

Belief attribution despite verbal interference

Baudouin Forgeot d'Arc^{1,2}, and Franck Ramus¹

¹Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, CNRS, DEC/ENS, Paris, France

²Child and Adolescent Psychiatry Unit, Robert Debré Hospital, APHP, Paris, France

False-belief (FB) tasks have been widely used to study the ability of individuals to represent the content of their conspecifics' mental states (theory of mind). However, the cognitive processes involved are still poorly understood, and it remains particularly debated whether language and inner speech are necessary for the attribution of beliefs to other agents. We present a completely non-verbal paradigm consisting of silent animated cartoons in five closely related conditions, systematically teasing apart different aspects of scene analysis and allowing the assessment of the attribution of beliefs, goals, and physical causation. In order to test the role of language in belief attribution, we used verbal shadowing as a dual task to inhibit inner speech. Data on 58 healthy adults indicate that verbal interference decreases overall performance, but has no specific effect on belief attribution. Participants remained able to attribute beliefs despite heavy concurrent demands on their verbal abilities. Our results are most consistent with the hypothesis that belief attribution is independent from inner speech.

Keywords: Theory of mind; False-belief task; Language; Belief attribution; Signal detection theory.

The ability to represent conspecifics as agents holding mental representations of the world, such as beliefs, has been labelled “theory of mind” 30 years ago (Premack & Woodruff, 1978) and has since been considered a key feature of human social cognition. It is generally assumed that behaviour prediction and interpretation heuristics in normal human adults can take into account both “desires” and “beliefs” (Leslie, Friedman, & German, 2004). However, little is known about the nature of the process of attributing beliefs to conspecifics. On the one hand, belief

attribution seems inherently propositional (“A believes that B . . .”) and therefore may seem to depend crucially on language and on inner speech as a mediator of verbal reasoning (Carruthers, 2002; de Villiers & de Villiers, 2000). On the other hand, belief attribution is also sometimes thought of as an intuitive, automatic form of reasoning that may be independent of language (Apperly, Samson, & Humphreys, 2009; Leslie et al., 2004). This paper investigates the role of inner speech in the ability to attribute beliefs.

Correspondence should be sent to Baudouin Forgeot d'Arc, Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, CNRS, DEC/ENS, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France. E-mail: forgeot@ens.fr

We thank Coralie Chevalier, Paul Roux, and Inga Vendelin for their help in testing participants, L. Robotham, A. Whitehouse, J. Proust, M. Pessiglione, and V. de Gardelle for suggestions and comments on a previous version of this paper, and APHP/CNRS for financial support.

False-belief (FB) paradigms (Wimmer & Perner, 1983) are among the most widely used tests of belief attribution. They are typically based on short stories with a similar structure. In a classical version (Baron-Cohen, Leslie, & Frith, 1985), an agent ("Sally") hides an object in one location, then leaves. Meanwhile the object is covertly removed and hidden in another location by another agent ("Ann"). The subject's task is to predict where the first agent is going to look for the object. This task has been considered to reflect the ability to predict someone else's action based on his/her false belief, demonstrating the representation of mental states that may differ from the actual state of the world and from the subject's own mental states (Dennett, 1978). Despite a wide use of this paradigm under various versions, in healthy and in clinical populations, in adults and in children, in humans and in animals (Baron-Cohen et al., 1985; Call & Tomasello, 2008; Wimmer & Perner, 1983), it is still debated whether it constitutes an appropriate test of the ability to attribute beliefs (Bloom & German, 2000): Much more is needed to solve the task than attributing beliefs, as shown by the developmental gap between early signs of sensitivity to others' beliefs during the second year of life (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007) and success in the classic FB task in the fourth (Wellman, Cross, & Watson, 2001). In particular, most FB paradigms are based on an explicit narration of the scenario accompanying the target scene and involve asking the subject an overt question (e.g., "Where will Sally look for her marble?"). This suggests that the performance of certain populations in this paradigm might be limited by language and/or communication abilities.

The role of language for reasoning

There is a whole literature on the role of language for reasoning (e.g., Carruthers, 2002; Hermer-Vazquez, Spelke, & Katsnelson, 1999; but see Bek, Blades, Siegal, & Varley, 2010), and a number of studies have investigated the extent to

which various forms of reasoning are mediated by inner speech or covert vocalization (Baldo et al., 2005). Experimental approaches have involved measuring muscular contractions of the articulators while performing reasoning tasks (Sokolov, 1972), studying reasoning in patients with various forms of aphasia (Varley, Siegal, & Want, 2001), and studying healthy participants under conditions in which inner speech is thought to be impossible or more difficult—for example, articulatory suppression (repeating a constant syllable) or verbal shadowing (repeating overheard speech). Results have shown that some forms of reasoning (but not all) seem to be particularly reliant on inner speech. For instance, remembering spatial locations seems unaffected by verbal shadowing if either geometric or nongeometric cues need to be encoded; on the other hand, performance is dramatically affected by shadowing when geometric and nongeometric cues need to be combined to define the location (Hermer-Vazquez et al., 1999). Some reasoning tasks, like the Wisconsin Card Sorting Task, seem to be particularly disrupted by articulatory suppression, although performance remains above chance (Baldo et al., 2005).

The role of language for belief attribution

FB tasks have been shown to be possible despite severe aphasia following brain stroke (Varley et al., 2001). This suggests a relative independence between belief attribution and linguistic processing. On the other hand, in people with autism, performance in FB tasks is generally low (Baron-Cohen et al., 1985). It has been reported that subjects with autism who show good performance "solve theory of mind tasks in an unusually conscious and logical way" and do so "in a verbally mediated fashion" (Happé, 1995). This suggests that reliance on verbal reasoning for belief attribution might in fact be a hallmark of high-functioning autistic individuals, while nonautistic subjects (whether aphasic or not) would by default rely on nonverbal, intuitive processes (Frith, 2003; Happé, 1995). Furthermore, studies showing sensitivity to false beliefs in preverbal infants (e.g.,

Buttelmann et al., 2009; Onishi & Baillargeon, 2005; Southgate et al., 2007) are also consistent with the idea that belief attribution can proceed independently from linguistic abilities, although the interpretation of such studies remains debated (Ruffman & Perner, 2005).

