



Children's reasoning about other's intentions: False-belief and counterfactual conditional inferences



Célia Rasga^{a,*}, Ana Cristina Quelhas^a, Ruth M.J. Byrne^b

^a William James Center for Research, ISPA-Instituto Universitário, Lisbon, Portugal

^b Trinity College Dublin, University of Dublin, Ireland

ARTICLE INFO

Article history:

Received 4 March 2016
Received in revised form 7 June 2016
Accepted 23 August 2016
Available online 29 August 2016

Keywords:

Intentions
Reasons
False beliefs
Counterfactuals
Reasoning

ABSTRACT

We report four experiments on children's reasoning about intentions using a new change-of-intentions task, in which an observer witnesses an actor carrying out an action, e.g., Mary hears her brother Tom say he wants to switch on the TV to watch a cartoon DVD. Mary goes away and the reason for the action changes, Tom's grandmother tells Tom to switch on the TV to watch the news. The experiments examine reasoning about false beliefs, e.g., What will Mary believe is the reason that Tom is switching on the TV?, and counterfactual reasoning, e.g., If Tom's grandmother hadn't asked Tom to switch on the TV to watch the news, what would have been the reason he was switching it on? Experiment 1 reveals three effects, first, children aged 6 years make more mistakes than those aged 8 years, second, they make more mistakes in false belief than counterfactual reasoning, and third, they make more mistakes for a desire changed to an obligation, compared to an obligation changed to a desire. Experiment 1B shows that the effects also occur for children aged 7 years compared to 9 years. Experiment 2 shows that the effects occur for unfamiliar make-believe content, and Experiment 3 shows that they occur in stories with a simpler structure. The implications for understanding the cognitive processes underlying children's reasoning about intentions are discussed.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

A child who is invited by another child to play on a swing may need to figure out the actor's intention – the actor may wish to be friends, or may expect to gain access to the child's toy, or may be about to trick the child in a prank, or may be following the instruction of an adult. The child may also need to track changes in the actor's intentions – the invitation may have arisen initially because the actor was instructed by an adult, but the actor may now wish to be friends. An important step in understanding other people's actions is reasoning about their intentions (e.g., Grant & Mills, 2011; Juhos, Quelhas, & Byrne, 2015; Walsh & Byrne, 2007). Intentions span a vast array of diverse reasons for actions, including internal reasons such as desires and urges, goals and values, and external reasons such as obligations and rules, social conventions and orders (e.g., Davidson, 1963; Von Wright, 1983). Adults believe that their own actions are determined primarily by their intentions (e.g., Libet, Gleason, Wright, & Pearl, 1983) and they evaluate the intentionality of other people's actions by considering their desires and knowledge (e.g., Malle & Knobe, 1997). Reasoning about people's intentions requires a 'theory of mind', that is, an understanding that others' mental states, such as their beliefs, desires, and knowledge, may differ from one's own (e.g.,

* Corresponding author.

E-mail address: celia.rasga@gmail.com (C. Rasga).

Premack & Woodruff, 1978). We aim to investigate children's reasoning about other people's intentions by examining their inferences that others may have false beliefs about an actor's intentions, and by examining their counterfactual inferences about how the actor's intentions could have been different. The four experiments we report contribute to the discussion about the relation between the development of false belief reasoning and counterfactual reasoning.

1.1. False beliefs and counterfactual inferences

By four to five years of age children understand that others may have false beliefs about the physical world (e.g., Wellman, Cross, & Watson, 2001). For example, in a standard false belief task children are asked to consider two puppets, Sally and Anne, who are in the kitchen; Sally places some chocolate in the cupboard, and she leaves; Anne takes the chocolate and moves it to the fridge; Sally returns. They are asked, where will Sally look for the chocolate? Children aged 3 years usually say Sally will look in the fridge, some children at the age of 4 years and most by the age of 5 years say Sally will look in the cupboard (e.g., Wimmer & Perner, 1983). Understanding that others may have false beliefs is an important milestone that marks children's ability to distinguish between the mental and physical world (e.g., Miller, 2009). It has been tested extensively for one's own and other's false beliefs, in various situations (e.g., Bloom & German, 2000; Baron-Cohen, Leslie, & Frith, 1985).

Children also develop the ability to reason about hypothetical situations and in particular to think about things that didn't happen. Counterfactual inferences, such as 'if Anne had not moved the chocolate, where would it be?' have been proposed to play an important role in the development of reasoning about other's false beliefs such as 'where does Sally think the chocolate is?' (e.g., Riggs, Peterson, Robinson, & Mitchell, 1998). Understanding false beliefs is correlated with counterfactual thinking (e.g. Peterson & Riggs, 1999; Riggs et al., 1998; see also Robinson & Beck, 2000), even when age, verbal intelligence, and other linguistic factors are controlled (e.g., Guajardo, Parker, & Turley-Ames, 2009; Müller, Miller, Michalczyk, & Karapinka, 2007; see also Perner, Sprung, & Steinkogler, 2004). The two sorts of inferences activate similar brain areas (e.g., Van Hoock et al., 2014) and children with autism exhibit difficulties with both false belief and counterfactual reasoning (e.g., Grant, Riggs, & Boucher, 2004; Peterson & Bowler, 2000; Scott, Baron-Cohen, & Leslie, 1999). Children may develop 'mindreading' abilities by deploying reasoning strategies that depend on counterfactual thoughts (e.g., Peterson & Riggs, 1999), including the ability to add or delete events from a representation of reality (e.g., Guajardo & Turley-Ames, 2004). They understand that Sally will think the chocolate is still in the cupboard because they can think that if Anne *had not* moved the chocolate to the fridge, it *would still* be in the cupboard. Counterfactual reasoning may comprise an important ingredient in false belief reasoning by enabling representational advances such as understanding that propositions refer to, or are about, the real world (e.g., Perner, 2000), or by enabling processing advances such as modifying one's own knowledge of a situation to simulate an alternative that accommodates the perspective of another person (e.g., Carlson & Moses, 2001; Peterson & Riggs, 1999).

However, even if counterfactual reasoning is necessary for false belief reasoning, it does not appear to be sufficient. Children's development of counterfactual reasoning begins to emerge early at 2–3 years but continues to develop throughout middle childhood even to young adolescence (e.g., Beck, Robinson, Carroll, & Apperly, 2006; Guttentag & Ferrell, 2004; Rafetseder, Schwitalla, & Perner, 2013; see also Beck & Riggs, 2014; Rafetseder & Perner, 2014). Since children aged 3 years can reason about simple counterfactual situations, such as simple causal and spatial inferences (e.g., German & Nichols, 2003; Harris, 2000; Perner et al., 2004) and yet still fail false belief tasks, their difficulties may arise from a third source that affects both false belief and counterfactual reasoning, such as executive function skills (e.g., German and Nichols, 2003; Guajardo et al., 2009; Müller et al., 2007; see also Beck, Riggs, & Gorniak, 2009). Counterfactuals such as 'if Anne had not moved the chocolate, it would still be in the cupboard' require reasoners to envisage two possibilities, the counterfactual conjecture 'Anne did not move the chocolate and it is still in the cupboard' and the presupposed or known facts 'Anne moved the chocolate and it is not in the cupboard' (see Byrne, 2016 for a review). Counterfactual reasoning and false belief reasoning both require executive function skills, including working memory e.g., holding in mind two representations simultaneously (e.g., Carlson & Moses, 2001; Müller et al., 2007), inhibitory control, e.g., suppressing attention to one representation, such as setting aside what is known about reality (e.g., Leslie, 1987; Robinson & Beck, 2000), and representational flexibility e.g., considering different perspectives about the same situation (e.g. Drayton, Turley-Ames, & Guajardo, 2011; Müller et al., 2007). But false belief reasoning may require further skill at tracking the status of each possibility as corresponding to each person's belief, the counterfactual conjecture 'Anne did not move the chocolate and it is still in the cupboard' corresponds to Sally's belief, and the presupposed or known facts, 'Anne moved the chocolate and it is not in the cupboard' corresponds to the child's knowledge of the situation.

