

Supplementary Material for

“False belief understanding for negative versus positive interactions in children and adults”

Table of Contents

Study 1: Script.....	2
Study 1: Participant Information	4
Study 1: Additional Analyses.....	7
Study 2: Stimuli.....	18
Study 2: Analyses for each individual study version	29
Study 2: Analyses pooling all three versions	38
Study 2: Supplementary information from main text	41

Study 1: Script

Nice Anne Condition

Hi (child's name)! Let me introduce you to some people.

This is Sally (Sally waves), and this is Anne (Anne waves). Sally is playing with her favourite ball (act playing with ball) when her mom calls her out to lunch. Sally has to go, so she runs off, dropping her ball in the basket on the way out (act dropping ball in basket).

Now Anne sees where Sally put her favourite ball and knows that the ball is supposed to go in the closet, and NOT in the basket. Anne thinks that Sally must have been in such a hurry that she put her ball in the wrong place! Anne is a very nice girl, so she wants to help Sally by moving the ball to the closet. Anne goes and gets the ball out of the basket, and to help Sally, puts it away in the closet. That was very nice. Now Sally comes back after lunch and wants to find her ball.

Where will Sally look?

Where does Sally think her ball is?

Should Anne and Sally be friends?

Is Anne a nice girl or not a nice girl?

Is Sally a nice girl or not a nice girl?

Mean Anne Condition

Hi (child's name)! Let me introduce you to some people.

This is Sally (Sally waves), and this is Anne (Anne waves). Sally is playing with her favourite ball (act playing with ball) when her mom calls her out to lunch. Sally has to go, so she runs off, dropping her ball in the basket on the way out (act dropping ball in basket).

Now Anne sees where Sally put her favourite ball and knows that Sally LOVES to play with it. Anne thinks that Sally must have been in such a hurry that she will definitely come back after lunch to play with her ball again. Anne is not a very nice girl, so she wants to trick Sally by moving the ball to the closet. Anne goes and gets the ball out of the basket, and to trick Sally, puts it away in the closet. That was not very nice. Now Sally comes back after lunch and wants to find her ball.

Where will Sally look?

Where does Sally think her ball is?

Should Anne and Sally be friends?

Is Anne a nice girl or not a nice girl?

Is Sally a nice girl or not a nice girl?

Study 1: Participant Information

The final sample consisted of 537 participants. Of the 635 participants that were recruited for Study 1, 98 were excluded due to: participant's age being outside our age range of interest (31), incompleteness of the task or declining to do the task (12), insufficient understanding of English (10), parental/other interference (9), lack of attention to the task (6), experimenter error (4), not understanding the task (4), having previously seen or completed the task (3), improper consent forms (2), having a developmental disorder (2), fussing out (2), or having data noted by the experimenter as unusable (12).

Refer to Tables S1-S4 for more details on our sample: Table S1 for number of participants by Age Category, Gender, and Condition; Table S2 for descriptive statistics regarding age across each condition; Table S3 for difference in age across condition for each Age Category; and Table S4 for exclusions broken down by criteria.

Table S1

Final Sample by Age Category, Gender, and Condition for Study 1

Age	Mean Anne			Nice Anne		
	Female	Male	Total	Female	Male	Total
Three	37	35	72	37	38	75
Four	66	64	130	71	65	136
Five	26	36	62	29	33	62

Table S2*Participant Age Summary Statistics by Condition and Age Category in Study 1*

Age	Mean Anne		Nice Anne	
	Mean	SD	Mean	SD
Three	3.58	0.31	3.53	0.27
Four	4.47	0.28	4.46	0.28
Five	5.49	0.27	5.48	0.28

Table S3*Age comparisons between Mean Anne and Nice Anne Conditions by Age Category in Study 1*

Age	t	df	p	Cohen's d
Three	0.982	140.46	0.328	0.16
Four	0.461	263.56	0.646	0.06
Five	0.200	121.93	0.842	0.04

Table S4*Exclusions by Criteria and Condition in Study 1*

Exclusion Criteria	Total Excluded	Mean Anne	Nice Anne
Age outside of age of interest	31	11	18
Incompletion or declining to do the task	12	7	5
Insufficient understanding of English	10	1	9
Parental / other interference	9	2	7
Lack of attention to the task	6	2	4
Experimenter error	4	3	1
Lack of understanding of the task	4	2	2
Previously saw or completed the task	3	3	0
Improper consent forms	2	1	1
Has developmental disorder	2	0	1
Fussed out	2	2	0
Data marked as unusable but reason was unspecified	12	5	7
Total	98	40	55

Note. The sum of the Mean Anne and Nice Anne columns does not equal Total Excluded because 3 participants were not put in a condition since they did not meet our criteria beforehand but wanted to experience a sample of the task anyway. Two were excluded because of age outside the age of interest; one was excluded because of a developmental disorder.

Study 1: Additional Analyses**Descriptive Statistics for Responses to Questions in Study 1**

In addition to the two main questions, “Where will Sally look?” and “Where does Sally think her ball is?”, we asked, “Should Anne and Sally be friends?”, “Is Anne a nice girl or not a nice girl?”, and “Is Sally a nice girl or not a nice girl?”. Proportion data for the latter three questions are depicted below (Fig. S1).