In contrast with this viewpoint, a large meta-analysis of studies investigating the relationship between language ability and belief attribution in children finds that general language abilities have a moderate to large effect on false-belief task performance, even after controlling for age (Milligan, Astington, & Dack, 2007). Furthermore, Newton and de Villiers (2007) recently used verbal shadowing in order to address the role of language in belief attribution in healthy adults. They compared performance in a nonverbal FB task with that in a closely matched true-belief (TB) task, where the object is displaced under the eyes of the agent, thereby leading to a correct belief about object location. They reported dramatic difficulties in the FB, but not in the TB condition, specifically during verbal shadowing, but not during a nonverbal interference task. They concluded in favour of a specific role of language in belief attribution.

In summary, whether one considers studies of belief attribution in normal or in pathological populations, results are not sufficiently consistent to draw conclusions on the role of verbal reasoning for belief attribution.

Interpretation of performance in FB and TB conditions

Historically, FB paradigms have been proposed because interpretation of performance in a TB condition is always ambiguous (Dennett, 1978): Indeed, when participants produce the correct response in a TB paradigm, it is not clear whether they correctly represented the mental state of the agent, or whether their response was directly driven by their own representation of the state of the world (which matches that of the agent). Because FB conditions introduce a mismatch between the agent's and the participant's mental states, correct performance implies that

the participant correctly represented the agent's mental state. However, this should not lead one to believe that FB conditions are an end in itself: Rather, they are a methodological means to dissociate the participant's and the agent's belief and therefore to unambiguously assess the ability to attribute beliefs—that is, assuming that the correct response in the FB condition could not be the result of any response bias. However, there is always a possibility that a participant might be biased to give the correct response for a different reason than the representation of the agent's false belief. This could be a perseveration bias (responding where the agent last stood), or this could be a perceptual bias (preferring one side of the display). For instance, in the classic Sally–Ann paradigm, the basket and the box are not identical objects, so some participants may have a preference for one over the other. More generally, left and right sides are never strictly equal, given that each participant has a dominant hand (to point at the target or press the response button). In the equally classic Smarties task (Perner, Frith, Leslie, & Leekam, 1989), where responses are not lateralized, the correct response (“Smarties”) to the test question (“What does XX think is in the box?”) could be obtained by a bias toward responding what is displayed on the box. Furthermore, the risk of introducing unwanted biases increases dramatically when scenarios are acted live by experimenters, as they often are. Thus, totally unbiased displays are illusory.

TB conditions, to the extent that they preserve all properties of the display but the epistemic status of the agent, would seem to be an appropriate control for the response biases in FB conditions (Newton & de Villiers, 2007; Sommer et al., 2007). This is correct, with the caveat that, because of inherent response biases (albeit identical in the two conditions), participants' raw performance in either condition can never be taken at face value. Only the contrast between the two conditions yields an informative, bias-free result. This is the basis of the paradigm we have designed for the present study.

Rationale of the present study

The design of the present study is based on the fundamental premise that the ability to attribute beliefs cannot be reliably inferred from the ability to attribute a false belief in a given condition, but rather by the ability to systematically modulate one's predictions depending on the epistemic status of the agent (i.e., true vs. false belief). Therefore, this ability is best measured, not by performance in a single condition, but by the specific pattern of performance across conditions manipulating the epistemic status of the agent. Furthermore, because of response biases, it cannot be properly inferred from the simple difference in raw performance between FB and TB conditions, but from *sensitivity* to differences between these two conditions (in the sense of signal detection theory).

Finally, since performance in such conditions may vary depending on many different factors (ability to attribute goals, ability to infer physical causation; e.g., Bloom & German, 2000), it is necessary to independently control these abilities in closely matched conditions (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2007).

Thus we introduce a new paradigm based on a variety of entirely nonverbal scenarios in several matched conditions. We use this paradigm with and without concurrent verbal shadowing, in order to assess the role of inner speech in the attribution of beliefs and goals and in the processing of physical causation. We assume that a sufficiently demanding verbal shadowing task engages phonological representations to such an extent that it is not possible to concurrently formulate propositions using inner speech. Thus, to the extent that reasoning is reliant on inner speech, we assume that it is blocked by verbal shadowing.

Method

Participants

Participants were 58 healthy native French-speaking adults (age: 19–31 years old, mean 22; 26 men, 32 women), naive to the aim of experiment. They gave informed consent and received monetary compensation for participation.

Stimuli

Based on various paradigms—appearance-reality (Gopnik & Astington, 1988), change-of-location (Wimmer & Perner, 1983), reality-unknown (Apperly et al., 2007; Call & Tomasello, 1999)—and control experiments (Zaitchik, 1990), we created 19 different scenarios. Each scenario was subjected to minimal variations in order to create five different conditions for each scenario (see Figure 1 and Appendix).

