“The game would have been better for me if … ”: children’s counterfactual thinking about their own performance in a game

Marta Stragà, Angela Faiella, Ingrid Santini and Donatella Ferrante

Department of Life Sciences, University of Trieste, Trieste, Italy

ABSTRACT

In two studies, we investigated for the first time the content of children’s counterfactual thoughts about their own experiences. Results showed that the majority of children aged 8-13 were able to produce valid counterfactuals regarding an event that happened to them, despite not achieving an adult-level ability. Comparing counterfactual and prefactual thinking, in Study 1 we found that children showed the same temporal asymmetry previously found in adults: They focused on the controllable features of their experience more in prefactual than counterfactual thinking. However, in Study 2, comparing counterfactuals produced by children and adults after a task in which making errors became salient, children produced more controllable counterfactuals (modifying their own errors) than adults, who still focused on uncontrollable features (as in Study 1). These results suggest that the ability to reason counterfactually in complex and real-life situations is not yet fully developed at age 8-13 years, affecting counterfactual content.

ARTICLE HISTORY Received 25 February 2022
Accepted 23 September 2022

KEYWORDS Counterfactual; prefactual; children; controllability

Introduction

The mental simulation of past and future scenarios allows individuals to understand the past, make predictions about the future, plan and regulate their behavior. A great deal of research has focused on how people imagine alternatives to past events (i.e. counterfactual thinking, for a review, Byrne, 2016). The ability to simulate what might have been involves the capacity to hold in mind both a representation of what is true, and a second representation of what might have been, but did not happen (e.g. Byrne, 2005).

Moreover, individuals need to understand that a past event is causally linked to the observed outcome and that a change in the past event would have led to a different outcome.[1] The development of counterfactual reasoning in children has been thoroughly investigated, but at what point in their development are children able to reason counterfactually in an adult-like manner is still subject to debate. Whereas some studies found evidence of such a capacity in 3-4 year old children (Harris et al., 1996; Riggs et al., 1998), more recent studies suggest that the tasks previously used did not involve the capacity to hold in mind multiple possibilities (Beck et al., 2006), and could reflect a more basic ability to reason about conditionals (Rafetseder et al., 2010; Rafetseder et al., 2013; Rafetseder & Perner, 2010; for a review, Rafetseder & Perner, 2014). Using different tasks, the emergence of adult-like counterfactual reasoning has been observed at a later age. Beck et al. (2006) showed that six-year old children, but not younger children, are able to acknowledge two possible scenarios when required to think about a single causal event that could bring about both the actual and the counterfactual outcomes. Other studies showed that even 5-6 year old children found it difficult to provide the correct answer to a counterfactual question which required them to imagine an alternative world in which the antecedent is true and only its causal consequences are modified accordingly, whereas all the rest remains the same (i.e. to take into account the nearest possible world constraint; Rafetseder et al., 2010). Moreover, Rafetseder and Perner (2014; see also Rafetseder et al., 2010; Rafetseder, et al., 2013; Rafetseder & Perner, 2010) claim that previous evidence indicated the use of basic conditional reasoning and domain-specific knowledge (i.e. general knowledge of causal regularities), that suffice to pass the test, but were not indicative of fully-fledged counterfactual reasoning. According to these authors, mature reasoning with counterfactuals (Rafetseder et al., 2021) can be observed in early adolescence (i.e. 12 years; Rafetseder et al., 2013) and evidence of such reasoning can be found in a minority of 6-year olds (Rafetseder & Perner, 2014). Nevertheless, employing less cognitively demanding scenarios, other studies showed that children were able to reason counterfactually in a task in which the application of basic conditional reasoning with counterfactuals would have led to incorrect answers at 6-7 years of age (McCormack et al., 2018; Nyhout et al., 2019) or even before if the causal structure is clear and involved the physical domain as opposed to human behavior (Nyhout & Ganea, 2019a). However, Beck and Rafetseder (2019) questioned this conclusion claiming that in the task employed by Nyhout and Ganea (2019a) updating the

[1]Here we focus on counterfactuals in which the mutation of the antecedent undoes the outcome, as most of the counterfactual research does, even if the term also includes the cases in which a change in the antecedent does not change the factual outcome (i.e. semifactuals; Goodman, 1983).

actual outcome is sufficient to provide a correct answer (thus involving basic reasoning with counterfactuals), whereas the former tasks (McCormack et al., 2018; Nyhout et al., 2019) require a simulation of a new version of the world which is entirely in the past.

In order to assess children's ability to engage in counterfactual reasoning, various tasks were used typically involving scenario-based paradigms (see Beck, 2020 for a review). After being presented with a series of events (usually acted out with puppets or portrayed with images), counterfactual conditional tasks require participants to imagine how the world would be in the present if an antecedent was false ("consequent task"; Beck et al., 2009; German & Nichols, 2003; Guajardo & Turley-Ames, 2004; Harris et al., 1996; Perner et al., 2004; Riggs et al., 1998) or to imagine how the outcome would have been different, generating different antecedents ("antecedent tasks", e.g. Harris et al., 1996; Guajardo & Turley-Ames, 2004). Guajardo and Turley-Ames (2004) showed that the performance on consequent and antecedent tasks are highly correlated and can be passed by young children. Open counterfactual tasks in which children are asked to answer an open question like "what else could have happened", are generally more difficult and they are successfully solved at a later age (Beck et al., 2006). Other studies used counterfactual syllogisms in which participants were presented with premises that they knew to be false (e.g. "all cats bark") and they were asked to derive the valid conclusion (e.g. Beck et al., 2009; Dias & Harris, 1988; Leevers & Harris, 2000). Recently, Guajardo et al. (2016) assessed the spontaneous generation of counterfactuals in 8 to 11 year old children while retelling a previously heard story. Authors tested whether outcome valence and outcome expectancy affected the counterfactuals that children generated, as previously found in adults. They found evidence of spontaneous counterfactual thinking in middle childhood and, consistent with data on adults, a greater tendency to engage in counterfactual thinking when the outcome was unfavorable for the protagonist of the story rather than favorable. All the above-mentioned developmental studies share a common feature: children reflected on events that were experienced by a third party and not by themselves.

Although scenario-based paradigms have been widely employed in counterfactual thinking research on adults, some studies have shown that the content of such thoughts can be very different when participants reflect on their actual failure instead of reading about the failure experienced by the protagonist of a story (Ferrante et al., 2013, Study 2; Girotto et al., 2007; Pighin et al., 2011). In particular, these studies showed that while readers of a story undid, in their counterfactual thoughts, a controllable event (i.e. a choice) which prevented the protagonist from having a good outcome regarding the task, actors who actually experienced the same failure focused on other events. This distinction does not depend on motivational factors (the effect occurred even when actors did not make blameworthy choices or another person made choices for them; Girotto et al., 2007), but likely derives from the differing availability of information with regard to the problem-solving phase. The personal experience of the actors may lead to a mental representation of the situation which is far more complex than the representation built on the small amount of information presented in the story, and the number of alternative antecedents that actors are able to come up with are many more than those available in the scenario-based paradigms which are usually limited to the events described. As far as we are aware, there has yet to be a study which investigates directly the generation of counterfactual thoughts about children's personal experiences, and given the actor-readers effect and the greater complexity that individuals face when reflecting on their own experience, the question whether children are able to generate counterfactual thoughts about their personal experience to a similar extent than adults is still open. Some insights on children's ability to reflect counterfactually about their personal experiences come from regret studies, in which the capacity to experience regret is inferred by comparing happiness ratings when children know vs. do not know that the option which they did not take would have resulted in a better outcome. If children report being sadder after knowing that they would have obtained a better outcome if they had chosen differently, they are assumed to be experiencing regret. Research found evidence of the emergence of such a feeling around 6-years of age (Weisberg & Beck, 2010; 2012), or even later (around 9 years of age, Rafetseder & Perner, 2012). Nevertheless, it is still under debate whether such studies actually measure regret (e.g. McCormack et al., 2016; O'Connor et al., 2014; Rafetseder & Perner, 2012; see also, Rafetseder & Perner, 2014). Moreover, and more importantly for our aims, none of these studies have directly assessed the content of counterfactual thoughts of children who felt such emotion.

Given the lack of studies in which children are asked to reflect on their personal experience and the few attempts aimed at identifying possible differences with respect to adults' counterfactual thinking (see, Guajardo et al. 2016; Payir & Guttentag, 2019), the main goal of the present paper is to test not only whether children in middle and late childhood are able to generate counterfactual thoughts about an event which happened to them, but also whether the content of such thoughts is similar or different in comparison with adults.

Following literature on counterfactual thinking in adults that mainly focuses on how people imagine a better past (i.e. upward counterfactual) and showed that such counterfactuals were more frequently generated with respect to imagining a worse past (i.e. downward counterfactual; e.g.

Petrocelli et al., 2012; Rim & Summerville 2014), we applied the same paradigm used with adults (e.g. Ferrante et al., 2013) in two studies in which after engaging in a task, participants were asked to complete a counterfactual sentence filling it in with an antecedent (i.e. complete the sentence "things would have been better for me if … "). In such a paradigm in which participants reflect on a personally experienced event, much richer information may be available (not only the information explicated in the scenario; see Girotto et al., 2007) and, as a consequence, participants may come up with a much broader range of possible antecedents compared to the tasks previously used with children in which they were limited to what is explicitly mentioned in the scenario or set up by the experimenter. This aspect, along with the experimental instructions, could make the generation of counterfactuals more difficult. Language skills which play a significant role in the ability to reason counterfactually (Beck et al., 2009), may be even more involved than in previous studies, and young children may have difficulty expressing their thoughts. For these reasons, we decided to use an older sample with respect to previous studies, recruiting children from middle to late childhood (on average 10 years of age), who should be able to pass standard tasks, while they may not fully possess the ability to construct complex counterfactual simulations (see, Rafetseder & Perner, 2014). We thus expected that the majority of children would be able to come up with a counterfactual thought but given the above-mentioned considerations and under these conditions, such an ability may not have been fully developed and possessed by all the participants. As regards the content, counterfactual literature based on adults showed that even if participants could virtually change an infinite number of antecedents that would undo the outcome (from their own actions to the laws of the universe), they tend to focus on a narrower range of possible features (e.g. changes in the laws of physics are very uncommon) when imagining almost the same types of counterfactuals (see Byrne, 2005, 2016). Developmental literature, however, mainly focuses on the emergence of counterfactual thinking and only a few studies investigate the presence of such fault lines in children, and most of these do so indirectly (Guttentag & Ferrell 2004; Meehan & Byrne, 2005; Nyhout & Ganea, 2020; Payir & Guttentag, 2016, 2019), finding that some biases may emerge late in childhood. The question as to whether children of our age range would focus on the same content as adults is thus still open.

