Psychological Research
https://doi.org/10.1007/s00426-018-1000-0

ORIGINAL ARTICLE

Spontaneous Theory of Mind is reduced for nonhuman-like agents as compared to human-like agents

Lara Bardi[1] · Charlotte Desmet[1] · Marcel Brass[1]

Received: 19 July 2017 / Accepted: 30 March 2018

Abstract
Theory of Mind research has shown that we spontaneously take into account other's beliefs. In the current study, we investigate, with a spontaneous Theory of Mind (ToM) task, if this belief representation also applies to nonhuman-like agents. In a series of three experiments, we show here that we do not spontaneously take into account beliefs of nonhuman-like others, or at least we do it to a lesser extent than for human and human-like agents. Further, the experience we have with the other agent, in our case a dog, does not modulate spontaneous ToM: the same pattern of results was obtained when dog owners and no owners were compared. However, when more attention was attracted to the dog behavior, participants' behavior was influenced by the beliefs of the dog. In sum, spontaneous belief representation seems to be primarily restricted to human and human-like agents, but can be facilitated when more attention is drawn to a nonhuman-like agent.

Introduction

To interact in a smooth manner with each other it is important that we can infer other people's intentions and beliefs. The human ability to do so has been termed Theory of Mind (ToM) (Premack & Woodruff, 1978). One of the common methods to assess ToM are false belief tasks. These tasks were typically developed to test ToM in children. In the Sally–Anne task, for example, children watch a scene where Sally places an object in a box and then leaves the scene. Anne stays in the room and while Sally is gone she puts the object in a basket. Children are then asked where Sally will look for the object upon reentering the room. The correct answer (box) shows that one can predict the behavior of Sally by taking into account her (false) beliefs of where the object is (Wimmer & Perner, 1983). Because participants are explicitly asked to reason about someone's mental state, these tasks can be referred to as explicit ToM tasks. More recently, spontaneous versions of this task (further called the spontaneous ToM task) have been developed (e.g., Senju, Southgate, White, Frith, 2009; Kovács, Teglas, Endress, 2010; Schneider, Bayliss, Becker, Dux, 2011; Schneider, Slaughter, Dux, 2017). This research shows that we represent other's beliefs even when we are not required to do so and other's mental states are irrelevant to our current goals.

One important and unanswered question is if we spontaneously take into account the belief of all other living creatures. In other words, do the beliefs of all others spontaneously influence our behavior or is this influence restricted to certain others, i.e., others that are more similar to ourselves? The question of how strongly we are influenced by other nonhuman agents has been systematically addressed in the domain of joint action, social attention and imitation. For example, it has been shown that joint action effects are stronger for biological compared to non-biological agents (Tsai & Brass, 2007) or that gaze cueing effects are stronger when observing humans compared to robots (Wiese, Wykowska, Zwickel, Müller, 2012). In addition, these effects can be altered by top down influences. For example, when priming participants with a video fragment of Pinocchio, joint action effects emerged for a wooden hand (Müller et al., 2011). Along the same line, gaze cueing effects were reduced when a human face was believed to be a mannequin and vice versa when a robot was believed to be controlled by a human (Wiese et al., 2012).

Finally, research revealed that imitation is sensitive to the similarity between the model and the observer: the

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00426-018-1000-0) contains supplementary material, which is available to authorized users.

✉ Lara Bardi
lara.bardi@ugent.be

[1] Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium

Published online: 16 April 2018

observation of human actions leads to stronger imitative responses than the observation of nonhuman actions (e.g., Kilner, Paulignan, Blakemore, 2003; Press, Gillmeister, Heyes, 2006; Klapper, Ramsey, Wigboldus, Cross, 2014). For instance, Press et al. (2006) presented participants with actions performed by a human actor or by a robot. Results revealed that an automatic tendency to imitate these actions was present for human as well as for robotic agents. Crucially, however, the imitation effect was reduced when observing robotic actions. Some studies report that imitation is also subject to top-down animacy beliefs modulation (e.g., Longo & Berenthal, 2009; Liepelt & Brass, 2010; Klapper et al., 2014), but others do not (Press et al., 2006). For example, Liepelt & Brass (2010) presented participants with movements of a hand that was placed in a glove. Participants either believed that the hand in the glove was a human hand or a wooden hand. The results revealed stronger imitation effects when participants believed that a human, compared to a wooden hand, executed the movements. On the other had, Press et al., (2006) found that imitation was modulated by stimulus properties (human vs. robotic hands) but not by the participant's belief about the animacy of the stimuli.