Each scenario is divided into four successive phases: beginning, change, suspense, and end. The beginning phase (common to all conditions) sets up the general situation and the main agent. In the change phase, the first condition (mentalistic/unseen change = false belief) displays a situation whereby a change occurs unbeknownst to the main agent. The second condition (Mentalistic/Seen change = true belief) displays exactly the same situation and change, but with the agent seeing the change. The third condition (mentalistic/no change) displays the same situation as the first condition, but with no change occurring. The fourth (mechanistic/change) and fifth (mechanistic/no change) conditions are until then identical to mentalistic/unseen change and mentalistic/no change conditions, respectively. In the suspense phase (common to all conditions), the main agent comes to the forefront and is about to perform one of two actions. The end phase displays one of two alternative ends representing either different behavioural outcomes (mentalistic ends) or different physical states (mechanistic ends). In the mentalistic conditions, a correct attribution of mental states (goals and beliefs) leads one to predict the same end for the unseen change and no change conditions, but the opposite end in the seen change condition. Similarly, in the mechanistic conditions, a correct understanding of physical causation leads one to predict opposite ends between the change and no change conditions. Based on these 19 scenarios in five conditions, 95 silent animated cartoons were produced using Macromedia Flash Professional 8 (see <http://www.lscp.net/persons/forgeot/stim/> for examples and detailed descriptions).

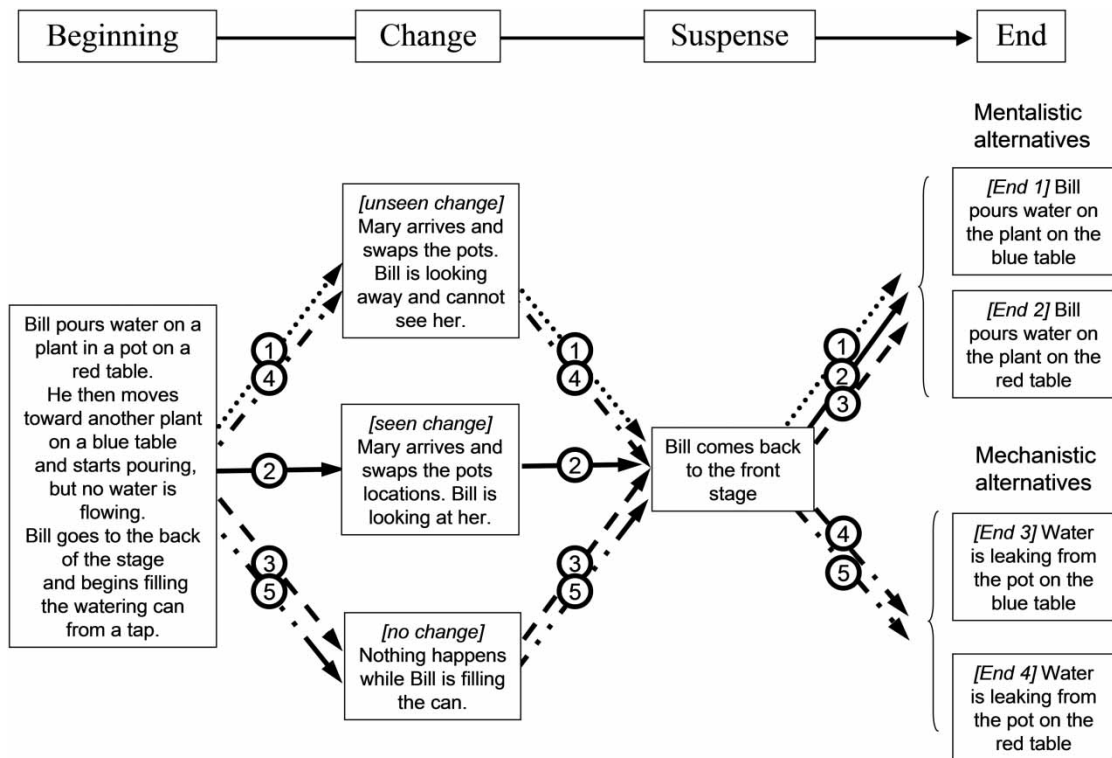


Figure 1. Structure of the animations illustrated with one particular scenario. All five conditions of this scenario share the same “beginning” phase, may vary during the “change” phase, then share the same “suspense” phase. In the “end” phase, subjects have to choose between two possible ends of the story, involving either a behaviour (mentalistic ends) or a physical event (mechanistic ends). Condition 1: mentalistic/unseen change (false belief); Condition 2: mentalistic/seen change (true belief); Condition 3: mentalistic/no change; Condition 4: mechanistic/change; Condition 5: mechanistic/no change.

Procedure

The stimuli were used in a forced-choice paradigm: Subjects were presented each animated film until the suspense scene froze. The two possible ends of the corresponding condition then appeared at the bottom left and bottom right corners of the screen. Subjects were instructed to choose the most plausible end for the story and gave their response on a response box. At the beginning of the test session, they were trained on this forced-choice task using different stimuli involving goals or physical causations but no false beliefs.

Stimuli were randomly grouped into five blocks of 19 trials, each block containing 1 trial per scenario and 3 to 4 trials per condition. Each

participant underwent two blocks out of five, with blocks counterbalanced across participants. Each participant thus underwent a total of 38 trials (6 to 8 per condition), lasting about 30 minutes in total.

Each participant underwent one block in silence and one block while performing a concurrent shadowing task (interference factor). The order of the two blocks was counterbalanced across participants. The stimuli to be shadowed were French sentences with a mean duration of 2.8 seconds and a mean speech rate of 6.1 syllables per second. They were concatenated with a 600-ms interval, which, in pilot trials, appeared to be the minimum duration possible while still allowing subjects to correctly perform the task. Sentences

were played in headphones. Subjects were instructed to continuously shadow the sentences—that is, to immediately repeat each word as it was being spoken. They were trained on the shadowing task immediately before taking the corresponding block, until they were able to perform it fluently for more than three sentences. Before starting the shadowing block, they were informed that correct performance of the shadowing was a condition for their response to be valid. Shadowing performance was coded online by the experimenter who was blind to each stimulus's condition. The experimenter pressed a button each time (and as long as) a participant interrupted the shadowing. Any button press longer than 300 ms was counted as an interruption of the shadowing.