Further examining the content of adult thinking, recent studies found a temporal asymmetry comparing counterfactual and prefactual thoughts. Thinking about how a previous performance could have been better resulted in less controllable modifications than thinking about how it will be better in the next attempt (Ferrante et al., 2013; Hammell & Chan, 2016; Mercier et al., 2017). These results are inconsistent with evidence collected via scenario-based paradigms in which participants tended to modify controllable features of the situation in their counterfactual thoughts (e.g. Girotto et al., 1991) and with the general assumption that counterfactual and prefactual thoughts should share the same function and thus the same content (Epstude et al., 2016; see also Epstude & Roese, 2008). Therefore, with the aim of comparing the content of thoughts between children and adults, in Study 1 we also compared counterfactual thinking with prefactual thinking.

Recent studies focusing on the development of temporal asymmetries found evidence of temporal emotion asymmetry (i.e. greater affect when thinking about the future than when thinking about the past) and of temporal distance asymmetry (i.e. future events feel closer than equidistant past events) in children aged 6-7 years, whereas evidence of temporal value asymmetry (i.e. future experiences are more valued than past experiences) was found at a later age, from age 9 to 10 years (Burns et al., 2019). As a consequence, even if children were able to reason counterfactually and prefactually, they might not focus on the same content as adults. Therefore, also the question as to whether children aged 8-12 years would show (as adults do) the previously-mentioned temporal asymmetry with regard to thought content is open.

Moreover, comparing counterfactual and prefactual thinking, an asymmetry in the ability to generate such thoughts may arise. Developmental literature suggests that future hypothetical thoughts are easier for children than counterfactual thoughts (Beck et al., 2006; Perner et al., 2004; Rafetseder et al., 2010; Riggs et al., 1998; Robinson & Beck, 2000). Counterfactual thinking requires that two mental representations are active at the same time (the real one and the alternative one, contrary to facts), whereas in future hypothetical thinking there is no need to negate or consider what it is known to be true and only one mental representation is sufficient (e.g. Beck et al., 2006; Rafetseder & Perner, 2010). Nevertheless, in the previously-mentioned studies with adults (Ferrante et al., 2013, Mercier et al., 2017), prefactual thoughts are explicitly linked to reality: participants were asked to think about how things will be better in the future with respect to what actually happened a few minutes prior. In such a task, the hypothetical future scenario must be different from the actual one, so two representations may need to be simultaneously active (the real and the alternative future). Thus, prefactual thinking in such a task may not be easier than counterfactual thinking, given that only the temporal prospective changes.

To summarize, we expected that the majority of the children would be able to generate counterfactual thoughts about their own experience, but not reaching the adult level. We also expected that prefactual thinking in our task may not be easier than counterfactual thinking. Finally, it is still open as to whether children are constrained by fault lines in their thoughts as adults are and if so, whether such fault lines are similar or different in comparison to adults.

Study 1

In Study 1 we compared counterfactual and prefactual thinking of children and adults when reflecting on their own performance. In particular, we investigated whether children were able to generate counterfactual and prefactual thoughts and whether, in doing so, they focused on the same content as adults, showing the previously-mentioned temporal asymmetry in the content (Ferrante et al., 2013; Mercier et al., 2017).

The data from Study 1 were collected as part of a larger longitudinal project that aimed to investigate how cognitive and social skills could be affected by structured physical activities. Given that measures of working memory capacity and selective attention capacity were collected for the project, a secondary aim was to assess whether individual differences in working memory and selective attention are related to the capacity to generate counterfactuals and prefactuals about personal experiences. There is evidence that specific cognitive abilities are needed to engage in counterfactual thinking. In particular, studies focused on executive functions that can affect the emergence of counterfactual thinking. There is compelling evidence that inhibitory control plays a role in this process (e.g. Beck et al., 2009; 2011) as well as cognitive flexibility (Burns et al., 2012), whereas evidence regarding working memory is not conclusive (e.g. Drayton et al., 2011; Guajardo et al., 2009, found an effect; Beck et al., 2009, did not find an effect). The number and complexity of the possible worlds (Byrne, 2016) and task-related information one has to keep in mind in order to answer correctly justify the relationship between working memory and the ability to reason counterfactually. Such a relationship has been found early in development (with 3-4 year olds; Drayton et al., 2011; Guajardo et al., 2009), while it has been suggested that it may not explain the difficulties faced by older children (Rafetseder et al., 2010, Rafetseder et al., 2013). Our study involved an older sample, so we might not expect a marked influence of working memory. However, as previously stated, we expected that a part of our sample would not be able to generate a counterfactual thought about a just-experienced event, therefore we wondered whether this failure may be partially accounted for by differences in working memory.

Method

Participants

One hundred and ninety-nine children and pre-adolescents (99 male and 100 female, age: $M = 10.11$, $SD = 1.40$, ages range from 8 to 13 years) participated in the study. One hundred and thirty-four children came from primary school (age: $M = 9.33$, $SD = 0.83$, ages range from 8 to 11) and sixty-five came from middle school (age: $M = 11.71$, $SD = 0.88$, ages range from 10 to 13). Participants were part of a large project in which children and adolescents received free sport lessons after school. The project aimed to investigate the possible cognitive and social benefits of sport activities. During the data collection related to the project, participants took part in the present study, and all of whom had the written consent of their parents/guardian to participate in the study. Participants came from diverse socio-economic backgrounds and different neighborhoods of a mid-sized town in Northern Italy.

Moreover, eighty-six undergraduate university students (43 male and 43 female; age: $M = 22.99$, $SD = 2.56$, ages range from 19 to 32) were recruited in the same town to take part in the study. The adult sample size was decided on using a power analysis assuming a medium effect size in a 2×2 chi-square test (power = .80) and considering that we expected no invalid thoughts in the adult sample (based on previous studies). Nevertheless, in the children's sample, we expected more variability in the responses, some invalid thoughts and some restatements (see the coding section), so we increased the sample size accordingly.

The procedure was approved by the Ethical Committee of the University of Trieste (report number 116).

Materials and procedure

Children were tested individually in a quiet area in the after-school clubs involved in the project or in their classrooms. They completed several measures in a random order and herein we report only the relevant measures for this study[2].

We assessed selective attention capacity using the selective visual attention sub-test of the Italian Battery for neuropsychological assessment (BVN 5-11, Bisiacchi et al., 2005). Participants were asked to find the target figure among a set of 64 figures (geometrical shapes arranged in eight rows and eight columns) within a one minute period. The target figure was present 12 times, resulting in a maximum score of 12. Before completing the task, we checked to see that the instructions were fully comprehended using a practice trial.

Working memory was assessed using a computerized version of the Corsi Block-tapping test (Kessels et al., 2000), implemented using the PEBL test battery (Mueller, 2012). In this task participants were presented with a set of spatially-arrayed square targets. In each trial the squares were lit up

[2]Data were collected during collective experimental sessions lasting about two hours, in which the individually administrated tasks were alternated with long breaks in a common area. The tasks, individually administrated, comprised: cognitive tasks (lasting about 20 minutes), self-report questionnaires (lasting about 20 minutes) and motor tasks (lasting about 15 minutes).

one at a time in sequence. When the sequence was finished, participants were asked to click on each square, replicating the order of the given sequence. The task started with a sequence of two squares. Participants were presented with two sequences of equal length. If at least one of the two sequences was recalled correctly, the length of the sequences was increased by one square. The task stopped when the participant failed to correctly reproduce two sequences of equal length. There were three practice trials before beginning in order to check that instructions had been comprehended. The length of the last correctly repeated sequence corresponded to the block span, whereas the total score was computed multiplying the block span by the number of correctly repeated sequences. According to Kessels et al. (2000) this last score is more reliable, thus we present the data using this score. The prefactual/counterfactual task referred specifically to participants' performance on the working memory task. Right after completing the memory task, which we referred to as a "game", participants were randomly assigned to the counterfactual ($n = 102$) or prefactual ($n = 97$) condition. First, they were asked to evaluate their performance on the memory task on a 5-point scale, ranging from 1 (*very bad*) to 5 (*very good*). Participants in the counterfactual condition were then asked to complete the sentence "The game would have been better for me if … ", whereas participants in the prefactual condition were asked to imagine that they could play the game again and to complete the sentence "The game will be better for me next time if … ". A space to complete the sentence was provided. The children were encouraged to write down the first thing that came to mind while thinking about how the outcome of the game would have been/will be better.

Adult participants were tested individually in a quiet room at the university. They were presented with the working memory task and were randomly assigned to the counterfactual ($n = 46$) or prefactual ($n = 40$) condition and they completed the same hypothetical thought task as the children. Given that the selective attention measure was devoted to exploring the possibility that attentional capacities are related to the ability to generate mature counterfactual thinking, adults did not complete the selective attention task.

