Cognition & Emotion

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pcem20

When the alternative would have been better: Counterfactual reasoning and the emergence of regret

Eva Rafetseder & Josef Perner

a Department of Psychology, University of Salzburg, Salzburg, Austria
b Centre for Neurocognitive Research, University of Salzburg, Salzburg, Austria

Available online: 06 Oct 2011

To cite this article: Eva Rafetseder & Josef Perner (2011): When the alternative would have been better: Counterfactual reasoning and the emergence of regret, Cognition & Emotion, DOI:10.1080/02699931.2011.619744

To link to this article: http://dx.doi.org/10.1080/02699931.2011.619744

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
When the alternative would have been better: Counterfactual reasoning and the emergence of regret

Eva Rafetseder¹ and Josef Perner¹,²

¹Department of Psychology, University of Salzburg, Salzburg, Austria
²Centre for Neurocognitive Research, University of Salzburg, Salzburg, Austria

Counterfactual reasoning about how events could have turned out better is associated with the feeling of regret. However, developmental studies show a discrepancy between the onset of counterfactual reasoning (at 3 years) and the feeling of regret (at 6 years). In four experiments we explored possible reasons. Experiment 1 (3- to 6-year-old children) and Experiment 2 (adult control) show that even when regret is assessed more directly than in previous studies (e.g., Amsel & Smalley, 2000) only adults but not children regret their decision. Experiment 3 (3- to 14-year-old children) suggests that double-questioning—asking children how happy they are with what they got before and after they had seen what they could have got—creates false positive indications of regret in the youngest children and that—when controlling for false positives—regret is not evident before 9 years. However, children before this age make a difference between attractive (three candies) and less attractive (one candy) items (Experiment 4; 6- to 8-year-old children). Taken together, this suggests that before 9 years of age children base their judgements solely on what they got without taking into account what they could have got.

Keywords: Counterfactual reasoning; Regret; Children; Adults.

In daily life people often reflect on how things would be different if an event in the past had turned out differently. When people create such alternatives to known facts, this is called counterfactual reasoning. In counterfactual reasoning actual states of affairs are compared to unrealised possibilities (Kahneman & Miller, 1986; Kahneman & Tversky, 1982). Markman, Gavanski,
Sherman, and McMullen (1993) differentiated between counterfactuals in which people come up with alternatives worse than reality, called down-ward counterfactuals, and those in which they come up with alternatives better than reality, called upward counterfactuals. Such comparisons can have affective consequences, i.e., reflecting on ways in which events might have turned out better or how negative events could have been avoided is strongly associated with feeling worse. Research has therefore focused on the relation between counterfactual reasoning and the feeling of regret (Gleicher et al., 1990; Kahneman & Miller, 1986; Kahneman & Tversky, 1982; Landman, 1987; Roese, 1997; Wells & Gavanski, 1989) as well as the cognitive processes and their neural bases that underlie regret (Camille et al., 2004; Coricelli et al., 2005; Coricelli, Dolan, & Sirigu, 2007; Nicolle, Bach, Driver, & Dolan, 2011; Nicolle, Bach, Frith, & Dolan, 2011; Nicolle, Fleming, Bach, Driver, & Dolan, 2011). Camille et al. (2004), for example, used a regret gambling task in which participants (patients with lesions of the orbitofrontal cortex, OFC, and control participants) had to choose between two "wheels of fortune". Each wheel had two possible outcomes. If participants chose wheel 1, they could win 200 units with 20% probability or lose 50 units with 80% probability. If they chose wheel 2, they could win or lose 50 units with equal probabilities. In the partial feedback condition, participants only got to know about the outcome of the chosen wheel, whereas in the complete feedback condition they were also shown the outcome of the unchosen wheel. Participants had to rate their affective state on a rating scale ranging from -50 (extremely sad) to +50 (extremely happy).

In the partial feedback condition, both patients and control participants were happy when winning and sad when losing. In the complete feedback condition, however, the two groups showed a different pattern of results. A win of 50 units induced a strong negative emotion in control participants when the other wheel won 200 units. This was perceived even more negatively than losing 50 units when the unchosen wheel lost 200 units. In contrast, the emotional rating of patients with a lesion of the OFC was not modulated at all by the outcome of the unchosen wheel. Moreover, while control participants chose the more advantageous wheel (in the sense of reducing opportunity for regret) after experiencing regret several times, patients with orbitofrontal lesions did not change their strategies. So, patients with lesions of the OFC seem to base their judgements on what has actually occurred. They do not experience or anticipate regret as control participants do, which indicates that the orbitofrontal cortex might play a crucial role in mediating the experience of regret (Camille et al., 2004).

In studies that explore the development of regret, children are usually presented with a choice of two possible courses of action, where the outcome either has an effect on them (Amsel & Smalley, 2000; Weisberg & Beck, 2010) or on other people (Beck & Crilly, 2009; Guttentag & Ferrell, 2004, 2008; Weisberg & Beck, 2010). After one course of action has been taken, children are supposed to judge its outcome by taking the counterfactual alternative into account. All these studies are based on the assumption that children who do not reason counterfactually will not experience or understand regret and they have good reasons to do so as neuropsychological and neuroimaging studies support the idea that regret is cognitively mediated by counterfactual reasoning (e.g., Camille et al., 2004).

Recent developmental research, however, does not point to a developmental linkage between counterfactual reasoning and feeling regret. Developmental studies that assessed the ability to reason counterfactually by telling a short story (e.g., "Carol made the clean floor all dirty with her shoes") and then asking a counterfactual question, "If Carol had taken her shoes off, would the floor be dirty or clean?" (Harris, German, & Mills, 1996), suggest that 3-year-old children are able to reason counterfactually. However, studies which assessed children's experience of regret found first signs only from 5 years on (Weisberg & Beck, 2010) and others even later, at 7 years (Amsel et al., 2003). Children's understanding of regret in other people seems to develop at around 7 years.
(Beck & Crilly, 2009; Guttentag & Ferrell, 2004; Weisberg & Beck, 2010), and their understanding of anticipatory regret not before the age of 9 years (Guttentag & Ferrell, 2008).

This age discrepancy seems puzzling. If regret does require counterfactual reasoning, one would expect that children who are able to reason counterfactually at the age of 3 years would also be able to experience regret at that age. So, what is it that makes for this discrepancy? One hypothesis is that by 3 years children are able to reason counterfactually and are able to feel counterfactual emotions like regret but the tasks used to elicit regret are needlessly complicated for assessing the feeling of regret in children.