Untangling the relationship between counterfactual reasoning and false belief reasoning has been compromised by an acknowledged limitation in previous studies, that reasoning about false beliefs has required participants to consider another person's mental states, e.g., 'where does Sally think the chocolate is?' whereas reasoning counterfactually has required them to consider only physical states, e.g., 'if Anne had not moved the chocolate where would it be?', and so the counterfactual question removes any mentalistic component of belief (e.g., Peterson & Riggs, 1999). Accuracy in first-order false belief tasks is correlated with accuracy even in an unrelated counterfactual reasoning task, for counterfactuals about either physical states or mental states, i.e., emotions (e.g., Guajardo et al., 2009). However, unaccounted variance in correlations between counterfactual and false belief reasoning has been attributed to false belief tasks making reference to mental states, unlike counterfactuals in first-order tasks (Guajardo & Turley-Ames, 2004). To remedy this discrepancy, we devised a novel change-

of-intentions task that required individuals to consider another person's mental states not only when they reasoned about their false beliefs but also when they reasoned about counterfactuals.

1.2. Change of intentions task

We report the results of four experiments in which children heard a variety of short stories such as the following: 'In the living room, Mary hears Tom say he wants to switch on the TV to watch a cartoon DVD. Mary goes into the bedroom. But while she is away, Tom's grandmother tells Tom to switch on the TV to watch the news. When Mary comes back into the living room, she sees Tom switching on the TV.' The stories referred to a range of everyday activities, such as picking up toys, opening a cupboard, using a spoon, looking in a backpack, switching on TV, writing on paper, going to the fridge, or walking to a lake (see the Supplemental materials). Each story was read to the child and illustrated by pictures, as the example in Fig. 1(A) shows.

In the example, Tom's reason for switching on the TV was initially an internal desire or goal (he wants to watch a cartoon DVD), but subsequently his reason is an externally imposed obligation (his grandmother tells him to switch on the news). The correct answer to a current-reason question, 'What is the reason that Tom is switching on the TV?' refers to the obligation reason, e.g., because his grandmother told him to, to watch the news. The correct answer to a false-belief question, 'What does Mary believe is the reason that Tom is switching on the TV?' refers to the desire reason, e.g., because he wants to watch a cartoon DVD. The correct answer to the counterfactual question, 'If Tom's grandmother hadn't told Tom to switch on the TV to watch the news, what would have been the reason that Tom was switching on the TV?' also refers to the desire reason. The new change-of-intentions task allows a comparison of false belief reasoning about mental states with counterfactual reasoning about mental states in a single task about the same content.

The change-of-intentions task is a second-order false belief task. A crucial step in the development of a theory of mind is reasoning about higher order false beliefs, such as a person's beliefs about another person's beliefs (e.g., Miller, 2009). Consider John and Mary, who are both independently told about the unexpected transfer of an object to a new location, e.g., an ice-cream truck moves to another road. Hence both know where the object is, but John mistakenly believes that Mary thinks the object is at its original location. Not until the ages of 6–7 years do children begin to show accuracy in answering the question 'Where does John think Mary will go for ice-cream?' (Perner & Wimmer, 1985). Similarly, 5 year old children make mistakes about false beliefs when the entity that changes location is a person, especially when the change of location is brought about by the person's own decision rather than by another's instruction (e.g., Rai & Mitchell, 2004; Symons, McLoughlin, Moore, & Morine, 1997). They also make mistakes about interpretative beliefs, such as one person believing the soup tastes good and another believing that it tastes bad (e.g., Carpendale & Chandler, 1996). Although the change in false belief understanding that occurs at 4–5 years is viewed as a landmark (e.g., Miller, 2009; Wellman et al., 2001), the development of a theory of mind continues from infancy to adolescence. Important precursors emerge in infancy, e.g., infants look at where an actor mistakenly thinks to find an object and they spontaneously act to help an actor, exhibiting an implicit understanding of false beliefs (e.g., Clements & Perner, 1994; see also Saxe, Tenenbaum, & Carey, 2005; Wellman, Lopez-Duran, LaBounty, & Hamilton, 2008). And even adolescents and adults experience difficulty in advanced theory of mind tasks such as interpreting instructions to move objects on a shelf from the perspective of a director who cannot see all of them (e.g., Dumontheil, Apperly, & Blakemore, 2010; see also Devine and Hughes, 2012; Epley, Morewedge, & Keysar, 2004; Keysar, Lin, & Barr, 2003). Reasoning about second-order beliefs (inferring a person's beliefs about another person's beliefs, knowledge, intentions, and so on recursively) may be as significant a developmental landmark as reasoning about first-order epistemic beliefs (inferring a person's beliefs about the physical world). However, it has received considerably less attention (e.g., Miller, 2009).

We examine this significant milestone in the new change-of-intentions task. First, we assess the developmental trajectory of children's reasoning about other's beliefs about intentions from the ages of 6 years through 7, 8 and 9 years. We take as our starting point children aged 6 years. At this age children begin to show accuracy in reasoning about a person's beliefs about another person's beliefs about the location of objects (e.g., Miller, 2009; Perner & Wimmer, 1985); but reasoning about a person's beliefs about another person's reasons for their actions may emerge somewhat later, since it requires reasoning wholly about mental states, without reference to the physical location of objects. Second we also examine their counterfactual reasoning about other people's intentions. We expect that younger children will make more correct counterfactual than false belief inferences even when both are about mental states. The counterfactual question provides helpful supports to aid children's tracking of the change of reasons by referencing the current facts in its antecedent e.g., 'If Tom's grandmother hadn't told Tom to switch on the TV to watch the news. . . ' prior to asking about the counterfactual possibility ' . . . what would have been the reason that Tom was switching on the TV?', whereas the false belief question, 'What does Mary believe is the reason that Tom is switching on the TV?' requires children to retrieve the alternative possibility.

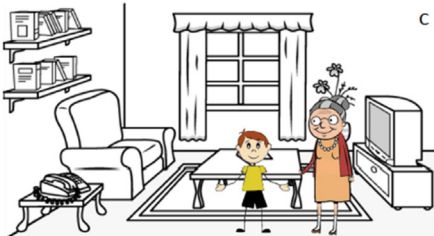
Third, we devised the new change-of-intentions task to also enable us to compare children's reasoning about changes from internal reasons such as goals and desires to external reasons such as obligations and social norms as illustrated in Fig. 1(A), compared to the opposite: reasoning about changes from external reasons to internal ones, as illustrated in Fig. 1(B). Children develop an understanding of other people's desires, beliefs and goals and gradually learn to distinguish between them (e.g., Feinfield, Lee, Flavell, Green, & Flavell, 1999; Gopnik & Slaughter, 1991; Perner & Roessler, 2010). By 5 years, children consider that a person's behaviour will be influenced by external norms, more so than by their preferences (e.g., Kalish & Shiverick, 2004). Children aged 4 years make more mistakes than those aged 7 years when they consider what an

A: Desire changed to Obligation

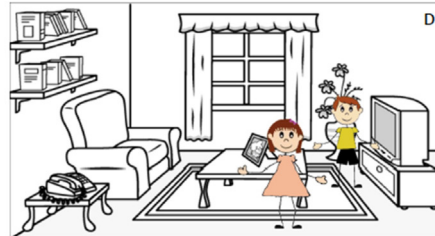
In the living room, Mary hears Tom say he wants to switch on the TV to watch a cartoon DVD.



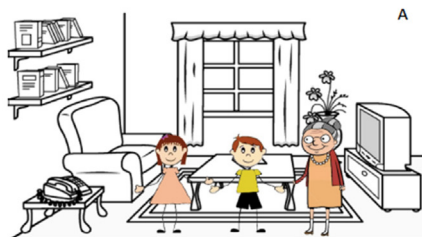
Mary goes into the bedroom.



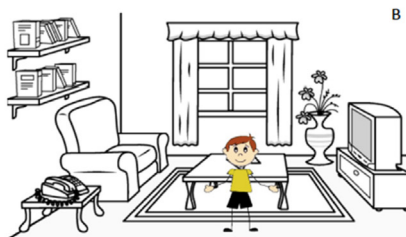
But while she is away, Tom's grandmother tells Tom to switch on the TV to watch the news.



When Mary comes back into the living room, she sees Tom switching on the TV.

B. Obligation changed to Desire:

In the living room, Mary hears Tom's grandmother tells Tom to switch on the TV to watch the news



Mary goes into the bedroom.