Results

Coding

In order to code each open-ended response, we first identified *valid* counterfactual/prefactual thoughts, namely, thoughts that undid the outcome by altering an antecedent causally linked to the outcome. Two types of responses that did not meet this criterion can be distinguished. First, responses may consist of irrelevant comments (e.g. "if I did not throw the ball"; "if good"), or could be left empty despite a prompt on the part of the experimenter. These responses were coded as *invalid* counterfactual/prefactual thoughts. Second, the thoughts participants reported may take the form of counterfactual conditionals, but they may simply consist of a restatement of the alternative outcome instead of focusing on how it could plausibly be achieved, thus not meeting the above-mentioned criterion for what we called a valid counterfactual/prefactual thoughts. A thought such as "the game would have been better if I had won" is similar to affirming "the game would have been better if it had been better", thus focusing only on describing what is a better performance and not on how it could be achieved. Similar reasoning may be applied to thoughts like: "if I will remember the sequences", "if I had remembered everything". Given the goal of the task, that is to remember the sequences, a thought such "the game would have been better if I had remembered the sequences" did not actually highlight a way in which performance could have been improved, but just modified the outcome. Nevertheless, this kind of response showed a better understanding of the request than the invalid counterfactual or prefactual thoughts: it retains the counterfactual conditional form and the antecedent is linked to the outcome. We coded such thoughts as *restatements*.

The valid counterfactual/prefactual thoughts were then coded into controllable and uncontrollable modifications following Ferrante et al. (2013), but also including a more specific coding scheme given our main goal to compare the thought content between children and adults. Examples of responses coded in each category are presented in Table 1. In particular, all the elements that participants could control and change if they were required to repeat the task were coded as *controllable*. Specifically, controllable thoughts included all modifications that referred to how participants would have handled/can handle the task differently. This category comprised two subcategories: concentration and strategies. The *concentration* sub-category included modifications related to concentration, attention or effort levels (e.g. "If I concentrated more"), and to the observation of the task's features (e.g. "If I watched the sequences more [carefully]"). The *strategies* sub-category referred to the strategies that can be adopted during the game (e.g. "If I follow the sequence using the mouse") and to the possibility of training (e.g. "if I practice").

All the elements that cannot be controlled and changed by the participants before or during a possible subsequent game were coded as *uncontrollable*. This category comprised both *external* features and *uncontrollable internal* features. In particular, modifications referring to external features were further divided into two sub-categories. Modifications of specific rules and features of the game (e.g. "If the game was slower", "If the same sequences were repeated twice", "If the squares drew the attention more"), and, more generally, of the difficulty of the game (e.g. "If the game was

Table 1. Examples of counterfactual and prefactual thoughts generated by participants in Study 1 arranged according to the coding scheme.

	Counterfactuals	Prefactuals
	The game would have been better for me if…	*The game will be better for me next time if…*
Controllable		
Concentration	• I had paid more attention • I had concentrated better	• I will put in more effort • I will observe better
Strategies	• I had imagined a geometric figure with the sequence of illuminated squares • I had used my finger to follow the movements	• I will mark the position with the mouse • I will watch the squares altogether and not one by one
Uncontrollable - external		
Task	• The sequences had been showed twice • The squares were colored differently	• The sequences will be easier • The last sequences will be slower
Context	• It had been quieter	• I will not have an exam tomorrow
Uncontrollable - internal		
Pre-existing abilities	• I'd had a better memory • I were better at this task	• I had a better memory • I would know how to do the task better
Contingent abilities	• I had seen better • I had not missed a square	• I will understand my errors • I will remember which square will be lit up
Psychophysical state	• I were less tired • I were more rested	• I will be more awake • I will be calmer
Restatement	• I had done everything right • I had answered all the sequences correctly	• I'll do better • I won't get it all wrong
Invalid	• I would be happy • I do good the score	• yes • good but bad

easier") were coded as a *task*. Modification of the environmental features or of the circumstances in which the study was carried out (e.g. "the room was quieter") were coded as *context*. Modifications referring to uncontrollable internal features were further divided into *pre-existing abilities*, *contingent abilities* and *psychophysical state*. Pre-existing abilities included modification to stable abilities that were already present before the study (e.g. "If I was better at this task"), whereas contingent abilities referred to specific abilities that participants lack during the task, but that they may happen to possess in another moment ("I had seen better"; "I will understand my errors"). The sub-category psychophysical state comprised elements like anxiety and tiredness (e.g. "If I was less tired").

Table 2. Frequency and percentage of responses according to the thought validity coding scheme, condition and age group in Study 1.

	Children				Adults			
	Counterfact.		Prefact.		Counterfact.		Prefact.	
	N	%	N	%	N	%	N	%
Valid	80	78.44	79	81.44	42	91.30	40	100.00
Restatement	15	14.71	8	8.25	4	8.70	0	0.00
Invalid	7	6.86	10	10.31	0	0.00	0	0.00
Total	**102**	**100.00**	**97**	**100.00**	**46**	**100.00**	**40**	**100.00**

Two independent judges coded the responses. Their agreement rate was 96.50%, Cohen's k = .95, $p < .001$. Disagreements were resolved by further discussion and the input of a third judge.

Ability to generate counterfactual/prefactual thoughts

The descriptive statistics according to our thought validity coding scheme are presented in Table 2.

12.06% of children and 4.65% of adults generated more than one thought to the open question. Given that we did not explicitly require them to generate more than one thought, only a few participants generated more than one, and following a procedure usually employed in previous studies (e.g. Kahneman & Tversky, 1982; Markman et al., 1995), we analyzed only the first thought that participants came up with[3]. 11.56% of the children generated counterfactual/prefactual thoughts that were coded as restatements and the 8.54% did not generate any counterfactual/prefactual thought (see Table 2). In order to analyze these data, we computed a thought validity score in which invalid counterfactual/prefactual thoughts were scored as 0, restatements as 1, and the remaining thoughts as 2. No difference between the counterfactual and prefactual conditions was found on this score in the children's sample (Mann-Whitney $U = 4845$, $p = .720$, $r_{rb} = .02$), as well as in the adult sample (Mann-Whitney $U = 840$, $p = .059$, $r_{rb} = .07$).

The adults generated no invalid counterfactual/prefactual thoughts and few restatements (see Table 2). The thought validity score significantly differed between children and adults, both in the counterfactual condition (children: $M = 1.72$, $SD = 0.59$, $Md = 2.00$; adults: $M = 1.91$, $SD = 0.28$, $Md = 2.00$; Mann-Whitney $U = 2030$, $p = .048$, $r_{rb} = .13$) and in the

[3]The thought validity score of the second/third thought (when generated) was consistent with the score of the first one generated in all adults ($n = 4$) and in 92% of the children (one child received a better score and one child received a worse score with respect to the first thought generated). As regards the controllable vs. uncontrollable coding, all adults and 71% of children who generated more than one thought were consistent in their responses (generating all controllable or all uncontrollable thoughts).

prefactual condition (children: $M = 1.71$, $SD = 0.64$, $Md = 2.00$; adults: $M = 2.00$, $SD = 0.00$, $Md = 2.00$; Mann-Whitney $U = 1580$, $p = .004$, $r_{rb} = .18$). To summarize, the majority of children were able to change the outcome in their thoughts altering an antecedent that was causally linked to the outcome. Nevertheless, such an ability does not appear to entirely overlap that of adults.

The relationship between age, the cognitive measures and the ability to generate counterfactual/prefactual thoughts in children was analyzed by means of ordinal logistic regression analysis. Given that very few adults reported restatements and none of them came up with invalid counterfactual/prefactual thoughts, we carried out this analysis only in the children's sample. We first verified the effect of demographic measures in a first block, regressing the thought validity score on age and gender (male = 1, female = 2)[4]. Next, in order to verify whether the cognitive measures scores predicted the thought validity score, controlling for demographic characteristics and condition, we added the condition (counterfactual = 1, prefactual = 2) in a second block, and working memory score and selective attention score in a third block. In the first block, age and gender were not significant predictors (age: $B = 0.21$, $SE = 0.14$, $z = 1.56$, $p = .118$, $OR = 1.24$; gender: $B = 0.29$, $SE = 0.36$, $z = 0.81$, $p = .420$, $OR = 1.33$). In the second block, the condition failed to predict the thought validity score ($B = 0.07$, $SE = 0.36$, $z = 0.20$, $p = .842$, $OR = 1.07$; difference between models: $\chi^2(1) = 0.04$, $p = .842$). In the third block, working memory score was a significant predictor of the thought validity score ($B = 0.04$, $SE = 0.01$, $z = 2.72$, $p = .007$, $OR = 1.04$) along with selective attention score ($B = 0.18$, $SE = 0.07$, $z = 2.47$, $p = .013$, $OR = 1.20$; difference between models: $\chi^2(2) = 17.86$, $p < .001$, $R^2_N = .10$). The higher the working memory score and the higher the selective attention score, the higher the probability to generate valid counterfactual/prefactual thought.

Content of counterfactual/prefactual thoughts

In order to analyze the content which participants focused on in their thoughts, in the subsequent analyses we considered only participants who scored 2 on the thought validity score (i.e. excluding restatements and invalid thoughts).

First, we replicated the temporal asymmetry previously found with adults (Ferrante et al., 2013; Mercier et al., 2017) both in children and adults (see Figure 1). The prefactual condition elicited significantly more controllable

[4]We followed a common practice in hierarchical regression analyses, in which demographic variables are usually entered first, and then substantive predictors are entered to assess their effects, above and beyond the demographics.