For instance, in Amsel and Smalley’s (2000) study (see also Amsel et al., 2003, for further details), children aged 5 years, 7 years and 10 years as well as college students had to turn over one of two cards lying face down. Only when the selected card matched the experimenter’s face-up card was it treated as an experimental trial. In that case participants were asked to turn over the unselected card, which was either higher or lower than the selected card, and were asked the counterfactual question: “How would you have felt if you had turned over the unchosen card?” All participants rated their feelings about the unselected card in the same way. They judged that they would have been happier having turned over the unselected card, if it would have won and less happy if it would have lost. Notably, in this part of the study children and adults did not have to take reality (their actually selected card) into account when they rated how they felt. They just had to compare the unselected card with the experimenter’s face-up card and rate their feelings about winning or losing. Then participants were asked the final question: “Now that you’ve seen the other card, how do you feel about the card you selected?” Here, 5-year-olds differed in their judgements from all the other participants. Their ratings of the selected card stayed the same as it was before they had seen the unselected card, while all the other participants changed their ratings of the selected card with respect to the unselected card. For this final question participants had to take the actual selected card into account. Only 7-year-olds and older participants but not preschoolers were influenced by the information of a better or worse counterfactual outcome of choosing the unselected card.

There is, however, a needless complication in this task which is due to the fact that participants experience regret right after they turn over the unselected card and realise that this card would have won. But, the question about regret (“Now that you’ve seen the other card, how do you feel about the card you selected?”) is asked much later, after other questions had been asked in between. This may destroy a natural tendency to feel less happy with a win in view of a better alternative than in ignorance of that alternative. It might then happen that children simply feel they should not be inconsistent and, therefore not change their rating from before on the same result. If this hypothesis is correct, then children younger than 7 should be able to show regret in procedures that assess the feeling of regret more directly, i.e., without repeating the same question.

However, it could very well be that children just don’t make these comparisons spontaneously although they are actually competent at counterfactual reasoning. In a study by Beck and Crilly (2009) 5-and 6-year-old children were shown stories (based on Guttentag & Ferrell, 2004) in which two characters (David and Bob) experienced the same negative outcome, e.g., falling off their bike after hitting a tree branch that had fallen across the red path. While for Bob the red path was his usual route, for David this was an unusual route. He usually goes on the yellow path and had only that day decided to go on the red path. Children were asked the regret question (“Who would be more upset about deciding to ride along the red path that day?”) and the open counterfactual question (“Could David and Bob have gone another way?”). The regret question was significantly more difficult than the open counterfactual question and no child passed the regret question without also passing the open counterfactual question. The authors concluded that children might fail to spontaneously consider the alternative world implied in the regret ques-
tion but when directly asked to do so, i.e., in the open counterfactual question, they can do it. However, it might be the case that it is not children’s application of counterfactual reasoning that directs them to answer correctly in the open counterfactual question but their linguistic expertise. They know that the word “other” simply means “other than where Bob and David have gone now”.

This connects with a further reason for why children are able to answer counterfactual questions long before showing regret. It might be that the so-called counterfactual tasks do not really require children to reason counterfactually. In other words, younger children’s correct answers to counterfactual questions might not be based on counterfactual reasoning by comparing what has actually happened with what could have happened. Such a comparison is, however, a necessary ingredient for experiencing regret (Guttentag & Ferrell, 2004). In response to counterfactual questions young children might use basic conditional reasoning instead, which yields the same answers as counterfactual reasoning, e.g., in the study by Harris et al. (Perner & Rafetseder, 2011; Rafetseder, Cristi-Vargas, & Perner, 2010). Basic conditional reasoning consists of applying regularities (e.g., If [whenever] one takes off one’s shoes, floors tend to stay clean) to an imagined event. In this case, children might treat counterfactual questions (e.g., “If Carol had taken her shoes off, would the floor be dirty or clean?”) just like an indicative future question (e.g., “If Carol takes her shoes off, will the floor be dirty or clean?”). According to Perner (2000) this ignores the requirement for counterfactual reasoning of relating the counterfactual stipulation (If Carol had taken her shoes off) to what actually happened (Carol left her dirty shoes on and, therefore, made the floor all dirty). Depending on the particular task children possibly arrive at the correct counterfactual answer via basic conditional reasoning, which makes their answers look like evidence for counterfactual reasoning. If the early evidence for counterfactual reasoning is, indeed, based on false positives—by applying basic conditional reasoning—then we would not expect these children to experience regret. Regret should only be shown by children who are able to reason counterfactually. For only they compare what they actually got with what they could have got by being able to relate counterfactual stipulations to actual events.

Also the counterfactual question “How would you have felt if you had turned over the unchosen card?” in Amsel and Smalley’s (2000) study may not have required counterfactual reasoning. As noted earlier, participants did not have to take the actually selected card into account in order to answer this question correctly. They only had to compare the not selected card with the experimenter’s card, judge who would have won (the use of the subjunctive focuses on the pretend character of this event) and rate their feelings about the outcome. That may explain why 5-year-olds were so good at answering this question but had a hard time answering the final question: “Now that you’ve seen the other card, how do you feel about the card you selected?”, which did require them to compare the real outcome (value of selected card) with the counterfactual outcome (value of the not selected card).

In four experiments we tried to control for all the alternative interpretations of existing experiments as stated above. In Experiment 1 children were shown two boxes and they had to choose one. They were told that they could keep the content. In the independent condition we opened the selected box, gave its content (e.g., one candy) to the children, opened the non-selected box (with, e.g., three candies) and asked them: “How happy are you with your one candy?” In the dependent condition the only difference was that we asked children twice how they felt about what they got: before and after they had seen the alternative (which may make the rating of the second question dependent on the rating of the first question). This is to control for the procedures used in both Amsel and Smalley’s (2000) and Weisberg and Beck’s (2010) studies, where participants were asked twice about the actual
outcome of the game, before and after the unselected option had been revealed to them.

**EXPERIMENT 1**

Children had the choice between two boxes. They were told that whatever box they chose they could keep the contents. In the dependent condition children were asked to rate their feelings about their reward twice, one time before (Baseline Question) and one time after (Dependent Test Question) they had seen what was in the unselected box. When they were asked the second time to rate their feelings, it was explicitly pointed out that they should do so in view of the contents of the non-chosen box: “Now that you know that you could have got ...”. In the independent condition children were asked to rate their feelings just once, after they had seen what was in the box they did not choose (Independent Test Question). They were asked the same question as the Baseline Question of the dependent condition (“How happy are you with ...?”), so no explicit reference to the contents of the unselected box was made.

If participants regret their decision then they should give a lower rating in response to both Test Questions (because they already know what they could have got) than to the Baseline Question, where they only know what they actually got. The most reliable comparison indicating regret is, however, the one between the Baseline Question and the Independent Test Question.