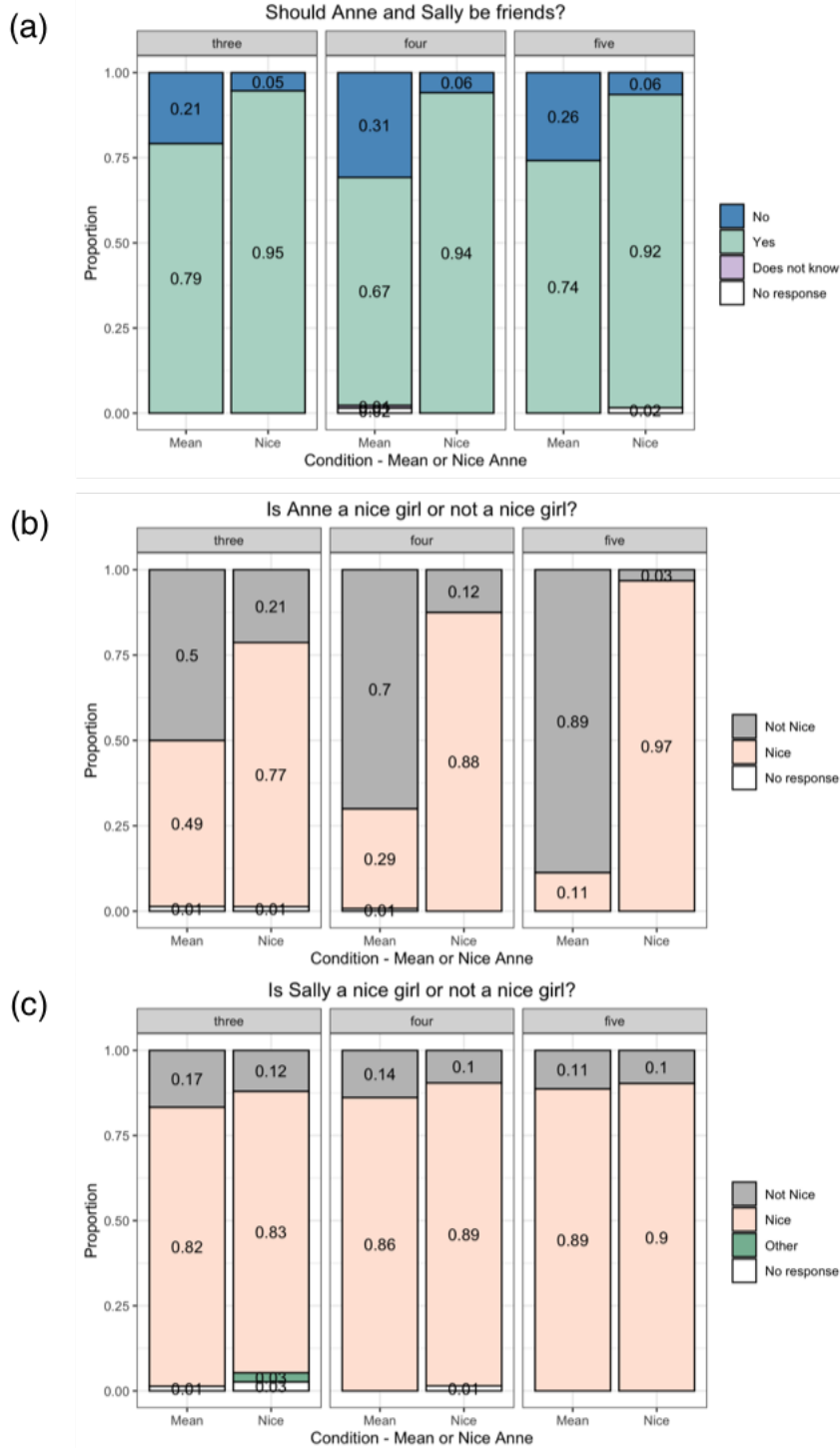


Fig. S1. Proportion of participants by response to auxiliary questions in Study 1

Examining Performance Across Conditions at Different Levels of Other Factors

In the main text, we revealed a main interaction of Condition, with no interactions between Condition and other factors (e.g., Age Category, Question Type, or Counterbalancing Order). Nevertheless, to address questions of how the conditions differed across levels of other factors, we conducted the following contrasts (effect sizes are reported in Table 1 in the main text):

Condition by Age Category

The log odds of responding correctly were numerically greater for *Mean Anne* than *Nice Anne* across each age group, though the difference was not significant for all age groups (3-year-olds: $z = 2.145, p = 0.032$; 4-year-olds: $z = 1.593, p = 0.111$; 5-year-olds: $z = 1.121, p = 0.262$). These results suggest that the general pattern of enhanced performance for *Mean Anne* holds across each age group but is only significant when pooled across age groups.

Condition by Question Type

The log odds of responding correctly were numerically greater in the *Mean Anne* condition than the *Nice Anne* condition for the question “Where does Sally think her ball is?” ($z = 1.627, p = 0.104$) and for the question “Where will Sally look?” ($z = 2.545, p = 0.011$), though the effect was significant only for the latter question. These results suggest that the general pattern of enhanced performance for *Mean Anne* holds across question type, though the size of the effect differs depending on the question.

Condition by Counterbalancing Order

The log odds of responding correctly were numerically greater in the *Mean Anne* condition than the *Nice Anne* condition for the first question presented ($z = 3.160, p = 0.002$) as

well as for the second question ($z = 1.034, p = 0.301$), though the effect was only significant for the first question presented. These results, again, suggest a general effect of Condition.

Other Significant Effects

Counterbalancing Order

There was a significant main effect of Counterbalancing Order ($\chi^2(1) = 28.979, p = 0.007$), wherein performance was better for the second question presented than the first question presented, regardless of which question was presented first.

Breakdown of performance by Age Category and Question Type

There was an interaction between Question Type and Age Category ($\chi^2(2) = 10.863, p = 0.004$): the effect of Question Type differed across the three age groups (Fig. S2). Pairwise contrasts performed at each age group revealed that the log odds of getting the “Where does Sally think her ball is?” question correct was significantly greater than the log odds of getting the “Where will Sally look?” question correct among 4-year-olds ($z = 4.438, p < 0.001$) and 5-year-olds ($z = 3.785, p < 0.001$), but not among 3-year-olds ($z = 0.495, p = 0.62$).