Data analysis

According to our approach, the ability to attribute beliefs is not directly indexed by performance in the mentalistic/unseen change condition, since the choice of a particular end is influenced by various response biases. Rather, the ability to attribute beliefs is better reflected by the ability to give both the correct answer in the mentalistic/unseen change condition and the opposite answer in the mentalistic/seen change condition (given that the stimuli, and therefore the biases, are strictly identical in the two conditions, apart from the epistemic status of the agent). Similarly, the ability to attribute goals is best reflected by giving both the correct answer in the mentalistic/seen change condition and the opposite answer in the mentalistic/no change condition. And the ability to infer physical causation is reflected by giving the correct answer in the mechanistic/change condition and the opposite answer in the mechanistic/no change condition.

In order to disentangle sensitivity from response bias, we performed analyses based on signal detection theory (MacMillan & Creelman, 2005). For belief attribution, we considered as a

hit the choice of the correct end of the story in the mentalistic/unseen change condition and as a false alarm the (incorrect) choice of the same end in the mentalistic/seen change condition. We thus computed a d' (belief sensitivity index). Similarly, we computed a goal sensitivity index based on the mentalistic/seen change and mentalistic/no change conditions and a physical causation sensitivity index based on the mechanistic/change and mechanistic/no change conditions.¹ We also computed criteria (reflecting response biases) for both mentalistic and mechanistic ends, based on mentalistic/unseen change and mentalistic/seen change conditions, and mechanistic/change and mechanistic/no change conditions, respectively.²

Predictions

According to the hypothesis that verbal reasoning is necessary specifically for belief attribution, the belief sensitivity index should fall to chance level ($d' = 0$) under concurrent verbal shadowing. In the case where verbal shadowing is not sufficient to entirely block all linguistic abilities, one would at least expect a significant decrease in this index, compared to the goal and physical causation sensitivity indices (Condition $\times$ Interference interaction). Indeed, according to proponents of the role of language for belief attribution, it is specifically the propositional nature of these mental states that requires language (Carruthers, 2002; de Villiers & de Villiers, 2000). Although these authors make claims only about belief attribution and not about attribution of goals and physical causation, the reasoning they use in the former case does not seem to apply in the latter.

On the other hand, if belief attribution can be performed without specific recourse to verbal reasoning, then the belief sensitivity index may drop because of the demands of the dual task (interference main effect), but not more so than the other two indices (no Condition $\times$ Interference interaction).

¹ $d' = Z(\text{hit rate}) - Z(\text{false alarm rate})$, with Z the inverse function of the normal distribution.

² $\text{Beta} = -Z(\text{false alarm rate})$.

Results

Two scenarios (of the *reality-unknown* type) were excluded due to incompatibility with the analysis procedure in one condition.³ They had been designed for a previous false-belief experiment, and blocks had been left unchanged in the current version. This left 17 scenarios for the analyses. We first conducted the analysis on the complete set of trials. Table 1 displays raw performance and response times on the forced-choice task in the different conditions, during silence and during shadowing.

We performed an analysis of variance (ANOVA) by subject (F_1) on percentage of correct responses, with interference (silence, verbal shadowing) and condition (1 to 5) as independent variables. We found significant effects of interference, $F_1(1, 57) = 58.59$, $p < 10^{-9}$, and condition, $F_1(4, 228) = 7.11$, $p < 10^{-4}$, but no Interference $\times$ Condition interaction, $F_1(4, 228) < 1$. Performance in all 10 conditions was significantly above chance level (all t values > 17.3 ; $p < 2.2 \times 10^{-16}$).

We performed the same analysis of variance on response times. Findings are similar to those on percentage of correct responses, with significant effects of interference, $F_1(1, 57) = 9.54$, $p = .003$, and condition, $F_1(4, 228) = 7.86$, $p = 5.9 \times 10^{-6}$, but no Interference $\times$ Condition interaction, $F_1(4, 228) = 1.92$, $p = .11$. Overall, accuracy was higher and reaction times were shorter in the shadowing than in the silence condition. Note that the profile of performance across conditions cannot be directly interpreted, as it is affected by response bias—hence the need for signal detection analysis.

We now turn to signal detection analyses. We computed belief, goal, and physical causation sensitivity indices as explained above (see Table 2 and

Figure 2). Indices were significantly greater than zero in all conditions (all t values > 7.29 , $p < 10^{-8}$). A first ANOVA by subject, with interference (silence, shadowing) and index (belief, goal, physical causation) as independent variables, revealed significant effects of interference, $F_1(1, 57) = 48.93$, $p < 10^{-8}$, and index, $F_1(2, 114) = 5.83$, $p = .004$, but no Interference $\times$ Index interaction, $F_1(2, 114) < 1$.

Because our scenarios are quite heterogeneous, indices were also computed by item (F_2). Indices were significantly greater than zero in all conditions (all t values > 5.41 , $p < 10^{-4}$). The ANOVA again showed significant effects of interference, $F_2(1, 16) = 82.36$, $p < 10^{-3}$, and index, $F_2(2, 32) = 3.38$, $p < .05$, but no Interference $\times$ Index interaction, $F_2(2, 32) = 2.16$, $p = .13$.

Response criteria were calculated for mentalistic and for mechanistic ends (Table 3), using responses in mentalistic/unseen and seen change, and mechanistic/change and no change conditions, respectively. None of the criteria were significantly different from 1, which indicates that subjects did not have a significant response bias on average across the different scenarios. An ANOVA showed no significant effect of index, $F_2(1, 16) = 1.75$, $p = .20$, and interference, $F_2(1, 16) < 1$, and no interaction, $F_2(2, 32) < 1$. The ANOVA was also run by subject and did not show any effect of index or interference (all F_1 values < 1).