But while she is away, Tom say he wants to switch on the TV, to watch a cartoon DVD.



When Mary comes into the living room, she sees Tom switching on the TV.

Fig. 1. Examples of the unexpected change in intentions task.

actor *should* do, in a change of locations task in which the rule to put toys in a classroom box was changed to put toys in a locker on a day when a student was absent – they expect the absent student to follow the new rule (e.g., Kalish & Cornelius, 2007). We expected that when young children begin to appreciate the importance of rules and obligations, they may loom so large that they find it difficult to reason accurately about changes in intentions. When young children are asked what John will think is the reason that Anne is picking up toys, they may tend to respond with the obligation reason 'because her mother told her to tidy the room' regardless of whether the original reason was the obligation, and it is now replaced with a desire, to find the ball to play with it, or whether the original reason was the desire, and it is now replaced with the obligation. Hence children may make more errors in the change of intentions task when a desire is replaced by an obligation.

2. Experiment 1

Our first aim was to test children's false-belief reasoning about an observer's beliefs about an actor's intentions in the new change-of-intentions task. We expected that children aged 6 years would make more mistakes than children aged 8 years. Our second aim was to examine their counterfactual reasoning about changes in an actor's intentions. We expected that the younger children would make more mistakes than the older ones, and also that the younger children would make more mistakes in their false belief inferences about intentions than their counterfactual inferences about intentions. Our third expectation was that children would make more mistakes when a desire changed to an obligation, compared to when an obligation changed to a desire, because they will focus on the obligation in their answers.

2.1. Method

2.1.1. Participants

The participants were 48 children recruited from a school, Colégio Pedro Arrupe, in Lisbon. There were 24 children aged 6 years, 13 boys and 11 girls (average age 6 years 5 months, age range from 6 years 4 months to 6 years 10 months); and 24 children aged 8 years, 12 boys and 12 girls (average age 8 years 5 months, age range from 8 years 3 months to 8 years 11 months). The children spoke Portuguese as their first language, they were predominantly from middle-class families, and none had language or learning difficulties or developmental disorders (in this and in each of the subsequent experiments). They were tested only if their parents first provided consent (85% of the parents did so in this experiment, and in the subsequent experiments 96%, 98% and 97% of parents did so) and if the child then agreed to participate (100% of the children did so).

2.1.2. Design

There were two groups of participants, 6 year olds and 8 year olds. Each participant was presented with 8 scenarios, 4 in which the reason for the action was initially a desire which was then changed to an obligation, and 4 in which the reason for the action was initially an obligation which was then changed to a desire. Each participant completed two reasoning tasks – a false-belief task, and a counterfactual task. The design was a 2 (age: 6 years or 8 years) by 2 (reasoning task: false-belief or counterfactual) by 2 (reason-change: desire-to-obligation or obligation-to-desire) design, with repeated measures on the last two factors.

2.1.3. Materials

We constructed 8 scenarios adapted from the change of location task, to examine false beliefs about reasons for actions (see the Supplemental materials for the full set of materials). Each story was presented to the children with pictures, and each one concerned simple actions, the desires were simple wants, e.g., wanting to watch cartoons, and the obligations were simple instructions given by an adult, e.g., the child was told to do his homework.¹ The 8 stories were constructed to comprise a change of reason from a desire to an obligation, and a second matched version of each story comprised a change of reason from an obligation to a desire, as the examples in Fig. 1A and B illustrate. Each participant received 4 desire-to-obligation stories and 4 obligation-to-desire stories. To control for any content effects, half of the participants in each age group received half of the scenarios in the desire-to-obligation condition (contents 1–4), and the other half in the obligation-to-desire condition (contents 5–8), the other half of the participants received the contents in the opposite assignment. Each participant received the 4 scenarios of each condition in a different random order. The pictorial materials were presented on a Macintosh Air laptop using *Open Sesame* software and narrated by the first author in a pre-recording played over headphones. Each picture was presented individually on the computer screen, with each accompanying sentence vocalised simultaneously over headphones. The participants were given three tasks after each story, (1) A current-reason question, e.g., 'What is the reason that Tom is switching on the TV?' (2) a false-belief inference, e.g., 'What does Mary believe is the reason that Tom is switching on the TV?' and (3) a counterfactual inference, e.g., 'If Tom's grandmother hadn't told Tom to switch on the TV, what would have been the reason that Tom was switching on the TV?' The three experimental tasks – current reality question, false belief inference, and counterfactual inference – were presented in a different randomized order to each child.

2.1.4. Screening tests

All children in all of the experiments completed two screening tests. They completed a verbal ability task to ensure that only children within the normal range of verbal ability were included, given that false belief tasks correlate with verbal ability. The vocabulary subtest from the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) was read aloud

¹ The stories were selected from a set of 10 stories piloted for ease of comprehension with 24 children who did not take part in the experiment – 12 children aged between 6 and 7 years and 12 children aged between 8 and 9 years. We excluded two stories on the basis of the results of the pilot study, which contained sentences that appeared to be ambiguous for children.

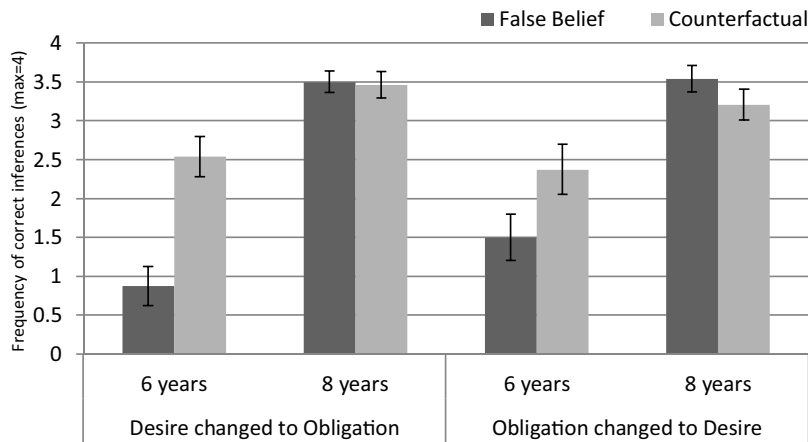


Fig. 2. Frequency of correct inferences by 6 and 8 year olds to false-belief and counterfactual inferences in Experiment 1. Error bars are standard error of the mean.

to children.² The criterion was set that if children scored below 1.3 SD of the norm for their age and gender, they would be excluded from the study, to ensure that all children who participated had at least a low average score of verbal ability. Children were given the test up to the point at which they reached the minimum criteria for inclusion in the current study. In fact, no children in any experiment needed to be excluded on the basis of this test.

They also completed a standard false belief task about a change of location of objects. They were shown puppets and toys and told the puppet's names. They were then asked a naming question (to say each doll's name). The experimenter enacted a scenario in which one puppet, Sally, placed a marble in her basket and left the scene. Another puppet Anne, then took the marble and put it in her own box. Sally then returned. The experimenter said aloud what the puppets were doing: "This is Sally and Anne. Sally has a basket and Anne has a box. Sally put a marble in her basket and then left the room. Anne then moved the marble from Sally's basket to her box. Sally then returned to the room." The child was asked a false-belief question, 'Where will Sally look for her marble?', a current-reality question, 'Where is the marble really?' and a memory question, 'Where was the marble in the beginning?'. All of the children in all of the experiments correctly answered the three questions in this test and so no children were excluded on the basis of this task either.

2.1.5. Procedure

The participants were tested individually in a quiet area of their school. They were asked to look and listen very carefully and they were told that they would be asked some questions about the stories. The two screening tasks were presented in the fixed order of verbal ability task first, and standard change-of-locations task second, and the screening tasks were completed before the participants carried out the change-of-intentions experimental tasks. They provided their responses verbally which were recorded by computer software and also by the experimenter. Participation took about 20 min, and children were given stickers or a pencil and asked not to talk about the questions with other children who had not yet participated.