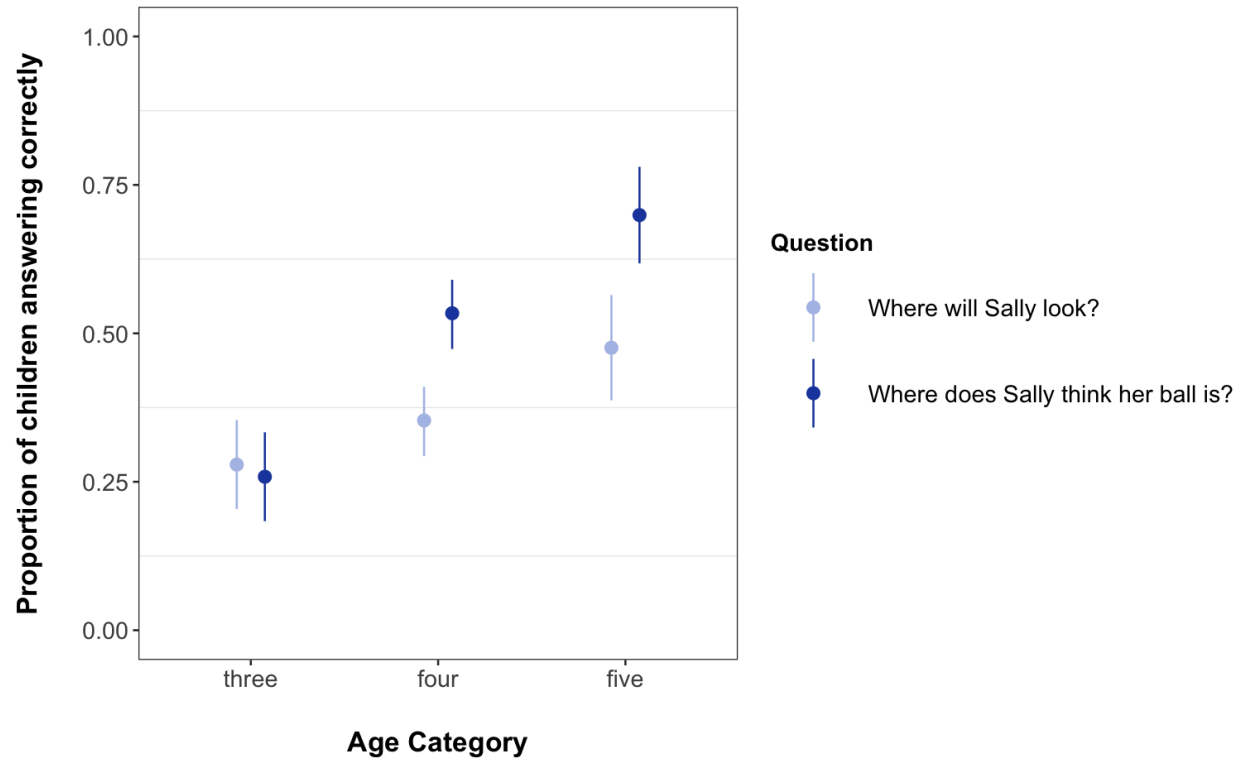


Figure S2. Proportion of participants answering correctly by age group and question. Bars denote bootstrapped 95% CI.

Analyses Limited to Participants Who Responded in a Congruent Manner

The following analyses restricted the data to those who responded to the question “Is Anne a nice girl or not a nice girl?” in a manner congruent to the condition to which they were assigned. This question acted as a comprehension check, but excluding people who did not get this question correct did not change the general pattern of results.

Results were similar to those found in the main text with all usable data. Likelihood ratio tests revealed no significant three-way interaction between Condition, Question Type, and Age Category ($\chi^2(2) = 2.456, p = 0.293$), no significant two-way interaction between Condition and Age Category ($\chi^2(2) = 2.363, p = 0.307$), no significant two-way interaction between Condition and Question Type ($\chi^2(1) = 2.277, p = 0.131$), no significant two-way interaction between Condition and Counterbalancing Order ($\chi^2(1) = 3.373, p = 0.066$), and no significant interaction between Question Type and Age Category ($\chi^2(2) = 5.844, p = 0.054$).

More importantly for our hypotheses, the main effect of Condition remained significant ($\chi^2(1) = 13.857, p < 0.001$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition. Refer to Table S5 for results of this model.

Table S5

Effects from the Full Model in Study 1, from Participants who Responded in a Congruent Manner, with Age as a Categorical Variable

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	0.85	-0.16	0.53 – 1.35	-0.63 – 0.30	-0.69	0.49
Condition [reference: Mean]	0.29	-1.24	0.16 – 0.53	-1.86 – -0.63	-3.96	< 0.001*
Question Type [reference: Look]	1.58	0.46	0.93 – 2.69	-0.08 – 0.99	1.68	0.09
Age Category [linear]	1.9	0.64	0.90 – 4.01	-0.10 – 1.39	1.69	0.09
Age Category [quadratic]	0.84	-0.18	0.46 – 1.52	-0.77 – 0.42	-0.59	0.56
Counterbalancing Order [reference: first]	1.59	0.46	0.96 – 2.62	-0.04 – 0.96	1.8	0.07
Gender [reference: Female]	0.67	-0.4	0.46 – 0.98	-0.78 – -0.02	-2.08	0.04*
Condition x Question Type	1.37	0.31	0.66 – 2.82	-0.41 – 1.04	0.85	0.40
Condition x Age Category [linear]	1.28	0.25	0.46 – 3.53	-0.77 – 1.26	0.47	0.64
Condition x Age Category [quadratic]	1.3	0.26	0.57 – 2.95	-0.56 – 1.08	0.62	0.54
Condition x Counterbalancing Order	1.87	0.62	0.95 – 3.68	-0.06 – 1.30	1.8	0.07
Question Type x Age Category [linear]	1.55	0.44	0.59 – 4.11	-0.54 – 1.41	0.88	0.38
Question Type x Age Category [quadratic]	1.1	0.1	0.51 – 2.40	-0.68 – 0.87	0.25	0.81
Condition x Question Type x Age Category [linear]	2.03	0.71	0.52 – 7.86	-0.65 – 2.06	1.02	0.31
Condition x Question Type x Age Category [quadratic]	0.49	-0.72	0.17 – 1.43	-1.79 – 0.36	-1.31	0.19
Random Effects						
σ^2	3.29					
τ_{00} Subject	0.94					
ICC	0.22					
N_{Subject}	419					
Observations	838					
Marginal R^2 / Conditional R^2	0.201 / 0.379					

Note. Reference levels are provided above. Polynomial functions (linear, quadratic) were fit to the levels of the Age Category variable because Age Category was an ordered factor. CI = 95% confidence interval. * $p < 0.05$, two-tailed

Analyses with Age as a Continuous Variable***With All Usable Data (Data Reported in the Main Text)***

Entering age as a continuous variable did not affect the pattern of results found in the main text. Similar to results reported in the main text, likelihood ratio tests revealed no significant three-way interaction between Condition, Question Type, and Age ($\chi^2(1) = 2.916, p = .088$), no significant two-way interaction between Condition and Age ($\chi^2(1) = 0.903, p = 0.342$), no significant two-way interaction between Condition and Question Type ($\chi^2(1) = 0.446, p = 0.505$), and no significant two-way interaction between Condition and Counterbalancing Order ($\chi^2(1) = 3.510, p = 0.061$). There was, again, a significant interaction between Question Type and Age Category ($\chi^2(1) = 10.290, p = 0.001$): the effect of Question Type differed by age. Importantly, the main effect of Condition remained significant ($\chi^2(1) = 6.593, p = 0.01$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition. Refer to Table S6 for results of this model.