In the introduction section we offered several explanations for the apparent discrepancy between children’s early ability to reason counterfactually and the late signs of showing regret. One hypothesis is that the procedure used by Amsel and Smalley (2000; Amsel et al., 2003) is unnecessarily distracting. It ruins the natural, spontaneous feeling of regret and, therefore, underestimates children’s ability to feel regret. If this were true, we would find signs of regret earlier than 7 years of age in our experiment, because our assessment is the most direct possible. This should be the case in particular for the Independent Question, since it avoids the distractive repetition of the Dependent Question.

Another explanation was that counterfactual tasks can be answered correctly without counterfactual reasoning but with a more general ability to reason with conditionals instead. If this hypothesis is correct children should not show signs of regret (at least in our independent condition, which controls for false positives due to repeated questioning) before the age of 6 years (and probably even older) as counterfactual reasoning—necessary for experiencing regret—does not seem to develop before that age (Rafetseder et al., 2010).

**Method**

*Participants.* The participants were 37 children (19 girls, 18 boys) from a nursery school in the city of Salzburg. Ages ranged from 3;0 (years; months) to 6;8 with a mean age of 5;1 and standard deviation of 10 months. For statistical analyses, age (in days) was treated as a continuous variable but for display purposes children were divided into two age-groups. These consisted of 16 children (age ranges from 3;0 to 4;11, $M = 4;4$, $SD = 6$ months) in the younger group and 21 children (age ranges from 5;0 to 6;8, $M = 5;8$, $SD = 6$ months) in the older group. Two children (one of the younger and one of the older age group) refused to answer the Independent Test Question but stayed in the sample for the analyses of the Dependent Test Question. We checked that range, mean age and standard deviation were the same for both age groups with and without these two children.

*Materials.* We used two identical looking boxes. In the middle of each box was a barrier. One side of each box contained one candy (less attractive item) and the other side contained five candies (attractive item). Each box could be opened on the left as well as on the right side. Regardless of which box was chosen by the child, it was always the side opened that contained the less attractive item. Additionally we used an 8-point scale
consisting of eight different smiley-faces ranging from being very happy (8) to being very sad (1).

**Design.** Each child was tested in two sessions approximately one week apart. Sessions lasted about 10 minutes. There were two different conditions (dependent and independent) and children were given one condition per session (in a fully counterbalanced order).

**Procedure.** First, children were introduced to the smiley-faces on which they could show how happy they were. They were asked to point at the happiest and then at the unhappiest face. In both conditions children were then given the choice between the two boxes. They were told that whatever box they choose they could keep the content. Each box had two chambers that contained one candy on one side and five candies on the other side. The child selected a box and then always the side with the less attractive item (one candy) was opened. In the dependent condition the child was always told: “Oh, there is one candy. It’s yours now”. Then they were asked the Baseline Question: “How happy are you with your one candy? Can you show me on our scale?” After the child had shown it on the scale, he/she was told: “Now I am going to show you what is in the other box. Oh, look! There are five candies in it”. Then the Dependent Test Question was asked: “Now that you know that you could have got five candies, how happy are you with your one candy? Can you show me on our scale?” In the independent condition regret was assessed more directly. After the selected box had been opened the child was told: “Oh, there is one candy. It’s yours now. Now I am going to show you what is in the other box. Oh, look! There are five candies in it”. Right after that the Independent Test Question (same as Baseline Question of the dependent condition) was asked: “How happy are you with your one candy? Can you show me on our scale?”

**Results**

For each question the emotion ratings were coded numerically, with a “very happy” rating coded as 8 and a “very unhappy” rating coded as 1.

We ran a mixed analysis of co-variance (ANCOVA) with Type of Question (baseline, dependent and independent) as within-subject factor, Order (dependent first vs. independent first) as between-subject factor and Age (in days) as covariate. Mauchly’s test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 7.98, p < .05$, therefore degrees of freedom were corrected using Huynh–Feldt estimates of sphericity ($\epsilon = .91$). The results show that there was no significant effect of which Type of Question was asked on the emotional ratings, $F(1.81, 58) = 0.82, p = .43, \eta^2_p = .03$. There was also no discernible systematic effect of the Order in which the two conditions followed each other, $F(1.81, 58) = 2.09, p = .14, \eta^2_p = .06$. The covariate, Age, was not significantly related to the emotional ratings $F(1.81, 58) = 0.62, p = .53, \eta^2_p = .02$.

Figure 1 shows the mean ratings for each question by age groups. Interestingly, in the younger group there is a tendency to rate the Dependent Test Question lower than the Baseline Question. Nine children (56%) gave lower ratings in response to the Dependent Test Question compared to the Baseline Question, but only three children (19%) did so for the Independent Test Question. In the older age group six children (29%) showed lower ratings in response to the Dependent Test Question compared to the Baseline Question and eight (38%) did so for the Independent Test Question. Only two children of each age group (younger: 13%; older: 10%) were less happy on both Test Questions in comparison to the Baseline Question.

**Discussion**

In Experiment 1 most of the children’s ratings were uninfluenced by knowledge about what they could have got, neither in the dependent nor in the independent condition. This is compatible with findings in other studies that children younger than 7 years of age do not experience regret themselves (Amsel et al., 2003) or understand regret in others (Beck & Crilly, 2009; Gutten tag & Ferrell, 2004; Weisberg & Beck, 2010). However, in the present study there was a
tendency for especially younger children to give a lower rating on the Dependent Test Question than on the Baseline Question but there was no such difference for the Independent Test Question. This finding is consistent with Weisberg and Beck (2010), who found a difference on a dependent design between ratings before and after revealing the alternative in 5-year-old children.

The reason why lower ratings are observed in the dependent but not in the independent condition may be twofold. As Beck and Crilly (2009) argued, children may fail to spontaneously take the alternative outcome into account but they are generally able to reason counterfactually. Alternatively, children at this age may not feel regret but may feel the need to change their answer when asked the same question a second time. The reason why children tend to give lower rather than higher ratings on the second rating may be caused by the near ceiling first ratings. Indeed, seven of the nine children (78%) in the younger age group and three of the six children (50%) in the older age group who lowered their ratings on the second Test Question had given top ratings “very happy” (8) on the Baseline Question. We tested for these two possibilities in Experiment 3.

Another question that arises is whether adults would show regret on our tasks by giving lower ratings on the Test than on the Baseline Question. It could be that this task is not appropriate for measuring regret at all. This issue was addressed in Experiment 2.

**EXPERIMENT 2**

The results of the first experiment indicated that even at 6 years few children, at best 38%, showed some evidence of regret. This raises the question of whether adults show regret on this kind of task. In the following experiment we, therefore, gave a slightly changed, age-appropriate version of this task to a group of adults. The adults were told that if they participated in a psychological experiment they would be allowed to take part in a gamble where they could win some money (instead of candies). In this gamble they had to choose two boxes out of five. The rest of the task was similar to the first Experiment.