2.2. Results and discussion

Participants were given a score of 1 if they gave the correct response, and a score of 0 if they gave an incorrect response, or an alternative answer that did not refer to the desire or obligation. A 2 (age: 6 years vs. 8 years) \times 2 (reasoning task: false belief vs. counterfactual) \times 2 (reason-change: obligation-to-desire vs. desire-to-obligation) ANOVA, with repeated measures on the second two factors on the correct responses showed a main effect of age, $F(1, 46) = 46.58, p < 0.001, \eta^2_p = 0.50$, a main effect of reasoning task, $F(1, 46) = 15.84, p < 0.001, \eta^2_p = 0.26$, and no main effect of reason-change, $F < 1$, as Fig. 2 shows. Age interacted with reasoning task, $F(1, 46) = 28.71, p < 0.001, \eta^2_p = 0.38$, reasoning task interacted with reason-change, $F(1, 46) = 5.95, p < 0.02, \eta^2_p = 0.11$, age did not interact with reason-change, $F(1, 46) = 1.23, p = 0.27$ and the three variables did not interact, $F(1, 46) = 1.27, p = 0.27$.

We decomposed the age by reasoning task interaction to show that 6-year olds made fewer correct inferences for the false-belief task than the counterfactual one, $t(23) = 5.198, p < 0.001, d = 1.06$, whereas 8-year olds made as many correct inferences for both tasks $t(23) = 1.567, p = 0.131$, as Fig. 2 shows. The 6-year olds made fewer correct inferences than the 8-year olds for false belief, $t(46) = 9.745, p < 0.001, d = 2.81$, and counterfactual inferences, $t(46) = 2.913, p < 0.006, d = 0.84$. The

² This test is part of the verbal comprehension subscale and measures children's knowledge of words and formation of concepts and shows internal consistency, test-retest reliability, and inter-rater reliability, as well as construct validity by correlating significantly with other intelligence measures (Wechsler, 2003). There are 36 items: four picture items and 32 verbal items.

Table 1

The percentages of correct inferences by children of different ages in each of the four experiments to both the false belief and counterfactual inferences, to neither of them, to the counterfactual only, or to the false belief inference only.

		Both Correct	Neither correct	Counterfactual only	False belief only
E1	6 years	28	36	34	2
	8 years	77	6	5	12
E1B	7 years	51	14	26	
	9 years	95	2	2	2
E2	6 years	35	32	22	11
	8 years	70	13	10	7
E3	6 years	17	60	21	3
	8 years	67	6	16	11

reasoning task by reason-change interaction showed that children made fewer correct inferences for the false-belief task than the counterfactual one in the desire-to-obligation condition, $t(47) = 3.995, p < 0.001, d = 0.58$, but not the obligation-to-desire condition, $t(47) = 1.316, p = 0.194$, and there were no other differences (largest $t = 1.594, p = 0.118$).

To rule out the possibility that the results are due to children misunderstanding the current reason, we carried out a supplementary analysis on only those responses for which the child gave the correct answer to the current-reason question, which corresponded to 98% of the full set of responses. It provided exactly the same results (see the Supplemental materials). Some of the scenarios contained mutually exclusive reasons for which the second reason necessarily replaces the first, e.g., switching on the TV to watch the news replaces switching it on to watch cartoons, but others contained reasons which could be conjunctive, e.g., picking up toys could be done both to tidy a room and to look for a ball. Nonetheless, children answered the false belief and counterfactual questions reliably above chance for every scenario (accuracy from 55% to 76%, binomial p from < 0.05 to < 0.001). To rule out the possibility that the results held for just some scenarios, we carried out a by-materials analysis; it showed exactly the same results as the by-participants one (see the Supplemental materials).

The performance of the children on the false belief task and the counterfactual reasoning task was correlated, $r(48) = 0.592, p < 0.0001$. The 6 year olds tended to give the correct response to both the false belief and the counterfactual task (28% of trials), or to neither (36%), or to only the counterfactual task (34%), they rarely gave the correct response to only the false belief task (2%), as Table 1 shows. The 8 year olds tended to give the correct response to both the false belief and the counterfactual task (77% of trials), they rarely gave the correct response to neither (6%), to only the counterfactual task (5%), or to only the false belief task (12%).

The experiment shows first, children aged 6 years make more mistakes than children aged 8 years in both false belief and counterfactual reasoning about intentions. The result extends the finding that higher order false belief reasoning is accomplished at a later time and as a separate step from epistemic first-order false belief reasoning (e.g., Miller, 2009) to false belief reasoning about intentions. The result also extends the finding that children's counterfactual reasoning continues to make significant developments during middle childhood (e.g., Beck et al., 2006; Rafetseder et al., 2013), to counterfactual reasoning about intentions. Second, children aged 6 years make more correct counterfactual inferences – even when they concern mental states – compared to false-belief inferences. The result extends the finding that children find counterfactual reasoning easier than false belief reasoning (e.g., Riggs et al., 1998) to the new change-of-intentions task in which not only false belief inferences but also counterfactual inferences refer to mental states. Third, children make more errors in false belief than counterfactual inferences when a desire is over-ridden by an obligation, compared to when an obligation is over-ridden by a desire. In the desire changed to obligation condition (e.g., Tom wants to watch a cartoon DVD; then Tom's grandmother tells him to switch on the news), children respond to the false belief inference by referring to the obligation, incorrectly. In the obligation changed to desire condition (e.g., Tom's grandmother tells him to switch on the news; then Tom says he wants to watch a cartoon DVD), children respond to the false belief inference again by referring to the obligation, this time correctly. The tendency to refer to the obligation may be ruled out in the counterfactual inference task, because the counterfactual contains an explicit reminder of the past reason in its antecedent 'what if <the second event> had not happened', and so it rules out the obligation as an answer (e.g., 'what if Tom's grandmother had *not* told him to switch on the TV to watch the news. . .'). We examine alternative explanations of this result in Experiments 3 and 4, but first we report an experiment designed to examine these three new effects in older children aged 7 years compared to 9 years.

3. Experiment 1B

The aim of the experiment was to trace the trajectory of the development of accurate reasoning about intentions further by examining whether children aged 7 years make errors on these tasks, given that 6 year olds made many errors; and conversely, whether children aged 9 years are accurate on these tasks, given that 8 year olds did not perform at ceiling.

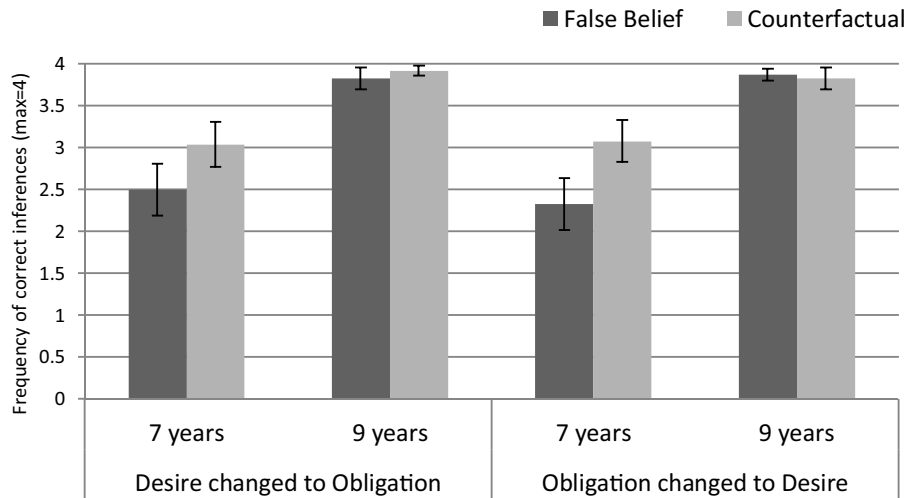


Fig. 3. Frequency of correct inferences by 7 and 9 year olds to false belief and counterfactual inferences in Experiment 1b. Error bars are standard error of the mean.

3.1. Method

3.1.1. Participants

The participants were a new set of 48 children from the Colégio Pedro Arrupe school, 24 7-year-olds, 12 boys and 12 girls (average age of 7 years 5 months, age range of 7 years 1 month to 7 years 11 months); and 24 9 year olds, 11 boys and 13 girls (average age of 9 years 5 months, age range of 9 years 1 month to 9 years 11 months).