Table S6*Effects from the Full Model in Study 1, with Age as a Continuous Variable: All Participants*

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	0.56	-0.49	0.38 – 0.84	-0.88 – -0.11	-2.50	0.01*
Condition [reference: Mean]	0.39	-0.94	0.23 – 0.66	-1.46 – -0.43	-3.58	<0.001*
Question Type [reference: Look]	1.27	0.29	0.82 – 1.95	-0.13 – 0.72	1.36	0.17
Age (scaled)	1.66	0.44	1.15 – 2.38	0.12 – 0.76	2.73	0.01*
Counterbalancing Order [reference: first]	1.77	0.57	1.15 – 2.72	0.14 – 1.00	2.61	0.01*
Gender [reference: Female]	0.81	-0.21	0.57 – 1.14	-0.55 – 0.13	-1.22	0.22
Condition x Question Type	1.38	0.39	0.74 – 2.55	-0.21 – 0.99	1.27	0.20
Condition x Age (scaled)	0.97	-0.03	0.58 – 1.62	-0.48 – 0.42	-0.12	0.90
Condition x Counterbalancing Order	1.80	0.59	0.98 – 3.30	-0.02 – 1.19	1.90	0.06
Question Type x Age (scaled)	1.44	0.32	0.90 – 2.32	-0.09 – 0.74	1.51	0.13
Condition x Question Type x Age (scaled)	1.56	0.39	0.78 – 3.11	-0.21 – 0.99	1.26	0.21
Random Effects						
σ^2	3.29					
τ_{00} Subject	1.11					
ICC	0.25					
N_{Subject}	537					
Observations	1073					
Marginal R^2 / Conditional R^2	0.180 / 0.387					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

With Data Limited to Participants Responding in a Congruent Manner

Similar to results in the main text, likelihood ratio tests revealed no significant three-way interaction between Condition, Question Type, and Age ($\chi^2(1) = 1.349, p = 0.246$), no significant two-way interaction between Condition and Age ($\chi^2(1) = 1.435, p = 0.231$), no significant two-way interaction between Condition and Question Type ($\chi^2(1) = 3.110, p = 0.078$), and no significant two-way interaction between Condition and Counterbalancing Order ($\chi^2(1) = 3.460, p = 0.063$). There was, again, a significant interaction between Question Type and Age ($\chi^2(1) = 5.622, p = 0.018$): the effect of Question Type differed by age. Importantly, the main effect of Condition remained significant ($\chi^2(1) = 12.511, p < 0.001$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition. Refer to Table S7 for results of this model.

Table S7

Effects from the Full Model in Study 1, with Age as a Continuous Variable: Only Participants who Responded in a Congruent Manner

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	0.79	-0.11	0.49 – 1.26	-0.55 – 0.33	-0.49	0.63
Condition [reference: Mean]	0.28	-1.26	0.15 – 0.52	-1.84 – -0.68	-4.26	<0.001*
Question Type [reference: Look]	1.48	0.45	0.87 – 2.52	-0.05 – 0.95	1.77	0.08
Age (scaled)	1.59	0.41	1.03 – 2.45	0.03 – 0.78	2.12	0.03*
Counterbalancing Order [reference: first]	1.58	0.46	0.96 – 2.60	-0.04 – 0.96	1.79	0.07
Gender [reference: Female]	0.66	-0.42	0.45 – 0.95	-0.80 – -0.05	-2.21	0.03*
Condition x Question Type	1.50	0.54	0.74 – 3.07	-0.14 – 1.21	1.56	0.12
Condition x Age (scaled)	1.05	0.04	0.59 – 1.87	-0.46 – 0.55	0.17	0.87
Condition x Counterbalancing Order	1.90	0.64	0.97 – 3.74	-0.03 – 1.32	1.86	0.06
Question Type x Age (scaled)	1.25	0.19	0.71 – 2.20	-0.30 – 0.69	0.76	0.45
Condition x Question Type x Age (scaled)	1.58	0.40	0.73 – 3.45	-0.28 – 1.08	1.16	0.25
Random Effects						
σ^2	3.29					
τ_{00} Subject	0.88					
ICC	0.21					
N_{Subject}	419					
Observations	838					
Marginal R^2 / Conditional R^2	0.207 / 0.375					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Study 2: Stimuli

1. **Scenario:** Sally put her ball in the box and left the room. Anne came in to the room, moved the ball from the box to the cabinet, and left the room.

Nice: Anne wanted to help Sally by putting the ball in the right place.

Mean: Anne wanted to play a trick on Sally by moving the ball.

Neutral: Anne wanted to leave the room to go play in the park.

Question: While Sally is away, she will think that her ball is in the box.

Answer: True

2. **Scenario:** Lucy put two piles of clothes on her bed: one to keep and one for donation. While Lucy was in the attic to look for more clothes, Christina took the pile Lucy wanted to keep and brought it downstairs.

Nice: Christina thought the pile of clothes she took was meant for donation.

Mean: Christina thought the pile had some really nice clothes she would want for herself.

Neutral: Christina thought the pile contained about twenty shirts.

Question: While Lucy is in the attic, she will think that her keep pile is still on her bed.

Answer: True

3. **Scenario:** Roxanne went to bed, leaving an unfinished puzzle in the living room so that she could solve it the next day. Max saw the puzzle and solved it that night.

Nice: Max wanted to help Roxanne solve the puzzle.

Mean: Max wanted to show Roxanne that he beat her to solving the puzzle.

Neutral: Max wanted to finish the puzzle even if it took the whole night.

Question: When Roxanne wakes up, she will think the puzzle is completed.

Answer: False

4. **Scenario:** Brad spent over 12 hours in the communal living area watching DVDs he borrowed from the library, and he planned on rewatching some of them the next day. While he was sleeping, his roommate Tracy returned all the DVDs to the library.

Nice: Tracy thought Brad was done with the DVDs and did it as a friendly gesture.

Mean: Tracy thought that she could finally have the living room to herself now that the DVDs were gone.

Neutral: Tracy thought it was interesting that all the movies were filmed in black and white.

Question: When Brad wakes up, he will think that the DVDs are still at the house.

Answer: True

5. **Scenario:** Davon cleared his desk and put all his items on the floor because he needed more desk space to do his work. When he left the house to run some errands, Nikki put all the items back on the table.

Nice: Nikki wanted to keep the floor clean so no one would trip on anything.

Mean: Nikki wanted to disrupt Davon's work so that he would be further behind.

Neutral: Nikki wanted to borrow some of the items from Davon later.

Question: While Davon is outside, he will think that his desk is cleared of any items.

Answer: True

6. **Scenario:** Trevor went to the pantry to get spices to make his soup less bland. While Trevor was away, Lauren, the instructor of the culinary class, added unlabeled spice to Trevor's soup.

Nice: Lauren wanted to improve the flavors of the bland soup.

Mean: Lauren wanted to sabotage Trevor by adding unknown spice.

Neutral: Lauren wanted to look at every student's soup.

Question: While Trevor is away, he will think that his soup needs more spice.

Answer: True

7. **Scenario:** Paula forgot the combination to her lock, so she left her locker unlocked and went in to class. Anthony saw the unlocked locker and locked it.

Nice: Anthony wanted to prevent people from stealing Paula's things.

Mean: Anthony wanted to prevent Paula from getting to her belongings.

Neutral: Anthony wanted to get to his next class on time so he walked briskly.

Question: While Paula is in class, she will think that the locker is locked.

Answer: False

8. **Scenario:** Stan needed to get to his fencing match soon, so he asked Duncan to request an Uber while he was getting ready to leave. Instead of requesting an Uber, Duncan requested a Lyft.

Nice: Duncan wanted Stan to get to his match soon and Lyft had an earlier arrival time than Uber.