**Method**

**Participants.** The participants were 39 adults aged between 18 and 84 years (mean age of 38 years and standard deviation of 20 years). The 20 adults who were exposed to the dependent condition had a mean age of 47 years (SD = 22 years). The 19 adults who were exposed to the independent condition had a mean age of 29 years (SD = 11 years). Participants were recruited via
the internet. They were informed that if they took part in a psychological experiment they could play a game where they could win money. Adults who agreed were invited to an office at the university where the game took place.

**Materials.** We used five identical looking boxes. Each box had two chambers one of which contained €0.5 and the other €50. For happiness ratings a 10-point scale (10 circles) ranging from happy (10), indicated with a happy smiley face, to unhappy (1), indicated with an unhappy smiley face, was used.

**Design.** Each adult was tested in one session that lasted about 5 minutes and was exposed to either the independent or the dependent condition.

**Procedure.** Participants were tested individually and were informed that they were taking part in a psychological experiment and that for a reward they could first participate in a gamble. They were then asked to select two boxes out of the five. The other three boxes were set aside and participants again had to select one of the two remaining boxes. Then always the side with the €0.5 was opened. In the dependent condition the participant was told: “Oh, here is €0.5. It is yours now.” Then the Baseline Question was asked: “How happy are you with this €0.5? Please tick the circle on the scale!” After participants had shown their happiness on the scale, they were told: “Now I am going to show you what is in the other box. There are €50 in it”. Then the Test Question was asked: “Now that you know that you could have got €50, how happy are you with your €0.5? Please tick the circle on the scale!” In the independent condition—after the selected box had been opened—the participant was told: “Oh, here is €0.5. It is yours now. Now I am going to show you what you would have got if you had chosen the other box. Oh, look! You would have got €50”. Right after that the Test Question was asked: “How happy are you with the €0.5? Please tick the circle on the scale!”

**Results**

For each question the emotion ratings were coded numerically, with a “very happy” rating coded as 10 and a “very unhappy” rating as 1. An independent-samples t-test revealed a highly significant difference between the mean of the Baseline Question ($M = 6.4$) and the mean of the Independent Test Question ($M = 3.6$), $t(37) = 3.58$, $p = .001$, Cohen’s $d = 1.18$. A paired-samples t-test showed a highly significant difference between the mean of the Baseline Question ($M = 6.4$) and the mean of the Dependent Test Question ($M = 3.3$), $t(19) = 5.64$, $p < .001$, Cohen’s $d = 1.26$. An independent-samples t-test for the mean of the Dependent Test Question and the mean of the Independent Test Question was not significant, $t(37) = 0.39$, $p = .70$, Cohen’s $d = 0.13$. So, adults clearly showed regret on both conditions and the conditions themselves did not differ.

**Discussion**

Experiment 2 examined whether adults would show regret on the task used with children in Experiment 1. Regardless of whether it was explicitly pointed out to them that they could have won €50 (“Now that you know that you could have got €50, how happy are you with the €0.5?”) or not (“How happy are you with the €0.5?”), their ratings were significantly lower on the Test Questions than on the Baseline Question. This indicates that adults show regret on this kind of task, which has also been found in other studies (Amsel & Smalley, 2000; Guttentag & Ferrell, 2004, 2008; Weisberg & Beck, 2010).

**EXPERIMENT 3**

In Experiment 2 adults showed regret on our tasks. It did not make any difference whether it was explicitly pointed out to them or left implicit that they could have got more than they really got. Two further questions arise from these two experiments. First, when do children start to show regret in the independent condition of our
task? To answer this question we gave the child version of the task to children aged between 3 and 14 years. Second, why did the younger group in Experiment 1 give lower ratings in response to the Dependent Test Question than the Baseline Question but not in response to the Independent Test Question? One hypothesis is that they did so because it was explicitly pointed out to them by saying: “Now that you know what you could have got . . .”. This made them compare what they got with what they could have got but they did not do so spontaneously in the independent condition (Beck & Crilly, 2009). Another possible explanation could be that double-questioning may have prompted the younger children especially to give a different answer than they had given to the Baseline Question. Since children tended to give very high baseline ratings they tended to lower their rating the second time they were asked.

The procedure of Experiment 3 was similar to the first experiment. However, this time not only quantitative material, i.e., a different number of candies, was used but also qualitative material, i.e., stickers of different attractiveness (see Experiment 4 for assurance that attractive stickers were rated better by children than less attractive stickers). Therefore it was possible to use different materials for the dependent and the independent conditions in order to minimise any influence of the condition administered first on the one administered second. Moreover, a finer scale was used, ranging from “very happy” (64) to “very unhappy” (1). The reason for that was to reduce the ceiling effect observed in Experiment 1. Using an adjustable vertical scale on the computer (top is very happy, bottom is very sad) made it less likely for children to stick to a particular face that they liked best. So, children had to decide for every question again how they really felt. Moreover, all three questions (Baseline, Dependent Test and Independent Test) were phrased exactly the same way: “How happy are you with your one candy?” By deleting the part “Now that you know what you could have got . . .” from the Dependent Test Question we could test whether the youngest children gave lower ratings to this question than to the Baseline Question in Experiment 1 because of the explicit hint that they should make a comparison or because they had been asked the same question a second time. If it was really the explicit hint that made them answer differently, then this time they should give the same rating on the Dependent Test Question as on Baseline.

### Method

**Participants.** The participants were 98 children (47 girls, 51 boys) from different nursery schools and youth centres in the cities of Salzburg and Munich. Ages ranged from 2;9 (years; months) to 14;4 with a mean age of 8;1 and standard deviation of 38 months. For statistical analyses, age (in days) was treated as a continuous variable but for display purpose children were divided into five age-groups. These consisted of 18 children (age range from 2;9 to 4;9, $M=4;0$, $SD=7$ months) in the youngest group, 26 children (age range from 5;1 to 6;7, $M=6;0$, $SD=6$ months) in the second group, 19 children (age range from 6;9 to 8;9, $M=8;1$, $SD=8$ months) in the third group, 19 children (age range from 9;0 to 11;1, $M=9;9$, $SD=7$ months) in the fourth group and 16 children (age range from 12;4 to 14;4, $M=13;6$, $SD=7$ months) in the oldest group. Two children (one 5- and one 11-year-old girl) were dropped from the analyses because the difference of their answers between baseline and independent question was more than three standard deviations away from the mean.