Arguably, the fact that belief, goal, and physical causation attribution remains efficient might be due to insufficient engagement in shadowing. This would have resulted in a trade-off between pause duration in the shadowing task and performance in the forced-choice task. For this reason, we considered separately trials in which the cumulated pause duration was equal or superior to 2,000 ms. A total of 24.2% of the

³The “reality unknown” scenarios were inspired from Apperly et al. (2007). Such scenarios were designed for the purpose of sparing participants the interference between their own belief about the object’s location and the agent’s belief. For example, an agent places a cat in one of two opaque boxes. Due to the agent’s position, the participant cannot see in which of the two boxes the cat is. Depending on the condition, either the two boxes are swapped during the agent’s absence (unseen change) or in presence of the agent (seen change), or the boxes remain at the same place (no change). The agent then heads toward one of the two boxes. In the end scene, the cat may come out of either box. In the “unseen change”, the cat is expected to come out of the box not chosen by the agent, whereas in the other two conditions, the cat is expected to come out of the chosen box.

Table 1. Mean percentages of correct responses and response times on the forced-choice task, in silence and under interference due to verbal shadowing

	<i>Mentalistic conditions</i>						<i>Mechanistic conditions</i>			
	<i>1: Unseen change</i>		<i>2: Seen change</i>		<i>3: No change</i>		<i>4: Change</i>		<i>5: No change</i>	
	%	RT	%	RT	%	RT	%	RT	%	RT
Silence	81.7 (26.9)	6,113 (2,930)	79.7 (23.7)	6,182 (2,766)	93.7 (14.5)	7,332 (3,592)	90.8 (17.5)	5,697 (1,695)	85.3 (19.6)	8,530 (3,602)
Verbal shadowing	70.5 (28.7)	4,655 (1,668)	62.4 (27.5)	5,797 (3,337)	81.0 (24.4)	5,647 (1,860)	71.1 (28.2)	4,561 (1,524)	69.9 (27.0)	6,058 (1,983)

Note: RTs (response times) in ms. Standard deviations in parentheses. All performance values are significantly above chance level (50%).

Table 2. Sensitivity indices for belief, goal, and physical causation in silence and under interference due to verbal shadowing

	<i>Belief</i>	<i>Goal</i>	<i>Physical causation</i>
Silence	1.82	2.32	2.44
Verbal shadowing	1.00	1.28	1.15

Note: All indices are significantly above 0.

trials met this criterion. Overall performance was lower in this subset than in trials with shorter interruption in shadowing ($t = 2.20$, $p < .05$), which suggests that interruptions in shadowing were not used to enhance task performance via verbal reasoning. Rather, these unsuccessful trials seem to reflect general processing limitations detrimental to both tasks. We nevertheless computed the analyses on the remaining subset of trials (those with few pauses). We found similar results, both on performance and indices, by subject and by item.⁴

⁴Participants who did not have at least one data point in each condition after the criterion for successful shadowing was applied were removed from this analysis. This left 46 participants out of 58 for the ANOVA by subject. None of the 17 scenarios had to be removed for the ANOVA by item. The ANOVA on percentage of correct responses showed significant effects of interference, $F_1(1, 45) = 33.81$, $p < 10^{-6}$; $F_2(4, 64) = 4.20$, $p < .01$, and condition, $F_1(4, 180) = 6.90$, $p < 10^{-4}$; $F_2(1, 16) = 36.94$, $p < 10^{-4}$, but no Interference $\times$ Condition interaction, $F_1(1, 45) = 1.82$, $p = .13$; $F_2(4, 64) = 1.54$, $p = .2$. The ANOVA on indices revealed a significant effect of interference, $F_1(1, 45) = 37.41$, $p < 10^{-6}$; $F_2(1, 16) = 48.45$, $p < 10^{-5}$, and significant and marginal effects of index, $F_1(2, 90) = 4.91$, $p < .05$; $F_2(2, 32) = 3.05$, $p = .06$. The Interference $\times$ Index interaction was significant, $F_1(2, 90) = 3.23$, $p < .05$; $F_2(2, 32) = 3.30$, $p < .05$. Although this interaction was not found in the main analysis, it actually shows a larger effect of verbal shadowing on physical causation detection than on belief and goal detection, which does not alter our conclusions in any way. Additional analyses were conducted after using a even more stringent criterion (i.e., 1,000 ms, $n = 39$). They showed similar results.

Overall, these results indicate that verbal shadowing generally decreased task performance. However, it did not affect the belief attribution index more than the other indices, neither did it reduce it to chance level. Moreover, it did not affect mean response bias.

Discussion

In this study, we used a novel paradigm to tap the ability to process beliefs, goals, and physical causation. We further used verbal shadowing as a way to test the role of inner speech in these abilities. Our results clearly show that heavy verbal interference has an effect on overall task performance, but no more effect on the ability to attribute beliefs than on abilities to make inferences about other mental states or about physical states. Furthermore, participants remained able to attribute beliefs better than chance even while performing verbal shadowing. This suggests that inner speech is not necessary for belief attribution,

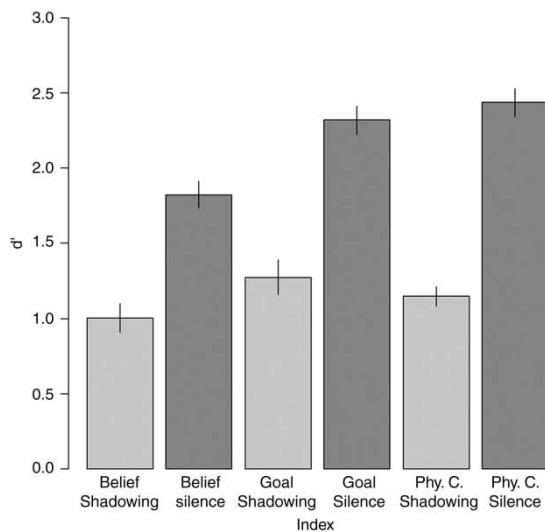


Figure 2. Sensitivity indices for belief, goal, and physical causation attribution, in silence (dark grey) and under verbal shadowing (light grey). Error bars represent standard errors of the mean by item.