3.1.2. Design, materials, and procedure

The design, materials, and procedure were the same as the previous experiment. The participants were asked a current-reason question, and false belief and counterfactual inferences.³

3.2. Results and discussion

A 2 (age: 7 years vs. 9 years) \times 2 (reasoning task: false-belief vs counterfactual), \times 2 (reason-change: obligation-to-desire vs desire-to-obligation) ANOVA, with repeated measures on the second two factors on the correct responses showed a main effect of age, $F(1, 46) = 23.15, p < 0.001, \eta^2_p = 0.34$, and reasoning task, $F(1, 46) = 5.99, p = 0.018, \eta^2_p = 0.12$, and no main effect of reason-change $F < 1$, as Fig. 3 shows, replicating the first experiment. Age interacted with reasoning task, $F(1, 46) = 5.26, p = 0.026, \eta^2_p = 0.10$, but this time not with reason-change, and once again reason-change and reasoning task did not interact, and nor did the three variables, $F < 1$ in each case.

We decomposed the age by reasoning task interaction to find that 7 year olds made fewer correct inferences for the false belief inference compared to the counterfactual one, $t(23) = 2.38, p < 0.03, d = 0.49$, whereas 9-year olds made as many correct inferences on the two tasks, $t(23) = 1, p = 0.328$, replicating the results of Experiment 1, as Fig. 3 shows. The 7-year old children made fewer correct inferences than the 9-year olds on false belief, $t(46) = 4.923, p < 0.001, d = 1.42$, and counterfactual inferences, $t(46) = 3.284, p < 0.002, d = 0.81$.

The performance of the children on the false belief and counterfactual reasoning tasks was correlated, $r(48) = 0.620, p < 0.001$ (see Table 1). Once again we carried out a supplementary analysis on only those responses for which the child gave the correct answer to the current-reason question, which was 96% of the full set of responses, which provided exactly the same results (see the Supplemental materials).

Given that the participants in the two experiments were drawn from similar populations, we carried out a further >0.49 , a main effect of reasoning task, $F(1, 92) = 18.93, p < 0.001, \eta^2_p = 0.17$, and no main effect of reason-change, $F < 1$. Age interacted with reasoning task, $F(1, 92) = 9.69, p < 0.0001, \eta^2_p = 0.24$, reasoning task did not interact with reason-change, $F(1, 92) = 1.70, p = 0.20$, age did not interact with reason-change, $F < 1$ and the three variables did not interact, $F(1, 92) = 1.42, p = 0.24$.

We decomposed the age by reasoning task interaction to show that 6-year olds made fewer correct inferences for the false-belief task than the counterfactual one, $t(23) = 4.656, p < 0.001, d = 0.95$, as did 7-year olds, $t(23) = 2.377, p = 0.026, d = 0.49$,

³ We included a fourth question about the conjectural future, e.g., 'What if next time Tom's grandmother does not tell him to switch on the TV to watch the news, what will be the reason that Tom is switching on the TV?' in Experiments 2, 3 and 4. However, we do not include responses to these questions in the analyses because such conjectural inferences about future intentional relations are often not deterministic, e.g., the reason Tom switches on the TV today is to watch a cartoon DVD, but tomorrow the reason may be to watch a football match, or to look at a music video, or any number of different reasons.

whereas 8-year olds made as many correct inferences for both tasks $t(23) = 1.194, p = 0.245$, as did 9-year olds, $t(23) = 1.000, p = 0.328$. The 6-year olds made fewer correct inferences compared to the 7-year olds for false belief, $t(46) = 3.514, p < 0.001, d = 1.01$, but not for counterfactual inferences, $t(46) = 1.831, p = 0.074, d = 0.53$; they made fewer correct inferences compared to the 8-year olds for both false belief, $t(46) = 10.035, p < 0.001, d = 2.89$ and counterfactual inferences, $t(46) = 3.038, p < 0.004, d = 0.88$, and fewer compared to the 9-year olds for both false belief, $t(46) = 12.794, p < 0.001, d = 3.69$ and counterfactual inferences, $t(46) = 5.423, p < 0.0001, d = 1.57$. The 7-year olds made fewer correct inferences compared to the 8-year olds, for false belief, $t(46) = 3.458, p = 0.001, d = 0.99$, but not counterfactual inferences, $t < 1$, and fewer correct inferences than the 9-year olds for both false belief, $t(46) = 4.923, p < 0.001, d = 1.42$, and counterfactual inferences, $t(46) = 3.284, p = 0.003, d = 0.95$. The 8-year olds made fewer correct inferences than the 9-year olds for false belief, $t(46) = 2.400, p = 0.021, d = 0.69$, and counterfactual inferences, $t(46) = 3.046, p = 0.004, d = 0.88$. Once again we carried out a supplementary analysis on only those responses for which the child gave the correct answer to the current-reason question, which provided exactly the same results (see the Supplemental materials).

The results of the experiment show that even at 7 years children continue to experience difficulty with second order false-belief and counterfactual reasoning about intentions, and they make many more mistakes than children aged 9 years, who perform close to ceiling. Although the children aged 7 years in this experiment performed better than the children aged 6 years in the previous experiment, their performance is still significantly less accurate than children aged 9 years. Second, children aged 7 years, like the 6 year olds in the previous experiment, perform better on counterfactual reasoning – even when it concerns mental states – compared to false-belief reasoning. Third, and importantly, by 7 years of age, children no longer exhibit a difference between the reason changes in the desire-to-obligation condition and the obligation-to-desire condition that were evident in the 6 year old children in the previous experiment. The 7 year olds appear to have overcome the tendency to focus on obligations in answering false belief and counterfactual inferences. It is notable that there is significant improvement in false belief reasoning from 6 years to 7 years, from 7 years to 8 years, and from 8 years to 9 years, with each age group showing greater accuracy than the earlier group; and that the changes in counterfactual reasoning between the four age groups are less pronounced between each age group. Experiment 2 tests whether the results extend to unfamiliar content.

4. Experiment 2

The aim of the experiment was to test whether the results generalise to unfamiliar content. The use of unfamiliar content allows us to rule out the explanation that the differences in the accuracy of false belief and counterfactual inferences in the previous experiments arise because of differences in the children's knowledge or experience rather than their reasoning skills. We modified the stories to refer to a fictional alien planet, Planet Pokron. Children as young as 4 years of age, can reason with make-believe content (e.g., Hawkins, Pea, Glick, & Scrihner, 1984). Pretend make-believe play emerges early and shares some of the cognitive skills required for false belief and counterfactual reasoning (e.g., Harris, Kavanaugh, Wellman, & Hickling, 1993), such as inhibiting the representation of reality (e.g., Richert & Lillard, 2002), and understanding mental representations (e.g., Bruell & Woolley, 1998; but see Harris, Lillard, & Perner, 1994). The stories concerned unfamiliar activities such as, 'The Orange Pok hears the Red Pok say he really wants to plant the kalap in the crater', and the obligations were uttered by figures described merely by their colour, e.g., 'The Purple Pok tells the Red Pok to fit the lesak in the crater'.

In addition, the make-believe content allows us to examine further the source of the tendency in the younger children to focus on obligations in answering false belief and counterfactual inferences. A possible explanation of the finding that children make more errors to the false belief inference than the counterfactual one in the desire-changed-to-obligation condition compared to the obligation-changed-to-desire condition is that they respond simply by describing what the actor child *should* do (e.g., Kalish & Cornelius, 2007). They may do so because the obligations were uttered by known authority figures, such as mothers or fathers, and because they corresponded to familiar norms, such as doing homework or eating fruit. If so, the tendency should be eliminated when the obligations are not uttered by recognizable authority figures and their content is unfamiliar, in the fictional alien scenarios. An alternative explanation is that children perceive obligations to be salient regardless of their familiarity or whomsoever utters them, because obligations require them to think about several possibilities, including what is permitted and what is forbidden. Envisaging more possibilities is costly but it results in children representing more information explicitly, compared to desires (e.g., Quelhas & Byrne, 2003). If so, the results should replicate those of the first experiment. The alien make-believe stories allow a test of these two alternative explanations.

4.1. Method

4.1.1. Participants

The participants were a new set of 48 children from Colégio Pedro Arrupe, in Lisbon. There were 24 6-year-olds, 12 boys and 12 girls (average age of 6 years 5 months, age range of 6 years 1 month to 6 years 11 months); and 24 8-year olds, 11 boys and 13 girls (average age of 8 years 5 months, age range of 8 years 1 month to 8 years 11 months).