**Materials.** We used two sets of identically looking boxes (set 1: two red boxes; set 2: two green boxes). Each of these boxes had two chambers that contained a relatively unattractive item (one candy/a black circle sticker: 1.5 cm in diameter; for the oldest group: €0.5/a not very attractive black fantasy animal that sticks to the skin and looks like a tattoo, later referred to as “skin-sticker”) and an attractive reward (three candies/a beautiful sticker: for boys we used a coloured excavator, size: 4 cm; for girls we used a coloured animal with glitter; size: 3.5 cm; for the oldest group: €5/a beautiful “skin-sticker”: for boys we used a
coloured dragon, size: 6 cm; for girls we used a coloured ornament; size: 7 cm).

In addition we used a 64-point scale (adjustable vertical scale on computer; top was very happy, bottom was very sad) that consisted of a 230 mm long vertical ruler with a colour-shape gradient (in the form of a bar) from blue (“unhappy”) to orange (“happy”). The “unhappy” part (coloured in blue) ranged from 1 to 32 and the “happy” part (coloured in orange) ranged from 33 to 64. Next to the blue part there was a smiley face with an unhappy expression, next to the orange part there was a smiley face with a happy expression. Ratings were made by moving the bar (with the wheel of the mouse or with the up and down arrows on the keyboard), which started between the blue and the orange part. By moving downward the bar became more and more blue and by moving upward the bar became more and more orange. More happiness was represented by a higher position of the bar (max. number: 64).

**Design.** Each child was tested in one session that lasted about 10 minutes. Both conditions were administered to every child in a fully counterbalanced order.

**Procedure.** Before the test started children were introduced to the scale. They were asked to point at the happy face and then at the unhappy one. Then they were told: “When my mum cooks my most favourite meal for me, I am happy. On the scale it looks like that.” (the bar of the scale was adjusted to 50). “When my mum tells me off, I am unhappy. On the scale it looks like that.” (the bar of the scale was adjusted to 15). “Now you can play with the scale. Show me how you feel when your mum cooks your most favourite meal. Okay, and now show me how you feel when your mum tells you off.” After the child had adjusted the bar to the orange part and then to the blue part the test started.

The child selected a box and then always the side with the one candy (or the not very beautiful sticker/skin-sticker/€0.5) was opened. In the dependent condition the children were told: “Oh, there is one candy. It’s yours now”. Then the Baseline Question was asked: “How happy are you with your one candy? Can you show me on our scale?” After the child had shown it on the scale, he/she was told: “Now I am going to show you what you would have got if you had chosen the other box. Oh, look! You would have got three candies”. Then the Test Question was asked: “How happy are you with your one candy?”

The independent condition was similar. However, no Baseline Question was asked. After the selected box had been opened the children were told: “Oh, there is one candy. It’s yours now. Now I am going to show you what you would have got if you had chosen the other box. Oh, look! You would have got three candies”. Right after that the Test Question was: “How happy are you with your one candy?”

**RESULTS**

For each question the emotion ratings were coded numerically, with a “very happy” rating coded as 64 and a “very unhappy” rating coded as 1.

We ran a mixed ANCOVA with Type of Question (baseline, dependent and independent) as within-subject factor, Order (dependent first vs. independent first) as between-subject factor and Age (in days) as covariate. Mauchly’s test indicated that the assumption of sphericity had been violated, \(\chi^2(2) = 6.13, p < .05\), therefore degrees of freedom were corrected using Huynh–Feldt estimates of sphericity (\(\varepsilon = .98\)). The results show that there was no significant effect of Type of Question, \(F(1.96, 186) = 0.45, p = .63, \eta_p^2 = .01\). There was also no discernible systematic effect of the Order in which the two conditions followed each other, \(F(1.96, 186) = 0.15, p = .86, \eta_p^2 = .002\). The covariate, Age, was not significantly related to the emotional ratings \(F(1.96, 186) = 2.17, p = .12, \eta_p^2 = .02\).

The reason for this lack of significance lies in the fact that age was treated as a linear predictor. Since the 3-year-old children gave lower ratings on the dependent question than their older peers (see Figure 2) the linear trend from younger to older is reduced and fails to register as significant.
When concentrating on the three oldest age groups, where ratings declined linearly, the same analysis resulted in a significant interaction between Condition and Age, $F(1.73, 102) = 3.64, p = .036, \eta^2_p = .07$.

Figure 2 presents children's mean ratings in response to the three different questions for each age group (for reasons of comparison we included the adult's response pattern from Experiment 2). A paired-samples $t$-test for the mean of the Baseline Question and the mean of the Independent Test Question revealed a significant difference for the two older age-groups, 9;0- to 11;1-year-olds: $t(18) = 2.42, p = .03$, Cohen's $d = 0.55$, 12;4- to 14;4-year-olds: $t(15) = 6.36, p < .001$, Cohen's $d = 1.59$, but not for the three younger age-groups. A similar picture emerged for the comparison of the Baseline Question with the Dependent Test Question. There was a significant difference in the 9;0- to 11;1-year-olds: $t(18) = 3.20, p = .005$, Cohen's $d = 0.73$, and in the 12;4- to 14;4-year-olds: $t(15) = 3.55, p = .003$, Cohen's $d = 0.89$. Unlike for the Independent Test Question this time also the youngest age group showed a significant difference, $t(17) = 2.15, p = .046$, Cohen's $d = 0.51$. This finding that 4-year-olds but not anymore 5½-year-olds show a difference remarkably parallels what was found in Experiment 1 with 4½- and 5½-year-olds. This strongly suggests that it is the double-questioning that causes a shift in response in the youngest children, which—due to the high baseline—tends to be lower rather than higher, i.e., a “false positive” if interpreted as a sign of regret.

The analysis so far suggests that the three youngest groups showed no reliable sign of regret, since there was no significant difference between baseline ratings and ratings in response to the Independent Test Question. However, this tempting conclusion may be premature. For, in all three groups, as Figure 2 shows, the mean ratings in response to the Independent Test Question were always lower than the baseline ratings. In fact, comparing the joint means for the three youngest age-groups a small (Cohen's $d = 0.32$) but significant, $t(62) = 2.54, p = .013$, difference was found between Baseline Question and Independent Test Question. This suggests either that all children showed regret but to a much smaller degree than the older children and adults or that only some children showed regret at adult levels. To find out we look at the distribution of differences.