From this we can conclude, that the more we perceive an agent as human-like, based on bottom-up characteristics or top down influences, the more the behavior of this other agent influences our own behavior. In these domains, however, it is tested how behaviour of others influence our own behaviour. Therefore, it is highly important to investigate the boundaries of this belief representational system.

In particular, the question arises whether we spontaneously represent beliefs of dissimilar others and specifically, of nonhuman agents and whether these representations directly influence our own behavior. Interestingly, it has been shown that even simple geometric shapes, having pattern of contingent movements, elicit the attribution of complex internal states, such as intentions and beliefs (Heider & Simmel, 1944; Castelli, Happé, Frith, Frith, 2000), suggesting that mental states attribution is not restricted to human-like agents. In spontaneous ToM tasks, which measure the effect of others' mental states on our own behavior, representations of others' beliefs have been reported in the spontaneous ToM task using a smurf as an agent and using Buzz Lightyear (a figure from the 'Toy Story' movies) as an agent (Kovács et al., 2010; Nijhof, Brass, Bardi, Wiersema, 2016; Bardi, Desmet, Nijhof, Wiersema, Brass, 2017a). While these comic figures are also nonhuman agents, it is very likely that we anthropomorphize them (e.g., Epley, Akalis, Waytz, Cacioppo, 2008) and we treat them as humans. So it remains an open question whether we spontaneously represent the belief of agents that we 'perceive' as humans or similar to us or we also represent beliefs of agents that we perceive as nonhumans.

To do deal with this issue, in the present work we will use a spontaneous ToM task and present an agent that is clearly dissimilar to humans, namely a dog. We used an adapted version of a spontaneous ToM task (Deschrijver, Bardi, Wiersema, Brass, 2016; Nijhof et al., 2016; Bardi et al., 2017a), originally developed by Kovács et al. (2010). Here, participants are presented with a video representing an agent who obtains certain knowledge about the location of an object, this being either behind an occluder or outside of the scene (belief formation phase). At the end of the video, the occluder is lowered and participants are requested to press a button if the object is present behind the occluder (outcome). The ball is behind the occluder in 50% of the cases (randomly with respect to the belief formation phase). Reaction times depend on the participant expectations: responses are faster when the participant expects the object to be present (P+ conditions) than when he/she does not (P− conditions). This difference is also referred to as reality-bias (Deschrijver et al., 2016; Bardi, Six, Brass, 2017b). More strikingly, responses are also shortened when the agent only (false belief condition) believes the object is present (P-A+), showing that participants' performance is also influenced by the other's expectations about the presence of the object. The difference in RTs when neither the participant nor the agent expects the object to be present (P-A−) and the conditions in which the agent expects the object to be present (P-A+), here referred as ToM index, reflects the pure influence of the agent's belief and is taken as an index of ToM processing (Kovács et al., 2010; Deschrijver et al., 2016; Nijhof et al., 2016, Bardi et al., 2017b). Previous neuroimaging studies support the idea that the P-A + condition is the critical condition reflecting belief representation. The TPJ, a key node in the "ToM network", was more active in incongruent trials (P−A+, P+A−) than in congruent trials (P−A−, P + A+). Importantly, this effect was driven by the P−A + condition (Bardi et al., 2017b; Kovács, Kühn, Gergely, Csibra, Brass, 2014). Moreover, the size of the ToM index was related to the symptomatology of autism in a group of individuals with the diagnosis of autism (Deschrijver et al., 2016) and in the neurotypical population (Nijhof, Brass, Wiersema, 2017).