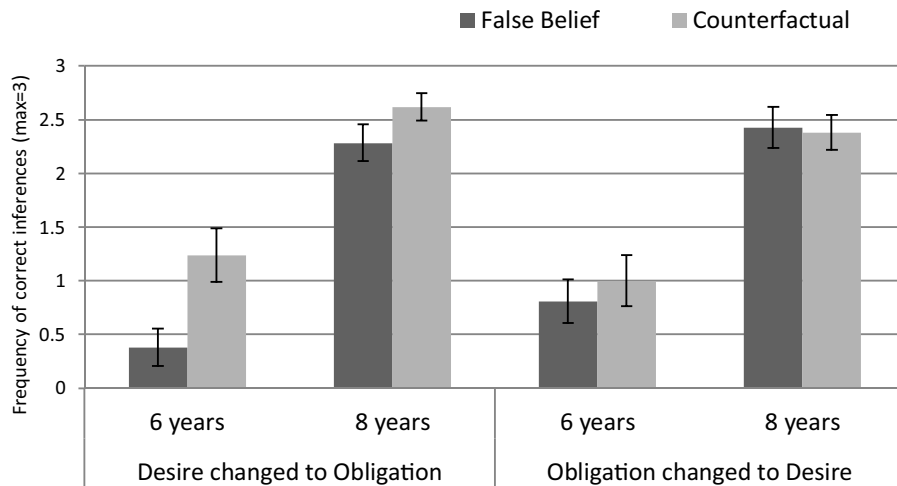


Fig. 4. Frequency of correct inferences by 6 and 8 year olds to the make-believe stories in Experiment 2. Error bars are standard error of the mean.

4.1.2. Materials, design and procedure

The design and procedure was the same as the previous experiments. The materials were 6 stories adapted from those used in the previous experiments to a make-believe planet (see the Supplemental materials). The participants were asked a current-reason question, and false belief and counterfactual inferences.

4.2. Results and discussion

A 2 (age: 6 years vs. 8 years) \times 2 (reasoning task: false belief vs. counterfactual) \times 2 (reason-change: desire-to-obligation vs. obligation-to-desire) ANOVA with repeated measures on the second two factors on the correct responses showed a main effect of age, $F(1, 40) = 108.16, p < 0.0001, \eta^2_p = 0.73$, and reasoning task, $F(1, 40) = 9.19, p = 0.004, \eta^2_p = 0.19$ and no main effect of reason-change, $F < 1$, replicating Experiment 1, as Fig. 4 shows. There was a marginal interaction of age and reasoning task, $F(1, 40) = 3.00, p = 0.091, \eta^2_p = 0.07$, reason-change interacted with reasoning task, $F(1, 40) = 4.36, p = 0.043, \eta^2_p = 0.10$, there was no interaction of reason-change and age, $F < 1$, and the three variables did not interact, $F < 1$.

We decomposed the marginal age by reasoning task interaction to show that 6 year olds made fewer correct inferences on the false belief task compared to the counterfactual one, $t(20) = 3.133, p < 0.005, d = 0.68$, whereas 8-year olds made as many correct inferences on the two tasks, $t(20) = 1, p = 0.329$, replicating the results of Experiment 1, as Fig. 4 shows. The 6 year olds made fewer correct inferences than the 8 year olds on the false belief, $t(40) = 8.92, p < 0.001, d = 2.75$, and counterfactual inferences, $t(40) = 7.86, p < 0.001, d = 2.42$. The reasoning task by reason-change interaction showed that children made fewer correct inferences for the false-belief task than the counterfactual one in the desire-to-obligation condition, $t(41) = 2.993, p = 0.004, d = 0.46$, but not the obligation-to-desire one, $t > 1$, as they had in the first experiment. No other differences were significant.

The performance of the children on the false belief task and the counterfactual reasoning task was correlated, $r(42) = 0.748, p < 0.001$ (see Table 1). Once again we carried out a supplementary analysis on only those responses for which the child gave the correct answer to the current-reason question, which was 94% of the full set of responses, which provided exactly the same results (see the Supplemental materials).

The results replicate each of the findings of the first experiment and extend them to unfamiliar make-believe content. They rule out the explanation that the differences in the accuracy of false belief and counterfactual inferences arise merely because of differences in the children's knowledge or experience rather than their reasoning skills. They corroborate the suggestion that children perceive obligations to be salient regardless of the familiarity of the regulations, or whether they are uttered by a known authority figure. We examine further the tendency of 6 and 8 year olds to focus on obligations in the next experiment.

5. Experiment 3

Children may tend to err on false belief inferences by saying what the actor child *should* do, regardless of what the initial reason is. They may answer with the obligation reason, and so make more incorrect answers to false belief tasks, whenever an initial reason is changed to an obligation, compared to when an initial reason is changed to a desire, irrespective of what the initial reason is. We tested the explanation by modifying the stories used in the first experiment to be simpler, and the initial reason given was a mere 'action plan' in which the character simply describes an action without providing a reason for it, e.g., 'In the bedroom, John hears Anne say she's going to put her ball in her wardrobe. John goes into the kitchen.' For

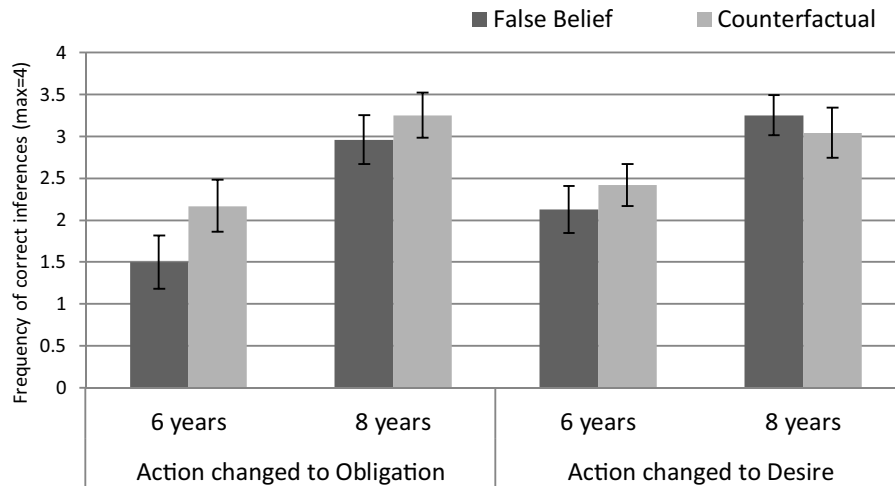


Fig. 5. Frequency of correct inferences by 6 and 8 year olds to the simpler stories in Experiment 3. Error bars are standard error of the mean.

the obligation condition, the story continued with an obligation, e.g., ‘But while he is away, Anne’s mother tells her to fetch her hat from her wardrobe’ and for the desire condition, the story continued with a desire, e.g., ‘But while he is away, Anne says she wants very much to fetch her hat from her wardrobe.’ Both stories concluded in the same way, e.g., ‘When John passes Anne’s bedroom again he sees her near the wardrobe.’ The correct answer to the false belief inference, ‘What will John believe is the reason that Anne is near the wardrobe?’ refers to the action plan, ‘to put her ball in the wardrobe’, in both conditions. If children have a tendency to respond by referring to the obligation, then they will make more incorrect responses in the obligation condition than the desire condition.

An alternative explanation is that children perceive obligations to be stronger than desires and so it is the relation between obligations and desires that gives rise to the tendency in the stories in the previous experiments. If so, there should be no difference in accuracy between the inferences in the obligation and desire conditions because the desire is not changed to an obligation or vice versa. The simplified action-changed-to-obligation/desire stories allow a test of these two alternative explanations.

5.1. Method

5.1.1. Participants

The participants were a new set of 48 children from the Colégio Pedro Arrupe school, 24 6-year-olds, 12 boys and 12 girls (average age of 6 years 6 months, age range of 6 years 4 months to 6 years 10 months); and 24 8 year olds, 10 boys and 14 girls (average age of 8 years 7 months, age range of 8 years 3 months to 8 years 10 months).

5.1.2. Design, materials and procedure

The design and procedure were the same as the previous experiments, and the materials were similar with the following modification: the actor provided an action plan rather than an explicit reason for his or her action in the opening scene. The participants were asked a current-reason question, and false belief and counterfactual inferences.