Table 3. Mean values of criterion beta for mechanistic and mentalistic ends

	Mentalistic	Mechanistic
Silence	1.07 (0.56)	1.07 (0.82)
Verbal shadowing	0.90 (0.28)	1.18 (0.92)

Note: Standard deviations in parentheses. No criterion was significantly different from 1.

and it is consistent with the idea that belief attribution is independent from verbal reasoning.

One might argue that the interference task is too easy and that verbal shadowing is insufficient to entirely block verbal reasoning. Of course it would be very difficult to entirely rule this out. Yet we did observe that verbal shadowing significantly decreased overall performance, as expected from a demanding dual task. Furthermore, using stringent criteria for successful performance of

the shadowing had the unexpected effect of increasing overall performance, without affecting the general pattern of results.

Could it be that participants may perform above chance without processing beliefs? As explained earlier, performance in each condition of each scenario may be individually explained in terms of behavioural rules like “goal processing”, “perceptual preference”, without referring to invisible representational mental states like beliefs. However, a consistent use of such rules across conditions, without belief attribution, would lead to null d' values (chance level), which is not what we observe.

Then why do our results differ from those obtained by Newton and de Villiers (2007; henceforth NdV)? First, we should stress that the goal of the present study was not to replicate that of NdV. As the general approach is different, there are important methodological differences that preclude a direct comparison between the results. One important difference is that in NdV's study, each subject underwent just 2 trials, while in ours there were 34 trials per subject. What is the impact of repeated trials in such a paradigm? One consequence may be a practice effect on both the main and the dual tasks, improving overall performance. Indeed, we did observe a practice effect over the trials,⁵ but analyses run separately on the first half of the trials showed similar results to those previously reported for the whole session.

Yet another difference is that NdV included a nonverbal interference control condition, while we did not. Instead we compared verbal interference with plain silence. In NdV, this control condition is rendered necessary by the fact that they observe a difference in performance between true- and false-belief conditions under verbal shadowing, which prompts the question of whether this effect is due to an interaction between the general demands of a dual task and condition

⁵Performance under shadowing was significantly lower in the first half of the trials (67%) than in the second half (74%), $F(1, 57) = 7.06$, $p = .01$. Mean duration of shadowing interruption in each trial was significantly higher during the first half of the trials (3,132 ms) than during the second half (2,774 ms), $F(1, 57) = 5.10$, $p = .03$. However, statistics were conducted on the two first trials of each condition and showed similar results.

difficulty. They control for this possibility by showing that another, nonverbal, dual task, with presumed similar demands, does not produce the same effect.⁶ In our study, there is no difference to be explained in the effects of verbal shadowing between mentalistic seen and unseen change conditions. Thus there is no need of a control condition other than "silence".

One major difference between NdV's and our results is in the use of a variety of scenarios. NdV used only one scenario involving an object changing locations. Given the impossibility for any given scenario to measure belief attribution independently of other cognitive processes, we consider that using a variety of scenarios with different computational demands and surface properties is a more reliable approach to assessing belief attribution. We used 17 different scenarios based on several well-established paradigms, and we know of no *a priori* reason to prefer one kind of scenario over another for the purpose of testing the role of inner speech for belief attribution. Our analysis by items supports the idea that the results we have obtained hold across a variety of stimuli and are not due to the specific properties of a given scenario.

Finally, it is worth pointing out that NdV's results are not very different from ours, in the sense that they too find that verbal shadowing does not entirely prevent belief attribution. Indeed, they report that more than 40% of their subjects pass the FB condition despite verbal shadowing. This cannot be interpreted as at or below chance level. Indeed, in a change-of-location paradigm, a failure to attribute beliefs is not expected to generate random responses, but systematic errors, as observed in young children (Wellman et al., 2001). Therefore, if NdV's subjects did not attribute (false) beliefs, they would systematically give the answer corresponding to the real location of the object (i.e., 0% in the FB condition, notwithstanding response noise). In other words, if

a signal detection analysis was applied to NdV's data, this would most probably show a belief sensitivity index significantly greater than zero. Therefore, consistent with our own data, we take NdV's results to provide evidence for the fact that verbal shadowing, although it decreases performance in belief attribution tasks, does not entirely prevent belief attribution. What our study crucially adds to theirs is that we compare this performance decrease between the attribution of beliefs, goals and physical causation, allowing us to test how specific to belief attribution the role of inner speech is.

Thus, we argue that the approach proposed here presents a number of strengths. First, by contrasting five closely related conditions teasing systematically apart different aspects of scene analysis, our paradigm allows us to distinguish the attribution of beliefs, goals, and physical causation in a way that has not been possible in previous paradigms. Secondly, by using a much greater variety of experimental materials, the results that we have obtained may achieve a greater degree of generalizability than those that can be obtained with just one scenario.

In conclusion, language does intervene in the performance of many tasks testing belief attribution, including entirely nonverbal tasks such as the one presented in this study. However, in our experiments, language seems to largely function as a nonspecific performance factor, for instance as an additional tool to keep relevant information in short-term memory. Taken together with results obtained on preverbal infants (Onishi & Baillargeon, 2005) and aphasic patients (Varley et al., 2001), our results are most consistent with the hypothesis that belief attribution is basically independent from language.

Original manuscript received 29 December 2009

Accepted revision received 10 June 2010

First published online 10 January 2011

⁶Incidentally, the assumption that NdV's verbal and nonverbal interference tasks have the same demands is questionable (as pointed out by Apperly et al., 2009). First, NdV's nonverbal interference task alternates between listening and repeating phases, rather than continuous shadowing. Secondly, the reproduction of 4/4 measures of beats leaves many more degrees of freedom than the reproduction of arbitrary sentences.