Figure 3 shows the distribution of the difference between the Baseline Question and the Independent Test Question for the three younger and the two older age groups. When a child rated the Baseline Question with 54 and the Independent Question with 50, then the computed
difference was +4 (note that the scale ranged from 0 to 64 and that children were not shown their prior rating on the Baseline Question—so a difference of 4 could be interpreted as the intention to give the same rating for both questions). Within the range of ±7 scores are symmetric around zero in the three younger age groups but not in the two older age groups. It is assumed that symmetry around zero can be interpreted as that these children did not intend to make a difference in their rating for the Baseline Question and for the Independent Test Question but their ratings just varied at random (group “±7”). If children’s rating on the Independent Test Question was higher than the one on the Baseline Question, they were grouped as “<−7” because this resulted in a negative value, e.g., when a child rated the one candy with 50 and after seeing that he/she could have got three candies, he/she rated the one candy with 58. These children might have interpreted the Test Questions as asking about how they would feel with having three candies. Children whose rating resulted in a positive value (the rating of the Independent Test Question was lower than the one of the Baseline Question) belonged to the group “> +7”.

Figure 4 shows how many children of each age group gave a similar rating (±7) of the Independent Test Question and the Baseline Question, how many gave a higher rating (<−7) and how many a lower rating (> +7). Fifty-four percent (n=19) of the older age groups but just 25% (n=16) of the younger age groups were unhap-

Figure 3. Distribution of the differences between baseline question and independent test question shown separately for the three younger versus the two older age groups (negative value: independent test question has been rated better than baseline question).
pier after they had seen what they could have got (Independent Test Question) than before (Baseline Question). Sixty-five percent \( (n = 41) \) of the 3- to 8-year-olds and 43% \( (n = 15) \) of the 9- to 14-year-olds stayed with their first rating. Taken together, only a minority of the younger children was less happy on the Independent Test Question than on the Baseline Question, most of them did not change their ratings. The picture reverses for the older children.

**Discussion**

Experiment 3 was designed to test (1) when children start to show regret in the independent condition and (2) whether the significant difference between Baseline Question and Dependent Test Question found in the younger children of Experiment 1 was due to double-questioning or to explicitly pointing out what they could have got. Results show that a significant difference between the Baseline Question and the Independent Test Question can first be found in 9- to 11-year-old children. However, 25% of the children in the younger age-groups gave lower ratings in response to the Independent Test Question than the Baseline Question. This is similar to Experiment 1, where 29% showed the same pattern.

Another issue that arose from Experiment 1 was whether younger children are responsive to double-questioning. This assumption could be confirmed in Experiment 3. Here, the Dependent Test Question was the same as the Independent Test Question. Children of the youngest group showed a significant difference between the Baseline Question and the Dependent Test Question but not between the Baseline Question and the Independent Test Question. This means that the difference that has been found between Dependent and Independent Test Questions in the youngest children of Experiment 1 is not due to the fact that the Dependent Test Question was formulated in the subjunctive and therefore giving children a hint to reason counterfactually. Instead, this effect might be due to the fact that the younger children especially felt compelled to change their ratings because of being asked a second time. Weisberg and Beck (2010) also might have found a difference between the initial question, “How do you feel after opening your box?” and the alternative question, “How do you feel about your box now?” in 5-year-old children because of asking them twice.

The final experiment was intended to deal with a more general problem, that even younger children might feel regret but are not able to use the scale properly (either because they misunderstand the use of the scale or they just cannot use the scale to show how they feel).

**EXPERIMENT 4**

The objective assessment of emotions is difficult because they cannot be measured directly. Therefore, one major concern about all these experiments is the face validity of the rating scale. The scale is supposed to measure how happy children are with what they got. It is assumed that children experience regret whenever they are not as happy as they were before after seeing what they could have got. In Experiment 3 we found a linear trend of age from 8 years on. One worry is that the 8-year-old children did not understand how to use the scale. In Experiment 4 we therefore tested 6- to 8-year-old children in order to show that these children are actually able to use the scale to convey how they feel outside the context of counterfactual emotions but, as in Experiment 3, will not show clear signs of regret in their ratings.

**Method**

**Participants.** Participants were 17 children (9 males and 8 females) from an after-school care club in the city of Salzburg. Ages ranged from 6;0 (years; months) to 8;9 and the mean age of children participating was 7;6 years \( (SD = 8 \text{ months}) \).

**Materials.** We used the same materials as in Experiment 3.

**Design.** Each child was tested in one session that lasted about 10 minutes. Every child was given either one candy (less attractive item) and a very
beautiful sticker (attractive item—for a description see materials section of Experiment 3) or three candies (attractive item) and a not very beautiful sticker (less attractive item). The order was fully counterbalanced. For each item they were asked (1) how they felt about it and (2) how they would feel if they could get the other item. Therefore, each child had to give four ratings in total (two about what they really got—one attractive and one less attractive item—and two about what they did not get—one attractive and one less attractive item).

Procedure. The introduction phase was the same as in Experiment 3. In the test phase the child selected one of two boxes, and depending on the condition, it got an attractive or a less attractive item. For instance if the child only received one candy the child was told: “Oh, there is one candy (relatively unattractive item). It’s yours now”, and the Baseline Question was asked: “How happy are you with your one candy? Can you show me on our scale?” After the child had shown it on the scale, it was told: “Assuming, that you would have got three candies (showing the child the attractive reward), how happy would you be? Can you show me on our scale?”

Results

In general, if children got the attractive reward they were significantly happier \( (M = 57) \) than when they got the less attractive item \( (M = 51) \), \( t(16) = 3.00, p = .008, \) Cohen’s \( d = 0.64 \). Even in their imagination children were much happier with the attractive item \( (M = 57) \)—after having received a less attractive item—than with the relatively unattractive item \( (M = 43) \)—after having received an attractive item, \( t(16) = 4.13, p = .001, \) Cohen’s \( d = 1.11 \). This means that children evaluated our items differently and that they were able to show this by using our scale.

GENERAL DISCUSSION

The aim of the present paper was to find out at which age children start to feel regret. So far, the developmental literature on regret suggests that children do not make a difference in their ratings—before and after they have seen of what they could have got—before the age of 5 years (Weisberg & Beck, 2010) or even later (Amsel & Smalley, 2000; Amsel et al., 2003), that they do not understand regret in others before the age of 7 years (Beck & Crilly, 2009; Guttentag & Ferrell, 2004; Weisberg & Beck, 2010) and that they can not anticipate regret before the age of 9 years (Guttentag & Ferrell, 2008). However, some studies on counterfactual reasoning suggest that children can think of a counterfactual outcome much earlier than 6 years (German & Nichols, 2003; Harris et al., 1996; Müller, Miller, Michalczyk, & Karapinka, 2007; Perner, Sprung, & Steinkogler, 2004; Riggs, Peterson, Robinson, & Mitchell, 1998). But other studies do not find counterfactual reasoning in children before 6 years of age (Beck, Robinson, Carroll, & Apperly, 2006, open counterfactuals; Rafetseder et al., 2010) implying that it must develop after 6 years of age.