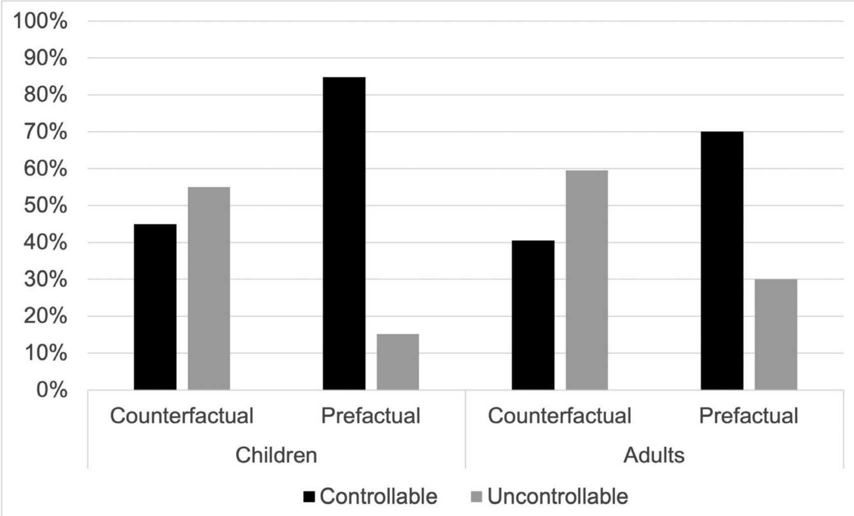


Figure 1. Study 1 percentage of controllable vs. uncontrollable thoughts according to condition (counterfactual vs. prefactual) and age group (children vs. adults).

modifications than the counterfactual condition, both in the children's sample (84.81% vs. 45.00%), $\chi^2(1, N = 159) = 27.61$, $p < .001$, Cramer's V = .42, and in the adults' sample (70.00% vs. 40.48%), $\chi^2(1, N = 82) = 7.21$, $p = .007$, Cramer's V = .30. The proportion of controllable and uncontrollable thoughts did not differ between the two groups in the counterfactual condition ($\chi^2(1, N = 122) = 0.23$, $p = .632$, Cramer's V = .04), whereas, in the prefactual condition, children, compared to adults, tended to generate more controllable thought, even if the difference did not reach the level of significance ($\chi^2(1, N = 119) = 3.62$, $p = .057$, Cramer's V = .17).

Considering the specific content participants focused on more in detail we analyzed the proportion of thoughts generated in each sub-category for controllable and uncontrollable categories separately (see Table 3).

Among controllable thoughts, children almost exclusively focused on concentration/attention level, whereas about 40% of adults focused on strategies, resulting in a significant difference between the two groups both in the counterfactual, $\chi^2(1, N = 53) = 10.65$, $p = .001$, Cramer's V = .45, and in the prefactual, $\chi^2(1, N = 95) = 16.84$, $p < .001$, Cramer's V = .42, conditions (see Table 3).

As regards the uncontrollable category, we distinguished external and internal uncontrollable features, which were further divided into two and three other sub-categories, respectively (task and context; pre-existing abilities, contingent abilities and psychophysical state; see Table 3). Nevertheless, most of these subcategories (with the exception of "task") had less than 5 cases in each cell. For this reason, we analyzed only the higher

Table 3. Frequency and percentage of responses according to the coding scheme, condition and age group among participants who generated a valid counterfactual/prefactual thought in Study 1.

		Children				Adults			
		Counterfact.		Prefact.		Counterfact.		Prefact.	
		N	%	N	%	N	%	N	%
Controllable	Concentr.	35	97.22	62	92.54	11	64.71	16	57.14
	Strategies	1	2.78	5	7.46	6	35.29	12	42.86
	Total controllable	**36**	**100.00**	**67**	**100.00**	**17**	**100.00**	**28**	**100.00**
Uncontrollable	*External*								
	Task	34	77.27	5	41.67	16	64.00	7	17.50
	Context	1	2.27	0	0.00	0	0.00	2	5.00
	Total external	**35**	**79.54**	**5**	**41.67**	**16**	**64.00**	**9**	**75.00**
	Internal								
	Pre-existing	3	6.82	4	33.33	3	12.00	1	8.33
	Contingent	4	9.09	1	8.33	1	4.00	0	0.00
	Psychophy.	2	4.55	2	16.67	5	20.00	2	16.67
	Total internal	**9**	**20.46**	**7**	**58.33**	**9**	**36.00**	**3**	**25.00**
	Total uncontrollable	**44**	**100.00**	**12**	**100.00**	**25**	**100.00**	**12**	**100.00**

Note. Counterfact. = counterfactual condition, Prefact. = prefactual condition, Concentr. = concentration, Pre-existing = pre-existing abilities, Contingent = contingent abilities, Psychophy. = Psychophysical state.

order categorization, that is external vs. internal uncontrollable category. No difference was found between children and adults in the proportion of internal and external uncontrollable modification, neither in counterfactual, $\chi^2(1, N = 69) = 2.00$, $p = .158$, Cramer's V = .17, nor in the prefactual, $\chi^2(1, N = 24) = 2.74$, $p = .098$, Cramer's V = .34, condition (see Table 3).

Discussion

The results of Study 1 show that the majority of the children in our sample were able to generate counterfactuals and prefactuals about their own performance without showing any differences between the two forms of hypothetical thinking. Nevertheless, children appeared not to possess this ability to the same extent as adults. Indeed, 20.10% of children (versus 4.65% of adults) did not report full-fledged counterfactual/prefactual thoughts, irrespective of their age. On the contrary, worse scores on the working memory test and on the selective attention test predicted the likelihood of producing invalid counterfactual/prefactual thoughts.

When a full-fledged counterfactual/prefactual thought was generated, in both the children's and adults' samples we replicated previous findings on the temporal asymmetry in hypothetical thinking (Ferrante et al., 2013; Mercier et al., 2017): participants focused more on controllable features in the prefactual than in the counterfactual condition. As regards the specific content of the thoughts that participants generated, children showed a slight but not significant tendency to generate more controllable thoughts than adults in the prefactual condition. However, almost no child reported thoughts about different strategies that they might have adopted/may adopt in the future to ensure a better performance.

The task in Study 1 is actually a cognitive test aimed at measuring a specific capacity (memory). Given that the task ended when participants failed to remember the correct sequence in two trials of equal length, participants may be prone to focus their attention on the outcome (the abrupt failure), and, as a consequence, they may be more likely to generate thoughts that restate the outcome or the goal of the task (e.g. "If I remember the sequences"). Consequently, such a task may have inflated the number of restatements, even if adults did not show the same pattern. Employing a different task in which making a mistake did not necessarily lead to failure, in Study 2 we aimed to further investigate the differences in counterfactual thinking between children and adults. Given that we found no evidence that prefactual thinking is easier than counterfactual thinking in the age groups we considered, and we found strong evidence of the temporal asymmetry also in children, in Study 2 we focused only on counterfactual thinking.

Study 2

In Study 2 we further investigated children's ability to produce counterfactuals about their own experience and the specific content they focus on with the aim of generalizing the findings of Study 1 using a different task. To this end, we used a task that offers a greater level of control and gives the possibility of continuing with the task despite any possible mistake (i.e. connecting a series of numbers and letters without lifting the pen off the paper within a given amount of time). Moreover, in order to check whether the number of restatements may have been inflated by the task ending when two mistakes were made, we manipulated the goal of the task: in one condition the goal was to finish the task as soon as possible (*definite* goal), whereas in the other condition the goal was to do their best in the given time (*indefinite* goal). In both conditions, participants were informed that there was time limit, but that they were not informed of how long this limit actually was. In fact, the time limit was individually adapted: when participants had completed approximately two-thirds of the path, they were told that the time expired. If the abrupt failure in Study 1 led participants to focus more on the outcome increasing the likelihood of restatements, we should find a greater number of restatements in the definite goal condition, in which they clearly failed the task, compared to the indefinite goal condition.

Moreover, allowing participants to repeat the game after generating counterfactual thoughts, we had the possibility to collect not only performance evaluation, but also predictions on future performance that could contribute explaining possible differences in thoughts generated by children and adults.

Method

Participants

One hundred and eighty-seven children (90 male and 97 female, age: $M = 9.32$, $SD = 1.03$, ages range from 8 to 12 years) not involved in the previous study participated in the study. One hundred fifty-three children came from primary school (age: $M = 8.95$, $SD = 0.74$, ages range from 8 to 10) and 34 came from the first year of middle school (age: $M = 10.94$, $SD = 0.34$, ages range from 10 to 12). All children had the written consent of their parents/guardian to participate in the study. Participants came from two different primary and middle schools in Southern Italy from diverse socio-economic backgrounds and different neighborhoods.

Moreover, 97 adults (49 male and 48 female; age: $M = 23.72$, $SD = 3.37$; ages range from 19 to 39), not involved in the previous study, were recruited to take part in the study. They came from various towns in Italy. Sample size was decided in the same way as in Study 1.

The procedure was approved by the Ethical Committee of the University of Trieste (report number 116).

Materials and procedure

Children were tested individually in a quiet area in their schools. Following Study 1, they first took part in a game. In particular, they were presented with a modified version of the Trail Making Test-B (TMT-B; Tombaugh, 2004). The original version of the TMT-B requires an individual to draw lines sequentially connecting 25 encircled numbers and letters distributed on a sheet of paper, alternating between numbers and letters (e.g. 1, A, 2, B, 3, C, etc.). In order to make the task equally challenging for children and adults, based on a pretest with a sample of children and a previous experiment using the same task, we created two versions of the task. In the children's version, the sheet included numbers from 1 to 11, and letters from A to L, whereas the adults' version included numbers from 1 to 17 and letters from A to R. Moreover, to reduce the possible difference between adults and children in the cognitive availability of the number-letter sequence (which should be higher in adults) we provided all participants with a sheet of paper next to the task sheet displaying the correct alternate order of the numbers and letters.

The instructions for the task varied according to the goal-condition to which participants were randomly assigned. In one condition, participants were told that their task was to connect all the elements, alternating numbers and letters in the shortest time possible (*definite goal* condition, $n = 92$ children and $n = 48$ adults). In the other condition they were told that their task was to connect as many elements as possible alternating numbers and letters before the time elapsed (*indefinite goal condition*, $n = 95$ children and $n = 49$ adults). In both conditions participants were told that a time limit was set, but that they would not be informed about what that limit was. In fact, when children reached the number 9 and adults reached the number 14, which corresponded to 80% of the task having been completed, the experimenter told participants (in both conditions) that the time had elapsed. In this way, all participants received the same feedback and the experimenter registered the amount of time they spent to reach the stopping point. Finally, participants were instructed to avoid making mistakes and not to lift the pen off the paper during the entire task, as this would be considered an error. The task was presented to participants as a game, and the experimenter read the instructions aloud with the children, whereas only written instructions were provided to adults. Children (not adults) performed a short practice trial before beginning the task (see the Appendix for detailed instructions).