REFERENCES

- Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W. L., & Humphreys, G. W. (2007). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal, "reality-unknown" false belief and false photograph tasks. *Cognition*, 103(2), 300–321.
- Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, 45(1), 190–201.
- Baldo, J. V., Dronkers, N. F., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language*, 92(3), 240–250.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21(1), 37–46.
- Bek, J., Blades, M., Siegal, M., & Varley, R. (2010). Language and spatial reorientation: Evidence from severe aphasia. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 36(3), 646–658.
- 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.
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112, 337–342.
- Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development*, 70(2), 381–395.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187–192.
- Carruthers, P. (2002). The cognitive functions of language. *Behavioural and Brain Sciences*, 25(6), 657–674; discussion, 674–725.
- Dennett, D. (1978). Beliefs about beliefs (commentary on Premack, et al.). *Behavioral and Brain Sciences*, 1, 568–570.
- de Villiers, J., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 189–226). Hove, UK: Psychology Press.
- Forgeot d'Arc, B. (2009). <http://www.lscp.net/persons/forgeot/stim>
- Frith, U. (2003). *Autism: Explaining the enigma* (2nd ed.). Oxford, UK: Blackwell.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 59(1), 26–37.
- Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66(3), 843–855.
- Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology*, 39(1), 3–36.
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind". *Trends in Cognitive Sciences*, 8(12), 528–533.
- MacMillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. London: Lawrence Erlbaum Associates.
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78(2), 622–646.
- Newton, A. M., & de Villiers, J. G. (2007). Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science: A Journal of the American Psychological Society*, 18(7), 574–579.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258.
- Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development*, 60(3), 688–700.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind. *Behavioral and Brain Sciences*, 1(4), 515–526.
- Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Sciences*, 9(10), 462–463.
- Sokolov, A. N. (1972). *Inner speech and thought*. New York, NY: Plenum Press.
- Sommer, M., Dohnel, K., Sodian, B., Meinhardt, J., Thoermer, C., & Hajak, G. (2007). Neural correlates of true and false belief reasoning. *NeuroImage*, 35(3), 1378–1384.
- Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science: A Journal of the American Psychological Society*, 18(7), 587–592.

Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, 7(6), 489–493.

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.

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.

Zaitchik, D. (1990). When representations conflict with reality: The preschooler’s problem with false beliefs and “false” photographs. *Cognition*, 35(1), 41–68.

APPENDIX

Plot and main properties of scenarios

Mn1 and Mn2: mentalistic ends; Mc1 and Mc2: mechanistic ends; elements in square brackets [] constitute the “change scene” (see Figure 1) and are *seen* or *unseen* by the character in the corresponding conditions; elements in braces {} show parts that are specific to mentalistic/mechanistic conditions. To view a colour version of this Appendix, please see the online issue of the Journal.

Title		Summary	Properties
Bird		Two cages stand on a table. Andy puts the bird in one of two covered cages [he makes a picture of the opened cages] and goes [taking the bird’s grains/plugging the camera on the computer]. [Sandy swaps the cages]. Andy {goes back with the grains, towards the cage on the right/ displays the picture on the computer, displaying the bird in the left cage}. (Mn1 & Mc1) The bird comes out of the left cage. (Mn2 & Mc2) The bird comes out of the right cage.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location <i>Remark:</i> reality unknown
Blotch		[Robert splashes red paint from his pot on the wall and lets the brush fall on the floor. Kevin comes in, holding a pot as well, and picks up the brush.] The two boys are standing near the red-blotched wall, Kevin holding the wet brush. Father comes. (Mn1) He scolds Robert. (Mn2) He scolds Kevin. (Mc1) Kevin pot contains red colour. (Mc2) Kevin pot contains blue colour.	<i>Object of the belief:</i> action attribution <i>Cause of the false belief:</i> appearance/ reality
Bottle		Dany puts a bottle of juice on the table and leaves. [Carry passes by and drinks all the juice from the bottle.] Dany is back with a glass. (Mn1) He takes the bottle on the table. (Mn2) He takes a new bottle on the shelf. (Mc1) The bottle is empty. (Mc2) The bottle is full.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality

(Continued overleaf)

Appendix. Continued.

<i>Title</i>		<i>Summary</i>	<i>Properties</i>
Cake		[Jean eats all the cake and moves away. Then comes Jacques. He seats down at the table near the empty plate.] Jean and Jacques are in the kitchen. Jacques is sitting near the empty plate, then comes their mother. (Mn1) She scolds Jean. (Mn2) She scolds Jacques. (Mc1) Jean has crumbs around his mouth. (Mc2) Jacques has crumbs around his mouth.	<i>Object of the belief:</i> action attribution <i>Cause of the false belief:</i> appearance/ reality
Castle		On the beach, all sandcastles wear flags on their tops. Jesse builds a new sandcastle. {Jesse makes a picture of the sandcastle.} He goes under the dodger and puts his shovel and bucket. [A wave breaks down the sandcastle.] Jesse is under the dodger and (Mn1) picks up a flag, (Mn2) picks up a shovel and bucket. (Mc1) The picture displays the sandcastle. (Mc2) The picture displays no sandcastle.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality
Cat		There is a box on a red table, beside a box on a green table. Andy puts the cat in one of two boxes {and makes a picture of the opened boxes} and goes {picking the cat's food/plugging the camera on the computer}. [Sandy swaps the boxes.] Andy {goes back with the food, towards the box on the red table/displays the picture on the computer, displaying the cat in the box on the green table}. (Mn1 & Mc1) The cat comes out of the box on the red table. (Mn2 & Mc2) The cat comes out of the box on the green table.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location <i>Remark:</i> reality unknown
Chicken		Pete puts a big pan and a chicken on a plate, on a table. [A dog comes in, takes the chicken and runs out.] (Mn1) Pete runs towards the door. (Mn2) Pete brings the sauce on the table. (Mc1) The chicken is on the plate. (Mc2) The chicken is not on the plate.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location
Egg		Kathy has just boiled an egg. She goes away. Vince comes around [he eats the egg and turns the shell] and goes away. Kathy is back. (Mn1) Kathy takes the fork and knife. (Mn2) Kathy takes the rubbish bin. (Mc1) The egg is full. (Mc2) The egg is empty.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality

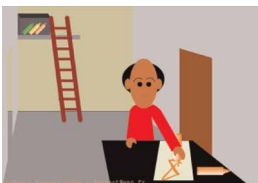
(Continued overleaf)

Appendix. Continued.

