetseder and Perner (2012), Rafetseder and Perner (2014) argued that a between-trial method is the most appropriate one to measure regret, whereas all of the other studies with young children have used only a within-trial method. We directly compared both of these methods in Experiment 1 in the current study to examine whether they showed different developmental patterns. For both the within-trial and between-trial measures, the developmental patterns were very similar, and under either method there was some evidence of a developmental increase in the likelihood that participants experienced regret but no evidence of developmental changes in relief. Our findings suggest that Rafetseder and Perner's pattern of developmental findings cannot be explained simply in terms of their use of a between-trial method for assessing regret. Of course, although the basic comparison we made in our between-trial method was in essence the same as that made in Rafetseder and Perner's between-trial method (i.e., a comparison of emotion ratings following either partial or complete feedback on separate trials), our task differed from that of Rafetseder and Perner considerably, with the latter involving a simple box-choosing task. These differences included not just the types of choice children needed to make but also the number of trials, the nature of the emotion scale used, the rewards involved, and so on. However, it is hard to see why these differences would make it more likely, rather than less likely, that we would observe counterfactual emotions in younger children; moreover, we note that in their own study, Rafetseder and Perner found no consistent evidence of regret in children under 9 years of age even using a within-trial method.

The main reason for comparing the between-trial and within-trial methods was to inform future studies of the development of counterfactual emotions. Although both of these methods yielded broadly similar results, the within-trial method requires administration of fewer trials and also seems to yield a somewhat clearer pattern of findings. In Experiment 1, we found that the difference in the

proportion of sadder responses given in regret trials versus relief trials was only marginally significant for the younger children in the between-trial method and that evidence for relief in the older group was significantly stronger using the within-trial method. One possible explanation of the clearer pattern of findings using the within-trial method may lie in the fact that in Two Rating trials we explicitly asked participants to make a comparative rating, that is, to report whether they felt the same as, happier than, or sadder than before they had received complete feedback. This may have encouraged participants to proactively compare the actual outcome with the counterfactual outcome. Moreover, these comparisons did not hinge on the sensitivity of the scale used to report emotions because we used the three-pronged arrow method of describing relative emotions (see also O'Connor et al., 2012; O'Connor et al., 2014; Weisberg & Beck, 2012). By contrast, between-trial comparisons involved comparing separate ratings on the same 7-point scale, and whether such comparisons differ will depend on the sensitivity of this scale. Our findings suggest that future developmental studies should use the within-trial method, at least for this type of task.

Conclusion

This study has demonstrated that children as young as 6 or 7 years do experience counterfactual emotions in the context of risky decision making; as such, it may pave the way for future studies that examine whether these emotions have an impact on children's decision making. Such studies are important because existing accounts of developmental changes in children's decision making (for reviews, see Dhami, Schlottmann, & Waldmann, 2013; Jacobs & Klaczynski, 2005) have typically not considered how counterfactual thinking and its emotional consequences have an impact on children's decisions, whereas it is widely argued that counterfactual cognition and its associated emotions play an important role in adult decision making (Connolly & Zeelenberg, 2002; Coricelli et al., 2007; Roese, 1999).

Acknowledgment

This research was supported by a grant from the Economic and Social Research Council (UK, Grant ES/K000411/1).

References

Barbey, A. K., Krueger, F., & Grafman, J. (2009). Structured event complexes in the prefrontal cortex support counterfactual representations for future planning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*, 1291–1300.

Bault, N., Coricelli, G., & Rustichini, A. (2008). Interdependent utilities: How social ranking affects choice behavior. *PLoS ONE, 3* (10), e3477.

Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives, 8*, 175–179.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research, 30*, 961–981.

Burnett, S., Bault, N., Coricelli, G., & Blakemore, S. J. (2010). Adolescents' heightened risk-seeking in a probabilistic gambling task. *Cognitive Development, 25*, 183–196.

Burns, P., Riggs, K. J., & Beck, S. R. (2012). Executive control and the experience of regret. *Journal of Experimental Child Psychology, 111*, 501–515.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science, 304*, 1167–1170.

Chua, H. F., Gonzalez, R., Taylor, S. F., Welsh, R. C., & Liberzon, I. (2009). Decision-related loss: Regret and disappointment. *NeuroImage, 47*, 2031–2040.

Connolly, T., & Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science, 11*, 212–216.

Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A., & Dolan, R. J. (2005). Regret and its avoidance: A neuroimaging study of choice behaviour. *Nature Neuroscience, 8*, 1255–1262.

Coricelli, G., Dolan, R. J., & Sirigu, A. (2007). Brain, emotion, and decision making: The paradigmatic example of regret. *Trends in Cognitive Sciences, 11*, 258–265.

Dhami, M. K., Schlottmann, A., & Waldmann, M. (2013). *Judgment and decision making as a skill: Learning, development, and evolution*.Cambridge, UK: Cambridge University Press.

Habib, M., Cassotti, M., Borst, G., Simon, G., Pineau, A., Houdé, O., et al (2012). Counterfactually mediated emotions: A developmental study of regret and relief in a probabilistic gambling task. *Journal of Experimental Child Psychology, 112*, 265–274.

Jacobs, J. E., & Klaczynski, P. A. (2005). *The development of decision making in children and adolescents*.Mahwah, NJ: Lawrence Erlbaum.

Levens, S. M., Larsen, J. T., Bruss, J., Tranel, D., Bechara, A., & Mellers, B. A. (2014). What might have been? The role of the ventromedial prefrontal cortex and lateral orbitofrontal cortex in counterfactual emotions and choice. *Neuropsychologia, 54*, 77–86.
Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal, 92*, 805–824.
McCormack, T., & Feeney, A. (2015). The development of the experience and anticipation of regret. *Cognition and Emotion, 29*, 266–280.
Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General, 128*, 332–345.
O'Connor, E., McCormack, T., Beck, S., & Feeney, A. (2015). Regret and adaptive decision making in young children. *Journal of Experimental Child Psychology, 135*, 86–92.
O'Connor, E., McCormack, T., & Feeney, A. (2012). The development of regret. *Journal of Experimental Child Psychology, 111*, 120–127.
O'Connor, E., McCormack, T., & Feeney, A. (2014). Do children who experience regret make better decisions? A developmental study of the behavioral consequences of regret. *Child Development, 85*, 1995–2010.
Rafetseder, E., & Perner, J. (2012). When the alternative would have been better: Counterfactual reasoning and the emergence of regret. *Cognition and Emotion, 26*, 800–819.
Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives, 8*(1), 54–58.
Roese, N. (1999). Counterfactual thinking and decision making. *Psychonomic Bulletin and Review, 6*, 570–578.
Sommer, T., Peters, J., Gläscher, J., & Büchel, C. (2009). Structure–function relationships in the processing of regret in the orbitofrontal cortex. *Brain Structure and Function, 213*, 535–551.
van Duijvenvoorde, A. C. K., Huizenga, H. M., & Jansen, B. R. J. (2014). What is and what could have been: Experiencing regret and relief across childhood. *Cognition and Emotion, 28*, 926–935.
Van Hoeck, N., Ma, N., Ampe, L., Baetens, K., Vandekerckhove, M., & Van Overwalle, F. (2013). Counterfactual thinking: An fMRI study on changing the past for a better future. *Social Cognitive and Affective Neuroscience, 8*, 556–564.
Weisberg, D. P., & Beck, S. R. (2010). Children's thinking about their own and others' regret and relief. *Journal of Experimental Child Psychology, 106*, 184–191.
Weisberg, D. P., & Beck, S. R. (2012). The development of children's regret and relief. *Cognition and Emotion, 26*, 820–835.
Zalla, T., Sirigu, A., Robic, S., Chaste, P., Leboyer, M., & Coricelli, G. (2014). Feelings of regret and disappointment in adults with high-functioning autism. *Cortex, 58*, 112–122.
Zeelenberg, M. (1999). Anticipated regret, expected feedback, and behavioral decision-making. *Journal of Behavioral Decision Making, 12*, 93–106.
Zeelenberg, M., Beattie, J., van der Pligt, J., & de Vries, N. K. (1996). Consequence of regret aversion: Effects of expected feedback on risky decision making. *Organizational Behavior and Human Decision Processes, 65*, 148–158.
Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology, 17*(1), 3–18.
Zeelenberg, M., van Dijk, W. W., & Manstead, A. S. R. (2000). Regret and responsibility resolved? Evaluating Ordóñez and Connolly's (2000) conclusions. *Organizational Behavior and Human Decision Processes, 81*, 143–154.