One reason for this discrepancy could be that the tasks used to elicit regret in children were too complicated, i.e., that the indirectness of the test questions simply suppressed or diluted children’s feeling of regret. So, a procedure that assesses regret much more directly should find regret earlier. However, in Experiment 1 we found that up to 6 years children did not show regret, even when they were asked about their feelings immediately after they had seen that they could have got more. Adults in contrast (Experiment 2) clearly regretted their decisions in a similar task. Finding no signs of regret in children younger than 6 years might therefore not be due to the task procedure. Moreover, in Experiment 3 children did not show a significant difference between the Baseline Question and the Independent Test Question before the age of 9 years. However, roughly 30% of the children younger than 9 years were less happy on the Independent Test Question than on the Baseline Question and when all three younger age groups were taken together then this small difference was also significant. This developmental trend is amenable to several explanations.
First, under the assumption that our task does assess regret it could be either that only some young children show regret or that all of them show regret but to a much smaller degree than adults. The results of Experiment 3, however, do not support this explanation. Although the mean ratings for the Independent Test Question were lower than the mean ratings for the Baseline Question (even in the younger age groups), more than 50% of the younger children gave similar ratings (±7) for the Independent Test Question and the Baseline Question, i.e., the differences varied symmetrically around zero suggesting that they had no intention to differ. From 9 years on the difference increases markedly. If all children—even the younger ones—showed regret then it would be mysterious that younger children are not as strongly affected by the knowledge of the alternative outcome as the older children and adults.

A parallel finding supports the idea that the ability to anticipate regret does not develop before around 9–10 years of age. In the study by Guttentag and Ferrell (2008) children were shown that they won a mediocre prize, e.g., a pencil, and knew that the box not chosen had contained either nothing or a much better prize, e.g., a stuffed animal. They were asked what they hoped was in the box not chosen. While the 9- to 10-year-olds and adults said they hoped that the alternative was less attractive (no prize at all) than what they got, the younger children (5- to 8-year-olds) wanted to find a bigger prize in the other container. They also were told about a puppet playing the same game who hoped to find the other box empty and they were asked to explain the puppet’s hope. Their explanations followed a similar age trend.

It could also be that children just don’t make these comparisons spontaneously although they are actually competent at counterfactual reasoning (Beck & Crilly, 2009; Guttentag & Ferrell, 2004). But, we think that this is not very likely, at least for the younger children, as it did not make any difference to them whether they were asked explicitly about what they got in Experiment 1 (“Now that you know that you could have got five candies, how happy are you with your one candy?”) or less explicitly in Experiment 3 (“How happy are you with your one candy?”).

However, we should not disregard the fact that about 30% of the younger children gave clearly lower ratings in the independent than the baseline condition. This result might, however, be misleading. While 24 children of the younger age groups in Experiment 3 rated the independent question lower than the baseline questions (also counting children who only went down by one, e.g., baseline: 53, independent 52), almost as many, 17 children, rated the independent question higher than the baseline question, which is not a significant difference, \( \chi^2(1, N=41) = 1.2, p = .27 \), and surely does not suggest a consistent effect of regret. So, what could be the reason for this small but significant difference? We can think of three possibilities.

Pointing out to children that there were also three candies on offer in the other box, may have different effects. For some this may shed a positive halo on the whole situation leading to even more positive ratings while for others it provides an unfavourable comparison anchor for what they got, “only” one candy. As a result the variance of ratings is larger than in the baseline condition. Since baseline ratings tend to be fairly close to ceiling (53 of 64) the increased variance tends to spread downwards where there are 52 units left to zero. As a result the mean ratings tend to be lower even though almost as many children give higher ratings as give lower ratings.

A slightly different explanation for the small difference observed for the youngest three age groups is that our task does not assess regret but frustration. It might not be necessary to reason counterfactually to show frustration. Frustration can be construed in two different ways. Children can be disappointed with what they got because they could have got more if the world had been more benevolent (Guttentag & Ferrell, 2008; Zeelenberg & Van Dijk, 2005). This would be based on counterfactual reasoning and is difficult to distinguish from regret. However, they can also feel frustrated, simply because they did not get the three candies in the other box.
This will also lower their ratings on the independent test question as a result of a confluence of two emotions: happiness about being given one candy mixed with the dissatisfaction of not getting three candies. This mix results in less satisfaction than having got one candy reflected in the baseline rating.

This could also explain the data in Amsel and Smalley’s (2000) study. When children were asked: “How would you have felt if you had turned over the unchosen card?”, it was not regret that was measured but how frustrated the children were about not winning. But when they were asked: “Now that you’ve seen the other card, how do you feel about the card you selected?”, they have to compare the consequences of their chosen card with the consequences of the not-chosen card. However, the explanation that our task just measures frustration is not very satisfying. For it begs the question of why younger children are not as strongly frustrated as are older children and adults.

The third explanation assumes that our task is sensitive to frustration and regret simultaneously. This could explain why the difference increases from 9 years on. Children up to about 9 years only show frustration, which results in the rather small difference between the Independent Test Question and the Baseline Question (see Figure 2). From around 9 years onwards more and more children start to regret not having chosen the other box, which enlarges the difference. Given that only regret but not frustration is an emotion that must be based on counterfactual reasoning, the data suggest that children younger than 9 years did not reason counterfactually in our task. How can this be brought into line with studies in which children show counterfactual reasoning before the age of six?

Beck et al. (2006), Beck, Riggs, and Gorniak (2010), Perner (2000), Perner and Rafetseder (2011), and Rafetseder et al. (2010) all claim that correct answers to counterfactual questions might not always be based on counterfactual reasoning but instead might be inferred from generalised knowledge. For example, children who were asked whether the floor would be dirty or clean if Carol had taken her shoes off (Harris et al., 1996) could answer the question by using general knowledge that whenever somebody takes shoes off, floors tend to stay clean. Beck et al. (2009) claimed that if it is possible for children to infer correct answers from general knowledge, e.g., “What if Nicolas’s friend hadn’t fallen into the rosebush. Would Nicolas be happy or sad?”, then it should be easier for them to answer the counterfactual question compared to when the correct answer can not be inferred from general knowledge, e.g., “What if the dog had not stolen the spade. Would the spade be in the sandpit or in the pond?” They could not find a difference in difficulty. However, although the correct answer to the latter counterfactual question can not be inferred from general knowledge, it does not require children to reason counterfactually. There are two possible locations where the spade can be, the sandpit and the pond. Children, who understand that the question asks them for a location where the spade is not at the moment, will arrive at the correct answer, the sandpit, simply by excluding the pond as a possible answer.