<i>Title</i>	<i>Summary</i>	<i>Properties</i>
Fish	 John puts a big pan and a fish on a plate, on a table. [A cat comes in, takes the fish and runs out.] (Mn1) John runs towards the door. (Mn2) John brings the sauce on the table. (Mc1) The fish is on the plate. (Mc2) The fish is not on the plate.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location
Flower theft	 [Phil waters his flowers and leaves. The first walker cuts the flowers and puts them in his jacket. Then comes a second walker, holding exactly the same flowers in his hand.] Both walkers are standing in front of Phil's ruined flower-bed. (Mn1) Phil blames the first walker. (Mn2) Phil blames the second walker. (Mc1) The first walker opens his jacket, and the cut flowers are visible. (Mc2) The first walker opens his jacket, and no flower is visible.	<i>Object of the belief:</i> action attribution <i>Cause of the false belief:</i> appearance/ reality
Hat	 Paul gets his hat blown away by the wind, to the chimney of the green house. Paul goes and picks up a scale. [The wind blows the hat to the chimney of the yellow house.] (Mn1) Paul brings the scale to the green house. (Mn2) Paul brings the scale to the yellow house. (Mc1) The hat falls down from the roof of the yellow house. (Mc2) The hat falls down from the roof of the green house.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location
Hide-and-seek	 Cindy and Ann play Hide and Seek. Cindy surreptitiously looks at Ann hiding under the table. [Ann move away, hiding behind the tree.] (Mn1) Sally is looking for Ann under the table. (Mn2) Sally is seeking Ann behind the tree. (Mc1) Ann is under the table. (Mc2) Ann is behind the tree.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location
Paint	 Steve is sitting on a bank. He stands up and walks to the news kiosk and buys a newspaper. [A maintenance man comes and paints the bank.] Steve comes back. (Mn1) Steve sits on the bank. (Mn2) Steve sits on a chair. (Mc1) When sitting on the bank, one's clothes are stained with paint. (Mc2) When sitting on the bank, one's clothes are not stained with paint.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality

(Continued overleaf)

Appendix. Continued.

Title	Summary	Properties
Pencil	 Mat doodles with the pencil he took on the table [but the paper remains clean. He takes another pencil in a box] and traces a doodle on the paper. He leaves with his paper and pencil. Liz takes a new sheet of paper. (Mn1) She takes the pencil on the table. (Mn2) She takes a new pencil in the box. (Mc1) She doodles on the paper with the pencil on the table, but the paper remains clean. (Mc2) She traces lines on the paper with the pencil on the table.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality
Pet shop	 Calvin and Bart are looking at the bird through the window of the pet shop. Bart goes away while Calvin is entering the shop. [Calvin looks at a golden fish and smiles and then looks at the bird very unhappily.] The rest of the scene is hidden by a car, stopped in front of the shop. Calvin comes out of the shop with a packet. (Mn1) Bart comes back holding a coop. (Mn2) Bart comes back holding a bowl. (Mc1) The packet contains a goldfish. (Mc2) The packet contains a bird.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality
Plants	 Cindy pours water on a plant in a pot on a red base. She then moves toward another plant on a blue base and starts pouring, but no water is flowing. Cindy goes to the back of the stage and begins filling the watering can from a tap. [Albert arrives and swaps the pots.] Cindy comes back to the front stage. (Mn1) Cindy pours water on the plant on the blue base. (Mn2) Cindy pours water on the plant on the red base. (Mc1) Water is leaking from the pot on the blue base. (Mc2) Water is leaking from the pot on the red base.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location
Rain	 Through the door, Tony looks at the rain [and makes a picture of the outdoors}. He walks through the corridor [and the sky is becoming blue]. Tony stands between an umbrella and a pair of sunglasses. (Mn1) He picks up the umbrella. (Mn2) He picks up the sunglasses. (Mc1) The picture displays the rain. (Mc2) The picture displays the sunshine.	<i>Object of the belief:</i> object properties <i>Cause of the false belief:</i> appearance/ reality

(Continued overleaf)

Appendix. Continued.

<i>Title</i>		<i>Summary</i>	<i>Properties</i>
Thief		Sally walks around, when a man steals her handbag from her and runs away through the yellow door. She goes to the telephone box and gives a call. [The man comes back from the yellow door and runs away through the blue door.] The policeman arrives. (Mn1) Sally points in the direction of the yellow door. (Mn2) Sally points in the direction of the blue door. (Mc1) The handbag lies behind the yellow door. (Mc2) The handbag lies behind the blue door.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location
Watering		Bill pours water on a plant in a pot on a red table. He then moves toward another plant on a blue table and starts pouring, but no water is flowing. Bill goes to the back of the stage and begins filling the watering can from a tap. [Mary arrives and swaps the pots.] Bill comes back to the front stage. (Mn1) Bill pours water on the plant on the blue table. (Mn2) Bill pours water on the plant on the red table. (Mc1) Water is leaking from the pot on the blue table. (Mc2) Water is leaking from the pot on the red table.	<i>Object of the belief:</i> object position <i>Cause of the false belief:</i> change of location