Mean: Duncan wanted Stan to get to his match late and Lyft had a later arrival time.

Neutral: Duncan wanted to use a promo code for a free ride.

Question: While Stan is getting ready, he will think an Uber will be there.

Answer: True

9. **Scenario:** Amy was told that the results of her medical tests would come in the mail that week. The mail arrived in a timely manner, but her daughter decided to hide the results from Amy for another two weeks.

Nice: Amy's daughter thought Amy had enough stress to deal with.

Mean: Amy's daughter thought that holding on to the results would prevent Amy from receiving treatment.

Neutral: Amy's daughter thought the results were difficult to interpret.

Question: When the week passes by, Amy will think that the medical results were delivered.

Answer: False

10. **Scenario:** Parker moved the ladder close to the air vent and crawled into the vent. Asher moved the ladder from under the vent back to where it typically is.

Nice: Asher thought the ladder was there because someone forgot to return it back.

Mean: Asher thought playing a trick on Parker would be funny.

Neutral: Asher thought it was hard not to scratch the floor when dragging the ladder.

Question: While Parker is in the vents, he will think the ladder is right under the vent.

Answer: True

11. **Scenario:** Stacy went to get some water after her workout at the treadmill, leaving behind her towel. Elena, who wanted to use the treadmill, placed the towel on a nearby elliptical.

Nice: Elena thought that placing the towel there would make it easy for Stacy to find her towel.

Mean: Elena thought moving Stacy's towel would annoy Stacy.

Neutral: Elena thought that she stayed at the gym long enough so she left soon after.

Question: While Stacy is away, she will think that her towel is on the elliptical.

Answer: False

12. **Scenario:** Ashley chopped some vegetables to use for cooking the next day. Her roommate Maggie saw the vegetables in the fridge and decided to cook them while Ashley was at work.

Nice: Maggie wanted to have Ashley's dinner ready for her when she returned.

Mean: Maggie wanted to shorten her cooking time by using Ashley's chopped vegetables.

Neutral: Maggie wanted her meal to contain vegetables and meat.

Question: While Ashley is at work, she will think her vegetables are in the fridge.

Answer: True

13. **Scenario:** Chris was moving out of his furnished apartment and was packing up items in the kitchen. Without any input from Chris, Andrew loaded the desk that came with the apartment into the truck.

Nice: Andrew wanted to help move bulkier items.

Mean: Andrew wanted to get Chris in trouble with the building superintendent.

Neutral: Andrew wanted a Gatorade after moving the desk.

Question: While Chris is in the kitchen, he will think the desk was moved to the truck.

Answer: False

14. **Scenario:** Brittany decided to take golf lessons and planned on signing up for them later that day. Sam signed up for the last spot not knowing that Brittany didn't sign up yet.

Nice: Sam thought that it would be great if he and Brittany took lessons together.

Mean: Sam thought that formal lessons would help him be better at golf than Brittany.

Neutral: Sam thought the lessons were really cheap given the frequency of the lessons.

Question: Before Brittany checks the sign-up sheet, she will think that there are no spots left.

Answer: False

15. **Scenario:** Joe, who was running late that morning, dropped his bag by the entryway and ran to the kitchen to pick up something he forgot. While Joe was in the kitchen, Steve put the bag away.

Nice: Steve thought Joe had left and didn't want the bag to clutter up the entryway.

Mean: Steve thought Joe would get frustrated about losing his bag.

Neutral: Steve thought Joe's bag was really heavy given its size.

Question: While Joe is in the kitchen, he will think his bag is in the entryway.

Answer: True

16. **Scenario:** Janice placed the heavy vase she had just purchased at the yard sale on the floor outside the public restroom. While Janice was in the restroom, Heather moved the vase to the vase section of the yard sale.

Nice: Heather thought the vase had been misplaced so she moved it there.

Mean: Heather thought doing this would really upset Janice.

Neutral: Heather thought the vase section of the yard sale contained a nice variety of different vases.

Question: While Janice is in the restroom, she will think her vase is at the vase section of the yard sale.

Answer: False

17. **Scenario:** Anna wanted to go mountain climbing so she bought some gear. Her husband, David, saw the gear on the table and brought it back to the store while Anna was at work.

Nice: David thought that Anna needed nicer gear for several safety reasons.

Mean: David thought that Anna doing mountain climbing would threaten his masculinity.

Neutral: David thought the store was so disorganized that it was hard to find specific items.

Question: While Anna is at work, she will think that her gear is on the table.

Answer: True

18. **Scenario:** Jeff wanted a nice photo for his dating profile, so he asked Allie to enhance a photo of him. Allie removed the background and added a plain white background.

Nice: Allie wanted to help create a decent photo for what she thought was for a passport.

Mean: Allie wanted to make Jeff's photo look boring and unappealing.

Neutral: Allie wanted to make other alterations but ran short on time.

Question: Before Jeff sees the photo, he will think the photo will have a plain white background.

Answer: False

19. **Scenario:** Erica removed the tire on her bike and went to the bike shop to get a new replacement tire. While she was away, Brian saw the tire and put the tire back on the bike.

Nice: Brian wanted to help Erica out with something that he's good at doing.

Mean: Brian wanted Erica to struggle with getting her bike in working order.

Neutral: Brian wanted to get a new bike like Erica's, but with bigger wheels and a lighter frame.

Question: While Erica is away, she will think the tire is removed from the bike.

Answer: True

20. **Scenario:** Brianna dislikes open umbrellas in indoor spaces, so she closed her wet umbrella and left it by the front door before heading to the living room. Amanda, who came in after Brianna, opened Brianna's umbrella and left it open by the front door.

Nice: Amanda thought opening it would get the umbrella to dry quicker.

Mean: Amanda thought opening the umbrella would annoy Brianna.

Neutral: Amanda thought the blue and orange swirls on the umbrella nicely complemented each other.

Question: While Brianna is in the living room, she will think the umbrella is closed.

Answer: True

21. **Scenario:** Amanda found Mike's paint can in the closet and left it out to be thrown away since the color was ugly. While she was away, Mike saw the paint can and painted their bedroom with it.

Nice: Mike thought it was the can Amanda wanted him to use to paint their room.

Mean: Mike thought the color that he picked was more welcoming than the one Amanda had wanted.

Neutral: Mike thought it would take 30 minutes to paint the room, but it actually took over two hours.

Question: While Amanda is away, she will think the bedroom will be newly painted.

Answer: False

22. **Scenario:** Maria left all of her clothes out as she was packing for vacation and left to get more new clothes. Dani saw that the clothes were out and put them away.

Nice: Dani wanted to help Maria keep her room clean.

Mean: Dani wanted to make packing more difficult for Maria.

Neutral: Dani wanted her sense of style to be as good as Maria's.