This means that we should use tasks in which counterfactual reasoning can be clearly differentiated from other kinds of reasoning such as basic conditional reasoning in order to avoid false positives (Perner & Rafetseder, 2011). When controlling for false positives 6-year-old children do not show counterfactual reasoning (Rafetseder et al., 2010). Pilot data with the same procedure point to a much later onset at around 11 years (Rafetseder & Perner, 2010). Using an even easier procedure, Schwitalla (2010) found no signs of counterfactual reasoning before 10 years. The study used two different conditions. In condition 1, children were shown a puppet (Susi) walking with her dirty shoes over the clean floor and were then asked: “If Susi had taken her shoes off, would the floor be dirty or clean?” The same story has been used in Harris et al. (1996) and Perner and Rafetseder (2011) are claiming that basic conditional reasoning (“Whenever somebody takes shoes off, floors tend to stay clean”) and counterfactual reasoning (“If Susi had taken her shoes off, then the floor would have stayed clean”) arrive at
the same correct answer “clean”. In condition 2 not only Susi but also Max were walking over the clean floor with their dirty shoes and children were asked the same question as in condition 1. While this time basic conditional reasoning leads to the wrong answer “clean”, counterfactual reasoning arrives at the correct answer “dirty”. Schwitalla (2010) found that in condition 1 the amount of correct answers was with 93% almost at ceiling for the 5-year-olds, while in condition 2 it dropped to 18%. It was not before the age of 10 years that more than 50% of the children showed systematically correct answers.

This suggests that the younger children especially did not apply counterfactual reasoning in our regret experiments. Instead they might have used the mixed strategy of being happy with one candy and being sad about not having three candies. But even for the older children we can not be sure whether they really show regret. It could be that they go down in their ratings more extremely but this could just be because they are more affected by not getting three candies.

One idea, to find out whether children were just frustrated in our regret task, is to study relief, which is closely related to regret. One feels relieved whenever a possible alternative decision could have resulted in a worse outcome than the actual outcome (Guttentag & Ferrell, 2008). In our regret study, one reason for children being not as happy on the Independent Test Question as on the Baseline Question could have been that children were just frustrated about not having three candies. So, these children did not show regret (as a result of reasoning counterfactually) but frustration. In a relief task where children win three candies and not having one candy would not result in being happier. If anything, children might be unhappy about not additionally having the one candy. A mix of being happy about having three candies and being unhappy about not having one candy should not result in a higher rating.

Another very important point is that Guttentag and Ferrell (2004) reported that 7-year-old children can understand when other people feel regret. The question is now how this fits with our claim that children don’t experience regret until around 9 years of age. We can but speculate. Guttentag and Ferrell used different stories, such as the one about Bob and David who both rode their bikes to school each morning (p. 773). Every day, Bob took the way to school that went around the right side of a pond. David always took the path that went around the left side of the pond. However, David decided one day to go on the right side. Unfortunately, that day a tree branch fell across the path that went around the right side of the pond. Bob and David hit the branch with their bikes and they fell off their bikes and were hurt. Children were asked the following question: “Who would be more upset about deciding to ride along the path that went around the right side of the pond that day? Bob, who rides on the path around the right side of the pond every day, or David, who usually rides around the pond on the left side but decided to ride around on the right side that day, or do you think they would feel the same?” From 7 years on children said that David felt worse than Bob.

One explanation for why children this young make a difference in the expected direction could hinge on the question being asked. The question was “Who is more upset about deciding …?” Since it is only really David who “decided” anything, Bob just did what he usually does, it is David who must be more upset about deciding to go around the right side.

However, Guttentag and Ferrell (2004) also used a story about commission/omission that might not be susceptible to the same criticism. Karen and Michelle did well on a test and that’s why their teacher gave them a choice between two
boxes and warned them that one prize was a lot better than the other prize. Both picked a box but were then asked by their teacher whether they really wanted to stick with their choice. While Michelle stuck with her choice, Karen switched to the other. Both girls ended up with the lesser prize of a pack of balloons. The test question was who felt worse about picking that box. Seven-year-olds systematically chose the girl who had switched.

A possible further alternative explanation for this effect—that applies to the unusual route as well as to the omission/commission scenario—could be that the effect is due to frustration or a mix of various emotions but not strictly speaking regret. David wanted to do something unusual one day, ride his bike on the right side of the pond. We don’t know why he decided to do so, but he likely had some good reason to do so, i.e., he must have expected some positive incentive. Hence the accident caused by the fallen tree not only had the upsetting consequences it had for Bob, it also deprived David (but not Bob) of the anticipated positive experience that made him decide to venture around the pond on an unusual route. Because of this additional frustration it is perfectly reasonable to say that David might feel worse than Bob without having to engage in counterfactual reasoning.

The same argument might also hold for the omission/commission scenario. Michelle and Karen make their choices but are then asked by the teacher whether they want to switch. Given that the teacher told them that one prize was much better than the other, children might assume that the teacher knows where the bigger prize is and with her question is trying to make them switch. Against this background Karen, who followed her teacher’s suggestion, would not only feel frustrated with her suboptimal prize but also angry with her teacher for setting her up. In contrast, Michelle, who resisted her teacher’s suggestion, will be frustrated about her outcome but will not be angry with her teacher, if anything she might be grateful to her teacher for trying to help her. So, in sum, Karen would be more upset (frustration + anger) than Michelle (frustration only).

Interestingly, this analysis does not apply very well to the second omission/commission scenario used by Guttentag and Ferrell (2004). Bill and Tim play baseball. While Tim decides to switch from the Lions to the Hawks, Bill decides to stay with the Hawks (and not to switch to the Lions). Unfortunately, the Hawks loose every game and the Lions appear to be the best team of the year. In contrast to the other scenario there is nobody who tried to persuade them to change and thus no one to be angry with. Only counterfactual reasoning about the boy’s emotions would lead to judging Tim as more upset than Bill. In line with our argument that 7-year-olds do not engage in counterfactual reasoning 7-year-old children showed only a minimal and much smaller rating difference in this scenario than in the Karen and Michelle story.

**Conclusion**

To conclude, we found that children before the age of 9 years did not show regret in our tasks even though we used a much more direct measurement than previous studies. We suggest that children younger than 9 years, who rate their happiness lower after they have seen what they could have got, are just frustrated about not having got those candies as well. Furthermore, some of the task procedures used in existing studies, e.g., double-questioning, might have produced false positive signs of regret especially in the youngest children. Our results strengthen the hypothesis that children under 9 to10 years do not reason counterfactually. It is possible that children, who are younger than that and who answer counterfactual questions correctly, might arrive at the correct answer without reasoning counterfactually.

---

1We thank an anonymous reviewer for pointing this out to us.
REFERENCES


philosophy and psychology (pp. 90–110). Oxford, UK: Oxford University Press.