After completing the task, all participants were presented with a questionnaire. First, participants were asked to evaluate their performance on a 5-point scale ranging from 1 (*very bad*) to 5 (*very good*). Next, participants were informed that they would be asked to tackle another version of the game again in a few minutes. They were then asked to think about their performance and to complete, in at least one way, the sentence: "The game would have been better for me if… ". Two boxes were provided for the answers. Next, they were asked to rate the extent to which they thought that: they could have performed better in the previous game, they would perform better in the next game, they would reach the end of the game in the next attempt, on a 5-point scale, ranging from 1 (*definitely not*) to 5 (*definitely yes*). Finally, they were asked whether they thought they were good at this kind of game, using a 5-point scale ranging from 1 (*not at all*) to 5 (*very much*). Children completed 5 more questions about their academic performance (i.e. their level of confidence in passing the academic year, general performance evaluation, average grade in Math, History and Italian).

After completing the questionnaire, participants took part in the second game. Participants were presented with the same game but the letters and

Table 4. Examples of counterfactual thoughts generated by participants in Study 2 arranged according to the coding scheme.

Category	Counterfactuals
	The game would have been better for me if …
Controllable	
Concentration	• I had put in more effort • I was going faster
Strategies	• I had checked the above sheet [with the alphabet] • I had had a look at the letters and numbers before starting connecting them
Errors	• I had not lifted the pen • I had not linked two letters in a row
Uncontrollable - external	
Task	• I could have lifted the pen • I had had more time
Context	• The room was quieter • I had performed the task in the morning
Uncontrollable - internal	
Pre-existing abilities	• My hand-eye coordination was faster • My hand was smaller
Contingent abilities	• I had seen where the letters were • I had better understand the instructions
Psychophysical state	• I were not tired • I were less anxious
Restatement	• I had won • I had linked more elements
Invalid	• the task is good but difficult • so so

numbers were arranged differently. The order of the first and second game was counterbalanced across participants. In order to avoid distress feelings in children, in the second game they were allowed to reach the end of the task and the experimenter recorded the time they spent to reach the point in which they were stopped in the first game and the overall time spent to complete the task. Instead, adults were given the same amount of time which they had spent to reach the point in which they were stopped in the first game, and the experimenter recorded the point they reached in the second game.

Results

Coding

We adopted the same coding scheme used in Study 1. Examples of responses for each category are reported in Table 4. Responses that contained irrelevant comments or missing responses were coded as *invalid*

Table 5. Frequency and percentage of responses according to the thought validity coding scheme and age group in Study 2.

	Children		Adults	
	N	%	N	%
Valid	149	79.68	95	97.94
Restatement	29	15.51	2	2.06
Invalid	9	4.81	0	0.00
Total	**187**	**100.00**	**97**	**100.00**

counterfactual thought. Tautological responses that described a better performance without highlighting how it could be achieved were coded as *restatement*. The remaining responses were coded as *valid* counterfactual thoughts and further divided into *controllable* or *uncontrollable* categories. The sub-categories for the controllable features were *concentration*, *strategies* (as in Study 1) and a new added subcategory: *errors*. Given the possibility of making a wrong connection or lifting the pen off the page, we added the errors subcategory which included all the responses in which participants highlighted that they would have avoided making mistakes (e.g. "If I did not make mistakes"), or to lift the pen off the paper (e.g. "If I did not lift the pen"). The *external uncontrollable* features included the same subcategories of Study 1 (*task* and *context*), as well as the *internal uncontrollable* subcategory (*pre-existing abilities*, *contingent abilities* and *psychophysical state*). Two independent judges coded the responses. Their agreement rate was 97.89%, Cohen's k = .97, $p = .001$, and disagreements were resolved by discussion and the input of a third judge.

Ability to generate counterfactual thoughts

The descriptive statistics according to the thought validity coding scheme are presented in Table 5.

34.22% of children and 21.65% of adults modified more than one thought to the open question. The fact that in Study 2 we provided participants with two boxes for their answers instead of one may have increased the probability of generating more than one thought. Nevertheless, the majority of both children and adults generated only one thought. Therefore, following Study 1 and previous studies (e.g. Kahneman & Tversky, 1982; Markman et al., 1995), we analyzed only the first thought that participants came up with[5]. 15.51% of the children's thoughts were

[5]The thought validity score of the second/third thought (when generated) was consistent with the score of the first one in all adults ($n = 21$) and in 80% of the children (only three children received a better score for the second thought, whereas the other eight children received a worse score). As regards the controllable vs. uncontrollable coding, 62% of adults and 84% of children who generated more than one thought were consistent in their responses.

coded as restatements, whereas only 2.06% of the adults' thoughts were coded similarly. 4.81% of the children (and no adults) reported invalid thoughts (see Table 5). Following Study 1, we computed the thought validity score in the same way (invalid = 0, restatement = 1, valid = 2). Children showed a significantly lower thought validity score ($M = 1.75$, $SD = 0.53$, $Md = 2.00$) than adults ($M = 1.98$, $SD = 0.14$, $Md = 2.00$; Mann-Whitney $U = 7405$, $p < .001$, $r_{rb} = .18$), showing a lower capacity to generate counterfactual thoughts, in line with Study 1.

Moreover, goal-condition had no effect on this score, neither in the children's sample (Mann-Whitney $U = 4103$, $p = .304$, $r_{rb} = .06$), nor in the adult's sample (Mann-Whitney $U = 1127$, $p = .155$, $r_{rb} = .04$). The finding that the clear failure in the definite goal condition did not decrease the thought validity score, along with a similar percentage of restatements in Study 2 with respect to Study 1 (15.51% and 14.71%, respectively), rules out the idea that restatements were inflated by the rules of the task in Study 1.

We ran an ordinal logistic regression analysis in order to test whether the thought validity score depended on demographic characteristics, performance on the game and self-reported academic performance. The analysis was run on the children's sample only, given that 98% of adults produced valid counterfactuals. We entered the predictors into three blocks. As in Study 1, in the first block, we entered demographic measures: children's age and gender (male = 0, female = 1). Neither gender ($B = 0.47$, $SE = 0.37$, $z = 1.26$, $p = .206$, $OR = 1.60$) nor age ($B = -0.07$, $SE = 0.18$, $z = 0.42$, $p = .677$, $OR = 0.93$) were significant predictors. In the second block, we entered the performance on the metrics of the first game (time spent, number of mistakes and number of pen lifts), in order to assess its effect controlling for demographic variables. This model fit the data better than the first model ($\chi^2(3) = 8.31$, $p = .040$; $R^2_N = 0.05$). The time spent on the task significantly predicted the thought validity score ($B = -0.02$, $SE = 0.01$, $z = 2.80$, $p = .005$, $OR = 0.98$): as the time spent by participants on the task increased, the thought validity score decreased. The number of mistakes participants made during the task ($B = 0.26$, $SE = 0.24$, $z = 1.09$, $p = .276$, $OR = 1.30$) and the number of pen lifts ($B = 0.20$, $SE = 0.12$, $z = 1.73$, $p = .082$, $OR = 1.22$) did not significantly affect the thought validity score. Gender and age were still not significant ($zs < 1.20$, $ps > .235$). In the last block we added the general evaluation of their academic performance and the average of their grades in school, in order to verify whether self-reported academic performance may predict thought validity score, controlling for demographic variables and performance metrics. This third model did not fit the data better than the second one ($\chi^2(2) = 4.98$, $p = .083$).

To summarize, children and pre-adolescents seem to possess the ability to reflect on how their own performance could have been different, but

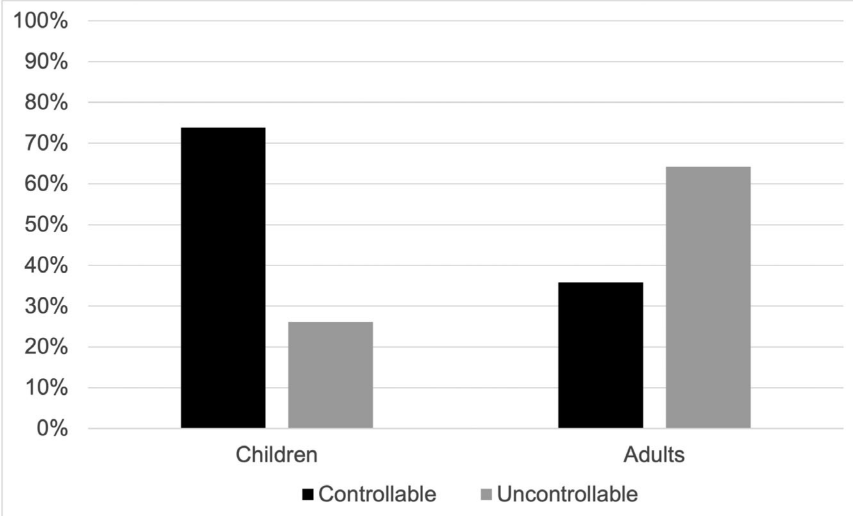


Figure 2. Study 2 percentage of controllable vs. uncontrollable counterfactual thoughts according to age group (children vs. adults).

they do not completely reach the adults' ability to do so, replicating the results of Study 1. This effect does not appear to be related to children's age, but to a more general ability: children who scored worse on the thought validity score spent a greater amount of time on the task.

Content of counterfactual thoughts

As in Study 1, the thought content analyses were based on participants who scored 2 on the thought validity score (i.e. excluding restatements and invalid thoughts).

The proportion of controllable and uncontrollable counterfactual thoughts strongly differed between the sample of children and adults, $\chi^2(1, N = 244) = 34.70$, $p < .001$, Cramer's V = .38 (see Figure 2). Differently from Study 1, children generated significantly more controllable counterfactuals than adults (73.83% vs. 35.79%, respectively). No differences were found between goal conditions (i.e. the definite and indefinite goals) neither in the children's sample, $\chi^2(1, N = 149) = 0.17$, $p = .680$, Cramer's V = .03, nor in the adults' sample, $\chi^2(1, N = 95) = 0.04$, $p = .843$, Cramer's V = .02.