5.2. Results and discussion

A 2 (age: 6 years vs 8 years) \times 2 (reasoning task: false-belief vs counterfactual) \times 2 (reason-change: action-changed-to-desire vs action-changed-to-obligation) ANOVA, with repeated measures on the second two factors, on the correct responses showed a main effect of age, $F(1, 46) = 12.49$, $p < 0.001$, $\eta^2_p = 0.21$, and this time no main effect of reasoning task for these new stories, $F(1, 46) = 1.598$, $p = 0.21$, and no main effect of reason-change, $F(1, 46) = 2.504$, $p = 0.12$. For these new stories, age did not interact with reasoning task, or with reason-change, $F < 1$ in both cases, but reasoning task and reason-change interacted, $F(2, 92) = 8.186$, $p = 0.006$, $\eta^2_p = 0.15$ as in the first experiment; and the three variables did not interact, $F(2, 92) = 0.167$, $p = 0.685$, as Fig. 5 shows.

We decomposed the reasoning task by reason-change interaction to show that in the obligation condition children made fewer correct inferences to the false-belief inference than the counterfactual one, $t(47) = 2.10$, $p = 0.042$, $d = 0.30$, there was no difference in the desire condition, $t < 1$, as Fig. 5 shows. They made fewer correct inferences in the obligation condition compared to the desire condition, for false belief inferences, $t(47) = 2.812$, $p = 0.007$, $d = 0.41$ but not for counterfactual inferences, $t > 1$. The pattern in this experiment is similar to, although less pronounced than, the pattern in the first experiment, as a comparison of Figs. 2 and 5 show.

The performance of the children on the false belief task and the counterfactual reasoning task was correlated, $r(46) = 0.996$, $p < 0.001$ (see Table 1). Once again we carried out a supplementary analysis on only those responses for which the child gave the correct answer to the current reason question (i.e., 99% of the full set of responses), which provided exactly the same results (see the Supplemental materials).

The experiment shows that children aged 6 years make fewer correct inferences than children aged 8 years, as the first experiment showed, even for these new stories in which a description of an action plan is changed to either an obligation or a desire. The children made more correct counterfactual inferences than false belief inferences in the obligation condition, and they tended to answer the false belief question by saying what the actor child *should* do, regardless of whether the first reason was a desire as in the first experiment, or an action plan as in this experiment. The result suggests that obligations may loom so large in young children's minds that they find it difficult to reason accurately about changes in intentions that concern them.

6. General discussion

How does a child who is invited to play on a swing by another child figure out the other child's intentions, such as that the child wishes to make friends, and track changes in the other child's reasons, such as that the other child initially issued the invitation because her mother instructed her to do so, but now the other child wishes to be friends? The results of the four experiments reported in this paper shed some light on the cognitive processes that develop from the ages of 6 years to 9 years as children acquire new reasoning skills about mental states. The primary contribution of the four experiments is the discovery that younger children aged 6 years make more correct counterfactual inferences compared to false-belief inferences, even when they concern the same sorts of mental states. The result suggests that even when children can make counterfactual inferences about mental states, they must still develop further skills to accomplish 'mind-reading' other people's mental states (e.g., Riggs et al., 1998).

A central component of this discovery is the new change-of-intentions task. It shows that children aged 6 and 7 years make more mistakes than children aged 8 and 9 years in false belief reasoning about intentions. The children in all of the experiments passed the standard false belief task about the unexpected change of location of a physical object. The finding of such a difficulty in false belief reasoning about intentions advances the suggestion that mastery of second-order false belief reasoning, such as reasoning about a person's beliefs about another person's beliefs, constitutes a significant cognitive feat (e.g., Miller, 2009; Perner & Wimmer, 1985). It is a further landmark development that is accomplished at an older age compared to first-order false belief reasoning about a person's beliefs about the physical world. The four experiments indicate that accuracy is not achieved until about 9 years in the case of reasoning about a person's beliefs about another person's reasons for their actions.

Children aged 6 and 7 years also made more mistakes than the older children in counterfactual reasoning about intentions. The result is consistent with the idea that children's counterfactual reasoning about intentions continues to make significant developments during middle childhood (e.g., Beck et al., 2006; Rafetseder et al., 2013). Counterfactual questions, such as, 'if Tom's grandmother had not told him to switch on the TV to watch the news, what would have been the reason he was switching on the TV?' provide an explicit reminder of the current facts and yet 6 and 7 year old children still found them difficult to answer accurately. Importantly, false belief reasoning about intentions was correlated with counterfactual reasoning about intentions in all four experiments. The change-of-intentions task remedies a limitation in previous studies that reasoning about false beliefs required participants to consider another person's mental states, e.g., 'where does Sally think the chocolate is?', whereas reasoning counterfactually removed any mentalistic component of belief by requiring them to consider only physical states, e.g., 'if Anne had not moved the chocolate where would it be?'. The four experiments showed that children were more accurate at counterfactual reasoning than false belief reasoning even when both were about another person's mental states in the novel change-of-intentions task.

Older children may be able to reason about false beliefs, e.g., Mary will think that the reason Tom was switching on the TV was to watch a cartoon DVD, in part because they can reason about counterfactuals, e.g., if Tom's grandmother had not told him to switch on the TV to watch the news, then the reason he was switching on the TV would have been to watch a cartoon DVD. Counterfactuals require reasoners to envisage two possibilities, the conjecture 'Tom's grandmother did not tell him to switch on the TV to watch the news and Tom's reason for switching on the TV is to watch a cartoon DVD' and the presupposed facts 'Tom's grandmother told him to switch on the TV to watch the news and Tom's reason for switching on the TV is not to watch a cartoon DVD' (see Byrne, 2016 for a review). Younger children experience difficulties in envisaging both reality, e.g., a toy mouse running down one slide, as well as its counterfactual alternative, the possibility in which the mouse had gone down the other slide (e.g., Beck et al., 2006); our results suggest that those difficulties may persist beyond the age of 6 years to 7 and even 8 years for envisaging different reasons for actions. Younger children may need to develop working memory skills to ensure that they can envisage two representations simultaneously (e.g., Müller et al., 2007), and to develop inhibitory control to suppress attention to what is known about reality (e.g., Robinson & Beck, 2000). But more fundamentally, younger children may need to make a conceptual leap from physical-causal reasoning about people's actions to teleological reasoning about them (e.g. Perner & Roessler, 2010; see also Atance, Metcalfe, Martin-Ordas, & Walker, 2014). This cognitive feat may be facilitated by the development of counterfactual reasoning to fine-tune skills in modifying a representation of reality to simulate an alternative to it.

An alternative explanation of the observed correlations between false belief and counterfactual reasoning is that the sorts of counterfactual questions used in earlier studies, e.g., ‘if Anne had not moved the chocolate where would it be?’ and in this study, e.g., ‘if Tom’s grandmother had not told him to switch on the TV to watch the news, what would have been the reason he was switching on the TV?’ are merely memory questions that require children to remember the initial location or the actor’s initial reason. If children cannot remember the initial location or reason, then they cannot provide the correct response to the false belief question. However, the data reported here render this explanation somewhat implausible. First, in the screening tests for each experiment, children completed the standard change-of-locations false belief task which included a memory question, ‘Where was the marble in the beginning?’ and none of the children had any difficulty in remembering the initial location, and even the 6 year olds performed at 100% accuracy (see the Supplemental materials). Second, if the younger children’s difficulty in the change-of-intentions task arises from a failure to remember the initial reason, then they should fail equally on the false belief and the counterfactual inference. Instead, although about one third of the younger children gave the correct response to both the false belief and the counterfactual task and about one third to neither, one third of them gave the correct response only to the counterfactual task. Of course, counterfactual thoughts require memory and in fact, implicate brain regions associated with the episodic recollection of specific past experiences, but they also require a comparison of the past situation with an imagined alternative, and so they implicate brain regions associated with the simulation of future events and goals (e.g., Schacter, Benoit, De Brigard, & Szpunar, 2015).