Question: While Maria is away, she will think her clothes are packed away.

Answer: False

23. **Scenario:** Maya was working on her portion of a complex multi-person project and took a break by taking a short nap. While Maya was napping, Brianna finished the project.

Nice: Brianna wanted to help Maya with the project.

Mean: Brianna wanted more credit for the work.

Neutral: Brianna wanted to take a nap, too.

Question: When Maya wakes up, she will think the project is finished.

Answer: False

24. **Scenario:** Juan checked his calendar and saw that his schedule was free at noon. While he was making plans to eat out for lunch, his secretary scheduled a meeting for Juan at noon.

Nice: His secretary thought this was a good business opportunity for Juan.

Mean: His secretary thought Juan would get mad for having to skip lunch yet again.

Neutral: His secretary thought the meeting should take place in Meeting Room B.

Question: Before Juan checks his calendar again, he will think that there is a meeting at noon.

Answer: False

25. **Scenario:** Kate, prior to chorus practice, came up with a song and showed it to Grace. While Kate was away, Grace memorized it and showed it to their teacher.

Nice: Grace wanted to get advice from the teacher on how to make Kate's song even better.

Mean: Grace wanted to take the credit for coming up with the song.

Neutral: Grace wanted to hit all the high notes without straining her voice.

Question: While Kate is away, she will think the teacher has already heard her song.

Answer: False

26. **Scenario:** Carla was baking a cake and asked Jane for vegetable oil. While Carla was looking for other ingredients in the pantry, Jane added sesame oil to the cake mix.

Nice: Jane thought there was no difference between sesame oil and vegetable oil.

Mean: Jane thought adding sesame oil to the mix would make Carla's cake taste weird.

Neutral: Jane thought the cake should be done in 40 minutes.

Question: While Carla is away, she will think the cake batter contains sesame oil.

Answer: False

27. **Scenario:** Caroline works as a fry cook flipping hamburgers. Her co-worker, Ali, saw that Caroline walked away from the grill and took the burger off the grill.

Nice: Ali wanted to prevent the burger from burning.

Mean: Ali wanted Caroline to get in trouble for walking away from the grill.

Neutral: Ali wanted to make sure she didn't burn her fingers while moving the burger.

Question: While Caroline is away, she will think the burger is on the grill.

Answer: True

28. **Scenario:** Jackson left his keys at home, so he kept his office door unlocked and open. While he was at a meeting, the janitor closed and locked the office door.

Nice: The janitor thought it was important to keep everything in the office safe.

Mean: The janitor thought locking the door would make it harder for Jackson to get into his office.

Neutral: The janitor thought the office door plaque needed to be updated.

Question: While Jackson is away, he will think his door is unlocked.

Answer: True

29. **Scenario:** Jenny found a nice spot at the party to put down her beer. While she was getting food, Evan moved her beer from the windowsill to the floor.

Nice: Evan thought the wind might knock down the beer from the windowsill.

Mean: Evan thought his beer should be the one by the windowsill instead.

Neutral: Evan thought there was no way the house would comfortably fit everyone.

Question: While Jenny is getting food, she will think that her beer is on the windowsill.

Answer: True

30. **Scenario:** Nicole put down twenty seashells she found at the beach and planned to pick them up after a swim. Jason moved the seashells further away from the water.

Nice: Jason wanted to keep the seashells from getting washed away.

Mean: Jason wanted to make Nicole think the seashells got washed away.

Neutral: Jason wanted to lie on the sand and remain far from the water.

Question: While Nicole is swimming, she will think the seashells are where she dropped them.

Answer: True

Study 2: Analyses for each individual study version

In the main text, we presented analyses pooling the two full versions of Study 2. Here, we break down the results for each study, including an initial version that we did not include in the main text for reasons specified below.

Study 2pre (not reported in main text)

Due to an experimenter error, this initial version of the study prevented participants from seeing the attention check. In the spirit of transparency, we include analyses involving this version of the study here.

We first performed a manipulation check by testing whether ratings of the niceness of the agent differed across conditions. Results reveal a main effect of Condition on ratings of niceness ($\chi^2(2) = 410.52, p < 0.001$). Post-hoc analyses reveal that agents were rated as nicer in the nice condition than in the neutral condition ($z = 13.097, p < 0.001$) and mean condition ($z = 18.829, p < 0.001$). Agents were also rated as nicer in the neutral condition than the mean condition ($z = 9.406, p < 0.001$).

In the full main model of this study, 2.9% of the variance in the response term was explained by the fixed factors in the model, and 35.4% of the variance was explained by both the fixed and random factors in our model. We did not find an effect of Condition on performance on the false belief question ($\chi^2(2) = 1.825, p = 0.402$).

Post-hoc comparisons show no significant pairwise differences in performance on the false belief question between the three conditions (mean versus nice: $z = -1.373, p = 0.355$; mean versus neutral: $z = -0.745, p = 0.724$; nice versus neutral: $z = 0.609, p = 0.815$). Response latencies also did not differ across the three conditions ($\chi^2(2) = 0.320, p = 0.852$).

Please refer to Tables S8-S9 for results of models for Study 2pre.

Table S8

Effects from the Full Model for Study 2pre: Response term is Response on False Belief Question

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	1.59	1.55	0.56 – 4.49	1.04 – 2.06	5.96	< 0.001*
Condition - Nice [reference: Mean]	1.2	0.19	0.92 – 1.57	-0.08 – 0.45	1.37	0.17
Condition - Neutral [reference: Mean]	1.11	0.1	0.85 – 1.44	-0.16 – 0.36	0.77	0.44
Gender – Male [reference: Female]	1.48	0.39	0.85 – 2.58	-0.16 – 0.95	1.38	0.17
Gender – Other [reference: Female]	1.26	0.23	0.11 – 14.75	-2.22 – 2.69	0.19	0.85
Age (scaled)	1.03	0.34	1.01 – 1.06	0.06 – 0.02	2.41	0.02*
Random Effects						
σ^2	3.29					
τ_{00} Subject	1.55					
τ_{00} Item	0.59					
ICC	0.39					
N_{Subject}	189					
N_{Item}	30					
Observations	5670					
Marginal R^2 / Conditional R^2	0.057 / 0.428					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Table S9

Effects from the Full Model for Study 2pre: Response Term is Response Time on False Belief Question