Following Study 1, we next analyzed more in detail the specific content of the counterfactual thoughts separately for controllable and uncontrollable categories (see Table 6 for details).

Among controllable modifications, the thought distribution into categories of concentration, strategies and errors differed between children and adults, $\chi^2(2, N = 144) = 22.11$, $p < .001$, Cramer's V = .39. In particular,

Table 6. Frequency and percentage of responses according to the coding scheme and age group among participants who generated a valid counterfactual thought in Study 2.

		Children		Adults	
		N	%	N	%
Controllable	Concentr.	49	44.55	21	61.77
	Strategies	8	7.27	10	29.41
	Errors	53	48.18	3	8.82
	Total controllable	**110**	**100.00**	**34**	**100.00**
Uncontrollable	*External*				
	Task	17	43.59	42	68.85
	Context	0	0.00	3	4.92
	Total external	**17**	**43.59**	**45**	**73.77**
	Internal				
	Pre-existing	3	7.69	4	6.56
	Contingent	17	43.59	5	8.19
	Psychophy.	2	5.13	7	11.48
	Total internal	**22**	**56.41**	**16**	**26.23**
	Total uncontrollable	**39**	**100.00**	**61**	**100.00**

Note. Concentr. = concentration, Pre-existing = pre-existing abilities, Contingent = contingent abilities, Psychophy. = Psychophysical state.

children focused much more on errors than adults: 48.18% of the children's controllable counterfactuals focused on this subcategory, whereas only 8.82% of the adults' controllable counterfactuals did so, $\chi^2(1, N = 144) = 16.93$, $p < .001$, Cramer's V = .34. On the other hand, children compared to adults focused less on strategies (7.27% vs. 29.41% of controllable thought, respectively), $\chi^2(1, N = 144) = 11.64$, $p < .001$, Cramer's V = .28, in line with Study 1. Finally, children and adults did not differ in the proportion of thoughts regarding concentration, $\chi^2(1, N = 144) = 3.08$, $p = .079$, Cramer's V = .15. As regards the uncontrollable category, following Study 1, we analyzed the proportion of external and internal uncontrollable thoughts among all the uncontrollable thoughts. Children tended to focus more on internal features in their uncontrollable thoughts than adults (56.41% vs. 26.23%), $\chi^2(1, N = 100) = 9.20$, $p = .002$, Cramer's V = .30.

Finally, no differences between goal conditions were found in any subcategories, neither in the sample of children (χ^2s < 3.03, *ps* $> .220$) nor in the adult sample (χ^2s < 3.32, *ps* $> .068$).

Taken together, these results suggest that the increased number of controllable thoughts in the sample of children with respect to Study 1 and the subsequent difference compared to the adult sample was driven by the possibility of focusing on the errors that participants made: 35.57% of children focus on their own errors (representing almost half of the children's controllable counterfactuals) while almost no adults did so (3.16%, representing about 9% of the adults' controllable counterfactuals).

The effect of performance and performance evaluation

Given that we did not find any differences on the dependent variables between goal conditions, we did not consider this variable in the following analyses.

Compared to adults, the larger proportion of the children's controllable counterfactuals focusing on avoiding errors might be driven by the fact that children lifted the pen off the page more frequently than adults (1.79 vs. 0.93, $t(282) = 4.01$, $p < .001$, $d = 0.50$), whereas they made a similar number of mistakes (0.63 vs. 0.45, $t(282) = 1.44$, $p = .151$, $d = 0.18$). To rule out this possibility, we tested by means of logistic regression whether the general content of the counterfactual thoughts (controllable = 1 vs. uncontrollable = 2) might be predicted by performance metrics (time spent, number of mistakes and number of pen lifts in the first game), and whether the difference between age groups would disappear after controlling for those variables. In particular, we considered: in the first block the age groups (children = 0 vs. adults = 1), the time spent in the first game (z standardized for children and adults separately, given that the adult version was longer), the number of mistakes and the number of pen lifts; in the second block the 2-way interactions between age groups and performance variables. The first model ($\chi^2(4) = 50.55$, $p < .001$, $R^2_N = .25$) showed that as the number of pen lifts increased, the probability of generating an uncontrollable counterfactual decreased ($B = -0.30$, $SE = 0.12$, $z = 2.57$, $p = .010$, $OR = 0.74$), while the number of mistakes ($B = -0.20$, $SE = 0.17$, $z = 1.17$, $p = .244$, $OR = 0.82$) and the time spent in the first game ($B = -0.22$, $SE = 0.18$, $z = 1.23$, $p = .219$, $OR = 0.81$) did not have any effect on thought content. The age group difference was still strong and significant ($B = 1.46$, $SE = 0.30$, $z = 4.83$, $p < .001$, $OR = 4.32$), suggesting that the increased number of controllable counterfactuals in children as opposed to adults was not due to a larger number of errors in this sample. The second model (with the 2-way interactions) did not fit the data better than the first one ($\chi^2(3) = 0.90$, $p = .825$), suggesting that performance metrics had no different effect on thought content in the two age groups. These results showed that when children and adults erroneously lifted the pen, both were more likely to generate controllable counterfactuals, but the difference between adults and children is still strong and significant after controlling for the number of pen lifts and other performance metrics.

As regards the retrospective self-report measures, children, compared to adults, better evaluated their first performance (3.91 vs. 3.30, $t(282) = 7.19$, $p < .001$, $d = 0.90$) and their ability at this kind of game (3.61 vs. 3.08, $t(282)$

$= 6.57$, $p < .001$, $d = 0.82$), but they were less confident they would have performed better in the previous game (4.10 vs. 4.38, $t(282) = 3.34$, $p < .001$, $d = 0.42$). In order to test whether such different evaluations might have affected the thought content, explaining the difference between adults and children on the proportion of controllable counterfactuals, we regressed the thought content onto all these variables in a logistic regression analysis (controllable = 1; uncontrollable = 2). The results of the model ($\chi^2(4) = 50.47$, $p < .001$, $R^2_N = .25$) showed that the better the performance was evaluated, the higher the probability to generate uncontrollable counterfactuals ($B = 0.88$, $SE = 0.24$, $z = 3.63$, $p < .001$, $OR = 2.42$), while the other ratings were not significant predictors (zs < 1.57, ps $> .11$). The effect of age groups remained strong and significant ($B = 2.01$, $SE = 0.35$, $z = 5.77$, $p < .001$, $OR = 7.48$).

As regards predictions about future performance, compared to adults, children had a higher level of confidence that they would finish the game on the second try (3.57 vs. 3.33, $t(282) = 2.22$, $p = .027$, $d = 0.28$), even if they were equally confident regarding future improvement (3.92 vs. 3.86, $t(282) = 0.63$, $p = .529$, $d = 0.08$).

Taken together, these results show that controllable counterfactuals were more likely as participants lifted the pen more often and their performance evaluation decreased. Nevertheless, the greater number of controllable counterfactuals in children cannot be explained by performance metrics and participant evaluations.

Discussion

In line with Study 1, the majority of children demonstrated the ability to reflect counterfactually on their own performance, but to a lesser extent than adults. As previously found in Study 1, the age of the children was not related to the counterfactual ability, and children who did not report a fully-fledged counterfactual thought also performed worse on the task, suggesting a relationship with the development of other abilities. These results indicate that the task used in Study 1 did not inflate the restatements, as in Study 2 we replicated the findings on the thought validity score using a different task, and framing the goal as definite or indefinite had no effect.

Unlike Study 1, we found a strong difference between children and adults regarding the number of controllable thoughts. Such a difference seems to be driven by the fact that children tended to focus on the errors they made, whereas adults preferred to focus on external features. Indeed, the proportion of controllable thoughts referring to concentration remained unchanged between children and adults, whereas the proportion of controllable thoughts referring to errors is much higher in children.

Interestingly, note that after making an error, it is possible to focus on the action that led to such a mistake (the controllable subcategory "errors") or on the rule that defines such an action as an error (i.e. uncontrollable feature referring to the task). Considering thoughts referring to pen lifts in both samples may give us a hint about the possibility that children focused on the action whereas adults on the rule behind it. Among all the children's mentions of lifting the pen off the paper ($n = 34$, 22.82%), the large majority focus on how they could have avoided doing so ($n = 33$), and just one child focused on the rule behind it (i.e. "if only I could have lifted the pen"). On the contrary, among all the adults' thoughts referring to pen lifts ($n = 11$, 11.58%), the large majority undid the rule that forbade pen lifts ($n = 9$) and did not undo the action of lifting the pen ($n = 2$; see the General Discussion).

We also found that a greater number of pen lifts prompted controllable counterfactuals. Moreover, controllable counterfactuals were more likely when participants were not satisfied with their performance.

General discussion

In two studies we investigated for the first time the capacity to generate valid counterfactual thoughts and the content of such thoughts when children reflected on events that actually happened to them. In particular, after taking part in a game, children reported their thoughts about how their performance would have been better. As expected, results showed that the majority of children aged 8 to 13 are able to report valid counterfactual thoughts, despite not reaching adults' ability: about 20% of children (21.6% in Study 1, 20.3% in Study 2), versus about 5% of adults (8.7% in Study 1, 2.1% in Study 2), provided a restatement or an invalid counterfactual thought. In both studies, the capacity to generate valid counterfactuals does not appear to be related to the children's age, but to general abilities. Indeed, we found that participants who scored worse on the working memory test and the selective attention test in Study 1 and those who performed worse in the game in Study 2[6], were more likely to generate an invalid counterfactual thought. Based on these results, it is possible that, on one hand, low cognitive skills impair comprehension of the counterfactual instructions (in which they were required to find an antecedent sufficient to undo the outcome) or, on the other hand, that the complex causal structure of the representation of a just-experienced event required a higher level of cognitive skills than those required in simpler tasks used in other studies

[6]Note that we cannot consider the score obtained on our adapted trail making test as a measure of cognitive flexibility given that the path, the instruction and the procedure were modified for the aim of the study.