Importantly, children make more errors in false belief than counterfactual inferences about a desire changed to an obligation, compared to about an obligation changed to a desire. Children respond to the false belief inference by referring to the obligation in both cases, as Experiment 1 and 2 shows. By 7 years of age, children no longer exhibit this effect, as Experiment 1B shows. Children aged 6 years tend to answer the false belief question by saying what children *should* do, regardless of whether the first reason was a desire as in Experiment 1, or an action plan as in Experiment 3. The result is consistent with the view that obligations loom large in young children’s minds (e.g., Kalish & Cornelius, 2007). Their salience does not appear to depend on familiarity or authority, as Experiment 2 showed. Instead, it may depend on the explicit representation of information about what is permitted and what is forbidden.

The results of the four experiments are consistent with the idea that higher order false belief reasoning about mental states is accomplished as a separate step from the earlier development of first order false belief reasoning about physical states. They are consistent with the idea that a key cognitive process in children’s development of a ‘theory of mind’ about other people’s mental states, in particular their reasons for their actions, is skill at counterfactual reasoning about mental states.

Acknowledgements

This research was supported by a PhD scholarship from Portugal’s Foundation for Science and Technology (FCT) to Célia Rasga (UID/PSI/04810), and a FCT grant to the William James Center for Research (UID.PSI.04810/2013) to Ana Cristina Quelhas. Some of the research was carried out while the first author was a visiting researcher in Trinity College Dublin for periods during 2012, 2013, and 2014. We thank Csongor Juhos and Marta Couto for their helpful comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.cogdev.2016.08.007>.

References

- Atance, C. M., Metcalf, J. L., Martin-Ordas, G., & Walker, C. L. (2014). Young children’s causal explanations are biased by post-action associative information. *Developmental Psychology*, *50*(12), 2675.
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition*, *21*(1), 7–46.
- Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives*, *8*(3), 175–179.
- 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*, 413–426.
- Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children’s counterfactual thinking and executive functions. *Thinking and Reasoning*, *15*, 337–354.
- Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25–B31.
- Bruell, M. J., & Woolley, J. (1998). Young children’s understanding of diversity in pretense. *Cognitive Development*, *13*, 257–277.
- Byrne, R. M. J. (2000). Counterfactual thought. *Cognition*, *77*(1), B25–B31.
- Bruell, M. J., & Woolley, J. (1998). Young children’s understanding of diversity in pretense. *Cognitive Development*, *13*, 257–277.
- Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*, 135–157.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children’s theory of mind. *Child Development*, *72*, 1032–1053.
- Carpendale, J. I., & Chandler, M. I. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*, 1686–1706.
- Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377–395.
- Davidson, D. (1963). Actions, reasons and causes. *Journal of Philosophy*, *60*, 685–700.
- Devine, R. T., & Hughes, C. (2012). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development*, *84*, 989–1003.
- 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*, 532–548.

- Dumontheil, I., Apperly, I., & Blakemore, S. J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, 13, 331–338.
- Epley, N., Morewedge, C., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40, 760–768.
- Feinfield, K. A., Lee, P. P., Flavell, E. R., Green, F. L., & Flavell, J. H. (1999). Young children's understanding of intention. *Cognitive Development*, 14(3), 463–486.
- German, T. P., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science*, 6(5), 514–523.
- Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development*, 62(1), 98–110.
- Grant, M. G., & Mills, C. M. (2011). Children's explanations of the intentions underlying others' behaviour. *British Journal of Developmental Psychology*, 29(3), 504–523.
- Grant, C. M., Riggs, K. J., & Boucher, J. (2004). Counterfactual and mental state reasoning in children with autism. *Journal of Autism and Developmental Disorders*, 34(2), 177–188.
- 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.
- Guajardo, N. R., Parker, J., & Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology*, 27(3), 681–702.
- Guttentag, R., & Ferrell, J. (2004). Reality compared with its alternatives: Age differences in judgments of regret and relief. *Developmental Psychology*, 40(5), 764–775.
- Harris, P. L., Kavanaugh, R. D., Wellman, H. M., & Hickling, A. K. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, i–107.
- Harris, P. L., Lillard, A. S., & Perner, J. (1994). Commentary: Triangulating pretense and belief. In C. Lewis, & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 287–293). Hillsdale, NJ: Erlbaum.
- Harris, P. L. (2000). *The work of the imagination*. Oxford: Blackwell.
- Hawkins, J., Pea, R. D., Glick, J., & Scribner, S. (1984). Merds that laugh don't like mushrooms: Evidence for deductive reasoning by preschoolers. *Developmental Psychology*, 20(4), 584–594.
- Juhos, C., Quelhas, A. C., & Byrne, R. M. J. (2015). Reasoning about intentions: Counterexamples to reasons for actions. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 41(1), 55–76.
- Kalish, C. W., & Cornelius, R. (2007). What is to be done? Children's ascriptions of conventional obligations. *Child Development*, 78(3), 859–878.
- Kalish, C. W., & Shiverick, S. M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development*, 19, 401–416.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41.
- Leslie, A. (1987). Pretence and representation: The origins of a 'theory of mind'. *Psychological Review*, 94, 412–426.
- Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain*, 106, 623–642.
- Müller, U., Miller, M. R., Michalczyk, K., & Karapinka, A. (2007). False belief understanding: The influence of person, grammatical mood, counterfactual reasoning and working memory. *British Journal of Developmental Psychology*, 25(4), 615–632.
- Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, 33, 101–121.
- Miller, S. A. (2009). Children's understanding of second-order mental states. *Psychological Bulletin*, 135(5), 749–773.
- Perner, J., & Roessler, J. (2010). Teleology and causal reasoning in children's theory of mind. In J. Aguilar, & A. A. Buckareff (Eds.), *Causing human action* (pp. 199–228). Cambridge, MA: MIT Press.
- Perner, J., & Wimmer, H. (1985). John thinks that Mary thinks that... attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, 39, 437–471.
- Perner, J., Sprung, M., & Steinkogler, B. (2004). Counterfactual conditionals and false belief: A developmental dissociation. *Cognitive Development*, 19(2), 179–201.
- Perner, J. (2000). About + belief + counterfactual. In P. Mitchell, & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 367–401). Hove: Psychology Press.
- Peterson, D. M., & Bowler, D. M. (2000). Counterfactual reasoning and false belief understanding in children with autism. *Autism*, 4(4), 391–405.
- Peterson, D. M., & Riggs, K. (1999). Adaptive modeling and mindreading. *Mind and Language*, 4, 80–112.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526.
- Quelhas, A. C., & Byrne, R. M. J. (2003). Reasoning with deontic counterfactual conditionals. *Thinking & Reasoning*, 9(1), 185–197.
- Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives*, 8(1), 54–58.
- Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, 114(3), 389–404.
- Rai, R., & Mitchell, P. (2004). Five-year-old children's difficulty with false belief when the sought entity is a person. *Journal of Experimental Child Psychology*, 89, 112–126.
- Richert, R. A., & Lillard, A. S. (2002). Children's understanding of the knowledge prerequisites of drawing and pretending. *Developmental Psychology*, 38(6), 1004–1015.
- 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.
- Robinson, E. J., & Beck, S. R. (2000). What is difficult about counterfactual reasoning? In P. Mitchell, & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 101–119). Hove, UK: Psychology Press.
- Saxe, R., Tenenbaum, J. B., & Carey, S. (2005). Secret agents inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, 16(12), 995–1001.
- Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of Learning and Memory*, 117, 14–21.
- Scott, F. J., Baron-Cohen, S., & Leslie, A. M. (1999). If pigs could fly: A test of counterfactual reasoning and pretense in children with autism. *British Journal of Developmental Psychology*, 17, 349–362.
- Symons, D., McLaughlin, E., Moore, C., & Morine, S. (1997). Integrating relationship constructs and emotional experience into false belief tasks in preschool children. *Journal of Experimental Child Psychology*, 67, 423–447.
- Van Hoek, N., Begtas, E., Steen, J., Kestemont, J., Vandekerckhove, M., & Van Overwalle, F. (2014). False belief and counterfactual reasoning in a social environment. *Neuroimage*, 90, 315–325.
- Von Wright, G. H. (1983). *Practical reason*. Oxford, UK: Basil Blackwell.
- Walsh, C. R., & Byrne, R. M. J. (2007). How people think if only... about reasons for actions. *Thinking and Reasoning*, 13(4), 461–483.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655–684.
- Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology*, 44(2), 618.
- Wechsler, D. (2003). Wechsler intelligence scale for children -WISC-IV. Psychological Corporation.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.