Predictors	Estimates	Std. Beta	CI	Standardized CI	z	p
(Intercept)	2.18	-0.03	-0.52 – 4.88	-0.21 – 0.15	-0.34	0.74
Condition - Nice [reference: Mean]	0.16	0.02	-0.43 – 0.74	-0.06 – 0.11	0.53	0.59
Condition - Neutral [reference: Mean]	0.13	0.02	-0.46 – 0.71	-0.07 – 0.10	0.43	0.67
Gender – Male [reference: Female]	0.22	0.03	-1.29 – 1.74	-0.19 – 0.25	0.29	0.77
Gender – Other [reference: Female]	-1.43	-0.2	-8.29 – 5.44	-1.19 – 0.78	-0.41	0.68
Age (scaled)	0.05	0.07	-0.02 – 0.11	-0.03 – 0.18	1.36	0.17
Random Effects						
σ^2	38.20					
τ_{00} Subject	10.6					
τ_{00} Item	0.18					
ICC	0.22					
N_{Subject}	86					
N_{Item}	30					
Observations	2580					
Marginal R^2 / Conditional R^2	0.006 / 0.225					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Study 2a

We first performed a manipulation check by testing whether ratings of the niceness of the agent differed across conditions. Results reveal a main effect of Condition on ratings of niceness ($\chi^2(2) = 416.53, p < 0.001$). Post-hoc analyses reveal that agents were rated as nicer in the nice condition than in the neutral condition ($z = 19.310, p < 0.001$) and mean condition ($z = 11.912, p < 0.001$). Agents were also rated as nicer in the neutral condition than the mean condition ($z = 11.157, p < 0.001$).

In the full main model of this study, 6.3% of the variance in the response term was explained by the fixed factors in the model and 42.7% of the variance was explained by both the fixed and random factors in our model. We tested for a main effect of Condition on responses to the false belief question. The log odds of providing a correct response was significantly different across the three conditions ($\chi^2(2) = 9.605, p = 0.008$).

Post-hoc analyses not specified in the pre-registration reveal significantly greater log odds of providing a correct response in the nice condition than in the mean condition ($z = 3.0166, p = 0.007$) and significantly greater log odds of providing a correct response in the nice condition than in the neutral condition ($z = 2.383, p = 0.045$) but not significantly different across the mean and neutral conditions ($z = 0-0.652, p = 0.791$). Response latencies, too, differed across the three conditions ($\chi^2(2) = 6.955, p = 0.031$), with response latencies longer for the mean condition than the nice condition ($z = 2.508, p = 0.033$).

Please refer to Tables S10-S11 for results of models for Study 2a.

Table S10*Effects from the Full Model for Study 2a: Response Term is Response on False Belief Question*

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	1.08	2.09	0.32 – 3.60	1.50 – 2.68	6.98	< 0.001*
Condition - Nice [reference: Mean]	1.53	0.43	1.16 – 2.03	0.15 – 0.71	3.02	< 0.001*
Condition - Neutral [reference: Mean]	1.09	0.09	0.84 – 1.43	-0.18 – 0.35	0.65	0.51
Gender – Male [reference: Female]	0.97	-0.03	0.53 – 1.77	-0.64 – 0.57	-0.11	0.91
Age (scaled)	1.06	0.57	1.03 – 1.09	0.27 – 0.87	3.76	< 0.001*
Random Effects						
σ^2	3.29					
τ_{00} Subject	1.56					
τ_{00} Item	0.54					
ICC	0.39					
N_{Subject}	94					
N_{Item}	30					
Observations	2820					
Marginal R^2 / Conditional R^2	0.063 / 0.427					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Table S11*Effects from the Full Model for Study 2a: Response Term is Response Time on False Belief Question*

Predictors	Estimates	Std. Beta	CI	Standardized CI	z	p
(Intercept)	6.59	-0.01	0.55 – 12.62	-0.14 – 0.13	-0.07	0.94
Condition - Nice [reference: Mean]	-2.23	-0.11	-3.97 – -0.49	-0.20 – -0.02	-2.51	0.01*
Condition - Neutral [reference: Mean]	-1.75	-0.09	-3.50 – -0.01	-0.17 – -0.00	-1.97	0.05*
Gender – Male [reference: Female]	2.14	0.11	-1.01 – 5.29	-0.05 – 0.26	1.33	0.18
Age (scaled)	-0.05	-0.03	-0.20 – 0.09	-0.10 – 0.05	-0.72	0.47
Random Effects						
σ^2	370.22					
τ_{00} Subject	41.12					
τ_{00} Item	1.22					
ICC	0.10					
N_{Subject}	94					
N_{Item}	30					
Observations	2820					
Marginal R^2 / Conditional R^2	0.006 / 0.108					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Study 2b

Our manipulation check revealed a main effect of Condition on ratings of niceness ($\chi^2(2) = 573.17, p < 0.001$). There was a greater likelihood of reporting the agent as nice in the nice condition than the mean or neutral condition and greater likelihood in the neutral condition than the mean condition (nice versus neutral: $z = 14.264, p < 0.001$; nice versus mean: $z = 22.283, p < 0.001$; neutral versus mean: $z = 12.618, p < 0.001$).

In the full main model of this study, 5.6% of the variance in the response term was explained by the fixed factors in the model and 43.2% of the variance was explained by both the fixed and random factors in our model. However, we did not find an effect of Condition on performance on the false belief question ($\chi^2(2) = 2.121, p = 0.346$).

Post-hoc comparisons show no significant pairwise differences in performance on the false belief question between the three conditions (mean versus nice: $z = 1.499, p = 0.291$; mean versus neutral: $z = 0.746, p = 0.736$; nice versus neutral: $z = -0.750, p = 0.733$). Response latencies also did not differ across the three conditions ($\chi^2(2) = 1.739, p = 0.419$).

Please refer to Tables S12-S13 for results of models for Study 2b.

Table S12*Effects from the Full Model for Study 2b: Response Term is Response on False Belief Question*

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	2.02	2.61	0.62 – 6.59	2.05 – 3.17	9.16	< 0.001*
Condition - Nice [reference: Mean]	0.8	-0.22	0.60 – 1.07	-0.51 – 0.07	-1.5	0.13
Condition - Neutral [reference: Mean]	0.89	-0.11	0.67 – 1.20	-0.40 – 0.18	-0.75	0.46
Gender – Male [reference: Female]	0.98	-0.02	0.55 – 1.76	-0.60 – 0.56	-0.06	0.95
Age (scaled)	1.05	0.56	1.02 – 1.08	0.26 – 0.86	3.69	< 0.001*
Random Effects						
σ^2	3.29					
τ_{00} Subject	1.52					
τ_{00} Item	0.66					
ICC	0.40					
N_{Subject}	95					
N_{Item}	30					
Observations	2850					
Marginal R^2 / Conditional R^2	0.056 / 0.432					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Table S13*Effects from the Full Model for Study 2b: Response Term is Response Time on False Belief Question*