(see Nyhout & Ganea, 2019b). However, the fact that the counterfactual thought task was based on children's performance on the same task used to measure cognitive abilities highlighted a possible confound that prevented us from drawing conclusive result. Future research may overcome this limitation using different tasks to measure cognitive skills and elicit counterfactual worlds. Nevertheless, it is worth noting that in the two studies we obtained similar findings on the ability to generate counterfactual thoughts about children's own experience, despite employing two different tasks.

In Study 1 we also showed that children engaging in prefactual thinking (i.e. imagining how the performance will be better in the future) experienced the same difficulties as children engaging in counterfactual thinking (18.56% of restatements or invalid prefactual thoughts), as we expected. This result may not seem to be in line with previous literature suggesting that prefactual thinking is easier than counterfactual thinking (e.g. Perner et al., 2004; Riggs et al., 1998; Robinson & Beck, 2000). However, as we pointed out in the introduction, in order to generate a valid prefactual thought in our study, participants needed not only to simulate a future scenario, but also to consider what had occurred just a few minutes before. Consequently, participants need to keep in mind two representations (i.e. the present and the future ones) just as well as in counterfactual thinking (i.e. the present and the alternative ones). Therefore, our results actually support previous findings; when prefactual thinking requires two simultaneously active representations, it is no longer easier than counterfactual thinking, in line with the idea that the previously found difference in difficulty between counterfactual and prefactual thinking lies in the representations needed to provide a correct answer (Rafetseder et al., 2010).

Our findings have shown that it seems more difficult generating counterfactual thoughts about a just-experienced event than about events presented in scenario-based paradigms. This difference may lie not so much in the perspective from which children build alternative worlds (first vs. third person), but in the complexity of a directly experienced event. First, studies with adults showed that perspective does not matter if participants were exposed to the same information: observers of actors solving a task modified the same antecedents as the actors themselves (whereas participants who read about the same event focus on other features; Pighin et al., 2011). Second, it appears that children reported evidence of regret about their own experience earlier than about the experiences of others (Weisberg & Beck, 2010). Therefore, it may be expected that first person counterfactuals are easier than third person counterfactuals (Beck, 2020). The difference may thus lie in the complexity of a real-world experience that may challenge children's capacity to consider a broader range of possible antecedents.

As regards our second aim, which was comparing the thought content between children and adults, we analysed controllable and uncontrollable modifications in children and adults experiencing the same event. In Study 1, children exhibited the same temporal asymmetry in hypothetical thoughts between controllable and uncontrollable modifications previously found with adults (Ferrante et al., 2013, Mercier et al., 2017): Both children and adults generated more controllable thoughts in the prefactual than in the counterfactual condition. Moreover, children's valid thought content did not differ with respect to adults, except for the fewer modifications focusing on the strategies that they would have adopted to obtain a better result. On the contrary, in Study 2 we found a strong difference in the counterfactual content between the two age groups: children focused on controllable modifications twice as much as adults.

From these findings two questions arise. First, why did children focus more on controllable modifications than adults in Study 2? Second, why did we find such a different result in the two studies? Analysing the content of controllable thoughts, we observed that while almost no adults mentioned errors in their thoughts, 36% of children undid the errors they had made, which seems to explain the difference in the controllable modifications between children and adults in Study 2. Focusing on error modifications seems to be a child-specific tendency, given that even controlling for the actual number of errors and the performance evaluation, the difference with adults remain strong and significant. We can only speculate about the basis of this tendency.

A first possibility may be directly related to the development of the ability to consider a broader range of possibilities when thinking counterfactually: children might not be able to fully consider the impact of the external context in a real-world experience such as the one we used in our study, in which a more sophisticated model of reality needs to be considered. In other words, the mechanism underlining this effect might be similar to the hypothesized explanation for the actors vs. readers effect (Girotto et al., 2007), in which readers of a scenario describing an individual's failure to solve a problem focused more on controllable modifications than individuals who actually failed at the same problem (i.e. actors). The readers' lower availability of information about the problem-solving phase may have led them to undo the salient factors in their counterfactual words, such as the protagonist's actions, whereas actors who had access to all the information about such a phase may choose from a broader set of factors one that would have been more likely to undo the negative result. Similarly, children's representation of the event may be simpler and less sophisticated, leading them to focus on salient factors instead of considering a large set of possibilities. Indeed, in Study 2, while children referring to errors in their thoughts simply undid the errors they had made, almost all adults that mentioned errors focused on one cause of their errors (i.e. the existence of the rule). Moreover, in both studies almost none of the children thought about alternative strategies that could have improved their performance (i.e. about 3% of the controllable modifications vs. 36% of the adult sample), a category of thoughts that may require a more complex representation of the experience. Taken together, these results suggest that some of the children's counterfactuals may result from basic conditional reasoning based on available general assumptions (e.g. errors decrease performance), that may derive from the errors' salience in the school context.

Another explanation relies on possible differences in blame attribution. Some insight may come from studies providing evidence that even when the ability to reason counterfactually has been developed, counterfactual emotions and some biases previously found in adults may emerge later in childhood. For instance, Payir and Guttentag (2016) showed that children over the age of 12 but not younger children applied counterfactual thinking as a consoling strategy to a similar extent as adults. Moreover, recently Payir and Guttentag (2019) found new evidence that some biases in counterfactual thinking may emerge later than others and not necessarily together with the ability to think counterfactually. In particular, the authors showed that children aged less than 11 years did not show the action/inaction bias, namely, they did not cast more blame on the character who acted than the character who did not act when both were responsible for a negative outcome. A similar mechanism may explain our results: children may not cast more blame onto themselves for their actions, and so they might be willing to focus on their errors. On the other hand, focusing on errors may lead adults to blame themselves, and so they shifted their attention to other external factors. Moreover, Payir and Guttentag (2019) showed that when 11 year olds and adults, but not younger children, took the prospective of others, they tended to focus on the actions of the protagonist, whereas when they made blame judgments themselves, they focused more on situational factors. Our results may be in line with this finding, but the question is still open as to whether the low number of controllable counterfactuals when adults reflect on their own experience depends on avoiding blame themselves (Girotto et al., 2007).

For another point of view, studies on developmental changes in children's concept of ability showed that the notion of ability as separated from effort emerges later in development (after 11 years of age), while effort tends to be seen by children aged 7 to 9 years as the primary cause of performance (Nicholls, 1978; Nicholls & Miller, 1984; Folmer et al., 2008; see also Muenks et al., 2018). As a consequence, children may have generated more controllable counterfactuals about their errors because of such a tendency to see the effort put into the game as the main cause of performance. Future research might disentangle which of the proposed explanations (complex mental model, blame attribution, effort vs. ability) better explains why children focus on different content with respect to adults, contributing to our knowledge regarding children's ability to think counterfactually.

As regards the second question, the reason why we find different results in the two studies, a possible explanation may be derived considering the different characteristics of the tasks used in Study 1 vs. Study 2. Whereas in Study 1 the game was a memory test in which the performance was given by the number of sequences correctly reproduced before failing, in Study 2 the goal was to connect the dots as fast as possible, and errors would impede performance but not lead to completely failing at the task. Therefore, the errors might become a valid candidate for the modification in the counterfactual world only in Study 2. In the task given in Study 1, altering the errors would result in a tautological thought: errors made in the first game means failing the task, so focusing on not making errors equals to stating "the game would have been better if I had not failed at the game". This may have led children who are able to generate counterfactual thinking to search for other factors that may have impacted their performance, such as external features. On the contrary, making less errors in the second game would result in a better and faster performance, so this antecedent became valid in Study 2 and was altered by 35% of the children. The tendency to focus on the mistakes they made may have been hidden in Study 1 by the task features that did not allow for the modifications of their own errors. Nevertheless, it is worth noting that the proportion of restatements in Study 1 was not inflated by the fact that valid counterfactuals could not be generated by undoing the errors made, given that restatements' proportion was identical in the two studies (about 15%).

In conclusion, our studies shed a new light into children's ability to think counterfactually, extending the investigation to more complex situations in which children have to reflect on their own experiences (and not only read about or observe the actions of an external character). Future research may continue exploring the content of counterfactual thoughts regarding children's first-hand experiences, ruling out the effect of perspective change in real world experiences (actors vs. observers) and further investigating the existence of fault lines in children's counterfactuals.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Marta Stragà http://orcid.org/0000-0003-2807-0628

References

Beck, S. R. (2020). The subject of children's counterfactual thoughts. *Psychology of Consciousness: Theory, Research, and Practice*, *7*(4), 340–350. https://doi.org/10.1037/cns0000235

Beck, S. R., Carroll, D. J., Brunsdon, V. E., & Gryg, C. K. (2011). Supporting children's counterfactual thinking with alternative modes of responding. *Journal of Experimental Child Psychology*, *108*(1), 190–202. https://doi.org/10.1016/j.jecp.2010.07.009

Beck, S. R., & Rafetseder, E. (2019). Are counterfactuals in and about time? *The Behavioral and Brain Sciences*, *42*, e245. https://doi.org/10.1017/S0140525X19000591

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking & Reasoning*, *15*(4), 337–354. https://doi.org/10.1080/13546780903135904

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development*, *77*(2), 413–426. https://doi.org/10.1111/j.1467-8624.2006.00879.x

Bisiacchi, P. S., Cendron, M., Gugliotta, M., Tressoldi, P. E., & Vio, C. (2005). *BVN 5-11: Batteria di valutazione neuropsicologica per l'età evolutiva*. Erickson.