Predictors	Estimates	Std. Beta	CI	Standardized CI	z	p
(Intercept)	2.56	0.02	-0.70 – 5.82	-0.07 – 0.10	0.37	0.71
Condition - Nice [reference: Mean]	0.98	0.05	-0.76 – 2.71	-0.04 – 0.14	1.1	0.27
Condition - Neutral [reference: Mean]	-0.06	0	-1.79 – 1.67	-0.09 – 0.09	-0.07	0.94
Gender – Male [reference: Female]	-1.08	-0.06	-2.68 – 0.52	-0.14 – 0.03	-1.32	0.19
Age (scaled)	0.04	0.02	-0.03 – 0.12	-0.02 – 0.06	1.13	0.26
Random Effects						
σ^2	371.22					
τ_{00} Subject	2.82					
τ_{00} Item	0.23					
ICC	0.01					
N_{Subject}	95					
N_{Item}	30					
Observations	2850					
Marginal R^2 / Conditional R^2	0.002 / 0.010					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Study 2: Analyses pooling all three versions

Our manipulation check revealed a main effect of Condition on ratings of niceness ($\chi^2(2) = 1417.94, p < 0.001$). There was a greater likelihood of reporting the agent as nice in the nice condition than the mean or neutral condition and greater likelihood in the neutral condition than the mean condition (nice versus neutral: $z = 22.946, p < 0.001$; nice versus mean: $z = 35.271, p < 0.001$; neutral versus mean: $z = 19.362, p < 0.001$).

In the full main model of this study, 4.5% of the variance in the response term was explained by the fixed factors in the model and 40.8% of the variance was explained by both the fixed and random factors in our model. However, we did not find an effect of Condition on performance on the false belief question ($\chi^2(2) = 3.335, p = 0.189$).

Post-hoc comparisons show no significant pairwise differences in performance on the false belief question between the three conditions (mean versus nice: $z = -1.764, p = 0.182$; mean versus neutral: $z = -0.337, p = 0.939$; nice versus neutral: $z = 1.427, p = 0.327$). Response latencies also did not differ across the three conditions ($\chi^2(2) = 1.671, p = 0.434$).

Please refer to Tables S14-S15 for results of models for the three combined versions.

Table S14

Effects from the Full model for Study 2pre, 2a, and 2b Combined: Response term is Response on False Belief Question

Predictors	Odds Ratios	Std. Beta	CI	Standardized CI	z	p
(Intercept)	1.50	2.09	0.74 – 3.02	1.70 – 2.47	10.69	< 0.001*
Condition - Nice [reference: Mean]	1.15	0.14	0.98 – 1.35	-0.02 – 0.30	1.76	0.08
Condition - Neutral [reference: Mean]	1.03	0.03	0.88 – 1.20	-0.13 – 0.18	0.34	0.74
Gender – Male [reference: Female]	1.11	0.11	0.79 – 1.56	-0.23 – 0.45	0.61	0.54
Gender – Other [reference: Female]	0.86	-0.15	0.06 – 11.90	-2.76 – 2.47	-0.11	0.91
Age (scaled)	1.05	0.50	1.03 – 1.06	0.33 – 0.67	5.73	< 0.001*
Random Effects						
σ^2	3.29					
τ_{00} Subject	1.50					
τ_{00} Item	0.52					
ICC	0.38					
N_{Subject}	275					
N_{Item}	30					
Observations	8250					
Marginal R^2 / Conditional R^2	0.045 / 0.408					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Table S15

Effects from the Full Model for Study 2pre, 2a, and 2b Combined: Response Term is Response Time on False Belief Question

Predictors	Estimates	Std. Beta	CI	Standardized CI	z	p
(Intercept)	3.66	0.00	1.20 – 6.13	-0.06 – 0.07	0.10	0.92
Condition - Nice [reference: Mean]	-0.36	-0.02	-1.23 – 0.50	-0.07 – 0.03	-0.82	0.41
Condition - Neutral [reference: Mean]	-0.56	-0.03	-1.43 – 0.30	-0.08 – 0.02	-1.27	0.20
Gender – Male [reference: Female]	0.42	0.02	-0.87 – 1.70	-0.05 – 0.10	0.64	0.52
Gender – Other [reference: Female]	-1.55	-0.09	-11.90 – 8.79	-0.70 – 0.52	-0.29	0.77
Age (scaled)	0.02	0.01	-0.04 – 0.07	-0.03 – 0.05	0.56	0.58
Random Effects						
σ^2	267.06					
τ_{00} Subject	18.67					
τ_{00} Item	0.44					
ICC	0.07					
N_{Subject}	275					
N_{Item}	30					
Observations	8250					
Marginal R^2 / Conditional R^2	0.000 / 0.067					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Study 2: Supplementary information from main text**Table S16**

Effects from the Full Model in Combined Study 2a and Study 2b: Response Term is Response Time to False Belief Question

Predictors	Estimates	Std. Beta	CI	Standardized CI	z	p
(Intercept)	4.44	0.01	1.01 – 7.88	-0.07 – 0.09	0.24	0.81
Condition - Nice [reference: Mean]	-0.6	-0.03	-1.83 – 0.63	-0.09 – 0.03	-0.95	0.34
Condition - Neutral [reference: Mean]	-0.88	-0.04	-2.11 – 0.35	-0.11 – 0.02	-1.41	0.16
Gender [reference: Female]	0.49	0.02	-1.25 – 2.23	-0.06 – 0.11	0.55	0.58
Age (scaled)	0.00	0	-0.08 – 0.08	-0.04 – 0.04	0.03	0.98
Random Effects						
σ^2	371.12					
τ_{00} Subject	22.47					
τ_{00} Item	0.62					
ICC	0.06					
N_{Subject}	189					
N_{Item}	30					
Observations	5670					
Marginal R^2 / Conditional R^2	0.000 / 0.059					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed

Table S17

Effects from the Full Model in Combined Study 2a and Study 2b: Response term is Response to False Belief Question, Removing Participants Performing at Ceiling

Predictors	Odds Ratio	Std. Beta	CI	Standardized CI	z	p
(Intercept)	1.61	1.45	0.81 – 3.20	1.06 – 1.84	7.33	< 0.001*
Condition - Nice [reference: Mean]	1.08	0.08	0.87 – 1.34	-0.14 – 0.29	0.7	0.48
Condition - Neutral [reference: Mean]	1.00	0.00	0.81 – 1.23	-0.21 – 0.20	-0.03	0.97
Gender [reference: Female]	1.09	0.08	0.79 – 1.49	-0.23 – 0.40	0.52	0.61
Age (scaled)	1.03	0.27	1.01 – 1.04	0.12 – 0.42	3.56	< 0.001*
Random Effects						
σ^2	3.29					
τ_{00} Subject	0.61					
τ_{00} Item	0.56					
ICC	0.26					
N_{Subject}	169					
N_{Item}	30					
Observations	3670					
Marginal R^2 / Conditional R^2	0.016 / 0.274					

Note. Reference levels are provided above. * $p < 0.05$, two-tailed