Burns, P., Riggs, K. J., & Beck, S. R. (2012). Executive control and the experience of regret. *Journal of Experimental Child Psychology*, *111*(3), 501–515. https://doi.org/10.1016/j.jecp.2011.10.003

Burns, P., McCormack, T., Jaroslawska, A., Fitzpatrick, Á., McGourty, J., & Caruso, E. M. (2019). The development of asymmetries in past and future thinking. *Journal of Experimental Psychology. General*, *148*(2), 272–288. https://doi.org/10.1037/xge0000464

Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157. https://doi.org/10.1146/annurev-psych-122414-033249

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

Dias, M. G., & Harris, P. L. (1988). The effect of make-believe play on deductive reasoning. *British Journal of Developmental Psychology*, *6*(3), 207–221. https://doi.org/10.1111/j.2044-835X.1988.tb01095.x

Drayton, S., Turley-Ames, K. J., & Guajardo, N. R. (2011). Counterfactual thinking and false belief: The role of executive function. *Journal of Experimental Child Psychology*, *108*(3), 532–548. https://doi.org/10.1016/j.jecp.2010.09.007

Epstude, K., & Roese, N. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *12*(2), 168–192. https://doi.org/10.1177/1088868308316091

Epstude, K., Scholl, A., & Roese, N. J. (2016). Prefactual thoughts: Mental simulations about what might happen. *Review of General Psychology*, *20*(1), 48–56. https://doi.org/10.1037/gpr0000064

Ferrante, D., Girotto, V., Stragà, M., & Walsh, C. (2013). Improving the past and the future: A temporal asymmetry in hypothetical thinking. *Journal of Experimental Psychology. General*, *142*(1), 23–27. https://doi.org/10.1037/a0027947

Folmer, A. S., Cole, D. A., Sigal, A. B., Benbow, L. D., Satterwhite, L. F., Swygert, K. E., & Ciesla, J. A. (2008). Age-related changes in children's understanding of effort and ability: Implications for attribution theory and motivation. *Journal of Experimental Child Psychology*, *99*(2), 114–134. https://doi.org/10.1016/j.jecp.2007.09.003

German, T. P., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science*, *6*(5), 514–523. https://doi.org/10.1111/1467-7687.00309

Girotto, V., Ferrante, D., Pighin, S., & Gonzalez, M. (2007). Post-decisional counterfactual thinking by actors and readers. *Psychological Science*, *18*(6), 510–515. https://doi.org/10.1111/j.1467-9280.2007.01931.x

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, *78*(1–3), 111–133. https://doi.org/10.1016/0001-6918(91)90007-M

Goodman, N. (1983). *Fact, fiction, and forecast*. Harvard University Press.

Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, *19*(1), 53–80. https://doi.org/10.1016/j.cogdev.2003.09.002

Guajardo, N. R., McNally, L. F., & Wright, A. (2016). Children's spontaneous counterfactuals: The roles of valence, expectancy, and cognitive flexibility. *Journal of Experimental Child Psychology*, *146*, 79–94. https://doi.org/10.1016/j.jecp.2016.01.009

Guajardo, N. R., Parker, J., Turley., & Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *The British Journal of Developmental Psychology*, *27*(Pt 3), 681–702. https://doi.org/10.1348/026151008X357886

Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: Age differences in judgments of regret and relief. *Developmental Psychology*, *40*(5), 764–775. https://doi.org/10.1037/0012-1649.40.5.764

Hammell, C., & Chan, A. Y. (2016). Improving physical task performance with counterfactual and prefactual thinking. *Plos One*, *11*(12), e0168181–12. https://doi.org/10.1371/journal.pone.0168181

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233–259. https://doi.org/10.1016/S0010-0277(96)00715-9

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.015

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258. https://doi.org/10.1207/S15324826AN0704_8

Leevers, H. J., & Harris, P. L. (2000). Counterfactual syllogistic reasoning in normal 4-yearolds, children with learning disabilities, and children with autism. *Journal of*

Experimental Child Psychology, 76(1), 64–87. https://doi.org/10.1006/jecp.1999.2542
Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1995). The impact of perceived control on the imagination of better and worse possible worlds. *Personality and Social Psychology Bulletin*, *21*(6), 588–595. https://doi.org/10.1177/0146167295216005
McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development*, *45*, 1–9. https://doi.org/10.1016/j.cogdev.2017.10.001
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*, *148*, 1–19. https://doi.org/10.1016/j.jecp.2016.02.008
Meehan, J. E., & Byrne, R. M. J. (2005). Children's counterfactual thinking: The temporal order effect. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1467–1473). Lawrence Erlbaum.
Mercier, H., Rolison, J., Stragà, M., Ferrante, D., Walsh, C., & Girotto, V. (2017). Questioning the preparatory function of counterfactual thinking. *Memory & Cognition*, *45*(2), 261–269. https://doi.org/10.3758/s13421-016-0660-5
Mueller, S. T. (2012). The psychology experiment building language (version 0.13).[Computer software]. http://pebl.sourceforge.net.
Muenks, K., Wigfield, A., & Eccles, J. S. (2018). I can do this! The development and calibration of children's expectations for success and competence beliefs. *Developmental Review*, *48*, 24–39. https://doi.org/10.1016/j.dr.2018.04.001
Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development*, *49*(3), 800–814. https://doi.org/10.2307/1128250
Nicholls, J. G., & Miller, A. T. (1984). Reasoning about the ability of self and others: A developmental study. *Child Development*, *55*(6), 1990–1999. https://doi.org/10.2307/1129774
Nyhout, A., & Ganea, P. A. (2019a). Mature counterfactual reasoning in 4-and 5-year-olds. *Cognition*, *183*, 57–66. https://doi.org/10.1016/j.cognition.2018.10.027
Nyhout, A., & Ganea, P. A. (2019b). The development of the counterfactual imagination. *Child Development Perspectives*, *13*(4), 254–259. https://doi.org/10.1111/cdep.12348
Nyhout, A., Henke, L., & Ganea, P. A. (2019). Children's counterfactual reasoning about causally overdetermined events. *Child Development*, *90*(2), 610–622. https://doi.org/10.1111/cdev.12913
Nyhout, A., & Ganea, P. A. (2020). What is and what never should have been: Children's causal and counterfactual judgments about the same events. *Journal of Experimental Child Psychology*, *192*, 104773. https://doi.org/10.1016/j.jecp.2019.104773
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*(5), 1995–2010. https://doi.org/10.1111/cdev.12253
Payir, A., & Guttentag, R. (2016). "It could have been worse": Developmental change in the use of a counterfactual consoling strategy. *Journal of Experimental Child Psychology*, *148*, 119–130. https://doi.org/10.1016/j.jecp.2016.04.001

Payir, A., & Guttentag, R. (2019). Counterfactual thinking and age differences in judgments of regret and blame. *Journal of Experimental Child Psychology*, *183*, 261–275. https://doi.org/10.1016/j.jecp.2019.02.007

Perner, J., Sprung, M., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development*, *19*(2), 179–201. https://doi.org/10.1016/j.cogdev.2003.12.001

Petrocelli, J. V., Seta, C. E., Seta, J. J., & Prince, L. B. (2012). "If only I could stop generating counterfactual thoughts": When counterfactual thinking interferes with academic performance. *Journal of Experimental Social Psychology*, *48*(5), 1117–1123. https://doi.org/10.1016/j.jesp.2012.03.017

Pighin, S., Byrne, R. M., Ferrante, D., Gonzalez, M., & Girotto, V. (2011). Counterfactual thoughts by experienced, observed and narrated events. *Thinking & Reasoning*, *17*(2), 197–211. https://doi.org/10.1080/13546783.2011.561598

Rafetseder, E., Cristi, Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of "nearest possible world". *Child Development*, *81*(1), 376–389. https://doi.org/10.1111/j.1467-8624.2009.01401.x

Rafetseder, E., O'Brien, C., Leahy, B., & Perner, J. (2021). Extended difficulties with counterfactuals persist in reasoning with false beliefs: Evidence for teleology-in-perspective. *Journal of Experimental Child Psychology*, *204*, 105058. https://doi.org/10.1016/j.jecp.2020.105058

Rafetseder, E., & Perner, J. (2010). Is reasoning from counterfactual antecedents evidence for counterfactual reasoning? *Thinking & Reasoning*, *16*(2), 131–155. https://doi.org/10.1080/13546783.2010.488074

Rafetseder, E., & Perner, J. (2012). When the alternative would have been better: Counterfactual reasoning and the emergence of regret. *Cognition & Emotion*, *26*(5), 800–819. https://doi.org/10.1080/02699931.2011.619744

Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives*, *8*(1), 54–58. https://doi.org/10.1111/cdep.12061

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*(3), 389–404. https://doi.org/10.1016/j.jecp.2012.10.010

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, *13*(1), 73–90. https://doi.org/10.1016/S0885-2014(98)90021-1

Rim, S., & Summerville, A. (2014). How far to the road not taken? The effect of psychological distance on counterfactual direction. *Personality and Social Psychology Bulletin*, *40*(3), 391–401. https://doi.org/10.1177/0146167213513304

Robinson, E. J., & Beck, S. (2000). What is difficult about counterfactual reasoning?. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 101–119). Psychology Press.

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203–214. https://doi.org/10.1016/S0887-6177(03)00039-8

Weisberg, D. P., & Beck, S. R. (2010). Children's thinking about their own and others' regret and relief. *Journal of Experimental Child Psychology*, *106*(2-3), 184–191. https://doi.org/10.1016/j.jecp.2010.02.005

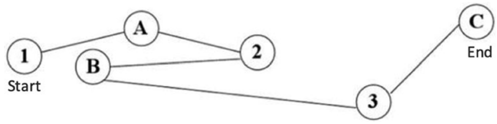
Weisberg, D. P., & Beck, S. R. (2012). The development of children's regret and relief. *Cognition & Emotion*, *26*(5), 820–835. https://doi.org/10.1080/02699931.2011.621933

Appendix

Study 2 instructions

Here we reported the written instructions of Study 2 related to the modified version of the trail-making task. Instructions were read aloud to children, whereas adults read them by themselves. The parts in italics referred to the indefinite condition, whereas the parts in italics reported in brackets concerned the definite condition.

On the sheet you will receive, you will find letters and numbers. You have to connect *as many numbers and letters as possible* [*all the numbers and letters*] in the shortest time possible, alternating numbers and letters in order. The line should follow this path: 1-A-2-B-3-C etc.



You have a time limit (which I won't tell you) to connect *as many items as possible* [*all the items*]. You must not lift the pen off the paper. If you realize you have skipped an item, go back to the previous point, and resume your path. I will keep track of any mistakes and of the number of times you lift the pen off the paper.

Complete the example before beginning! [children only]