If we take into account the beliefs of all biological agents, we expect that participants' performance will be influenced by the belief of the agent when this is a dog, resulting in a significant ToM index. On the contrary, if we do not represent the belief of a dog, we expect to see a reduced or absent ToM index, possibly together with a larger reality bias. Second, the use of a dog as the agent in our task, will allow us to investigate if the familiarity we have with the type of agent, affects the belief representation. More specifically, we recruited both dog owners and participants that do not own a dog and investigate whether they show a different pattern of results (Experiment 1). One may hypothesize that dog owners, having an extensive interaction with a dog, will perceive a dog as more similar to themselves and, therefore, will take into account the beliefs of a dog more strongly than participants who do not own a dog. In Experiment 2 we directly compared the spontaneous ToM effects when a dog is presented as an agent with the case where a human-like agent is presented. Finally, in Experiment 3 we investigate if addressing more attention to the dog has an influence on belief representation.

Experiment 1

Method

Participants

Thirty-one participants took part in this experiment (six males). Fourteen participants were dog owners. They were all University students and were paid 10 euros in turn for participation. The local ethical committee approved the study.

Design and stimuli

Eight different video clips were presented in the experiment. They all showed a dog, a rolling ball and an occluder and were presented with Presentation software (NeuroBehavioural Systems, Albany, CA, USA). During the first phase of the video (the belief formation phase) the participant and the agent watch the ball make a trajectory. Dependent on the last seen position of the ball the participant expects the ball to be behind the occluder or not. Moreover, when the ball rolled out of the scene before the occluder was lowered, the participant expects the ball to be absent (P−), whereas if the ball rolled behind the occluder before it was lowered, participants should expect the ball to be present (P+). The belief of the agent was manipulated by varying the time at which the agent left the scene. When the agent left the scene after the ball reached its final position, then the agent has a true belief about the ball's presence (i.e., the same belief as the participant). There are two situations in which this is the case: both the participant and agent believe that the ball is present (P+A+) or both the participant and the agent believe that the ball is absent (P−A−). When the agent left the scene before the ball reached its final position, then the agent has a false belief about the ball's presence (i.e., a different belief than the participant). Again, two situations can occur that match this false belief situation: when the participant expects the ball to be absent and the agent believes the ball to be present (P−A+) and when the participant believes the ball to be present and the agent believes the ball to be absent (P+A−). During the second phase of the video (outcome phase), the occluder falls down and in 50% of the cases there is a ball (irrespective of what happened in the belief formation phase). 8 possible videos were presented resulting from a combination of the factors Participant [the participant expects the ball to be there (P+) or not (P−)] by Agent [the agent believes the ball is present (A+) or absent (A−)] by Outcome [the ball is present (B+) or absent (B−) at the end of the movie]. This combination results in 4 true belief videos (P−A−B−, P−A−B+, P+A+B− and P+A+B+) and 4 false belief videos (P−A+B−, P−A+B+, P+A−B−, P+A−B+). See Fig. 1 for an overview. Every video clip was repeated 5 times through the experiment, this resulted in 40 experimental trials. Before the experiment started, participants completed four practice trials. During these trials, feedback on participants' performance was given. Participants were required to watch the video clips attentively and to press a right response button with their right index finger (key B of the keyboard) if they detected a ball at the end of the video clip. To make sure participants were attentive they also had to make a button press when the agent left the scene. A left key press was required here (key V of the keyboard).

The video clips were generated with a 3D modeling software (3D Studio Max 4.2, Autodesk, Discreet). The videos were all 18 s long. During the first part of the video the belief of the agent was formed (belief formation phase). All video clips started with a dog letting a ball role behind an occluder by pushing its paw against it. This lasted 5 s. Moreover, dependent on the condition the movies continued as follows: In the P+A+ condition, the ball rolls out of the scene and then returns behind the occluder (12 s). Then the dog leaves the scene (14 s) and reappears. During the absence of the dog, the ball stays behind the occluder. This results in a true belief situation where both the participant and the agent believe the ball to be present. In the P−A− condition, the ball rolls from behind the occluder then returns behind the occluder and finally rolls out of the scene (at 12 s measured from the start of the video clip). After this has happened, the dog leaves the scene (at 14 s), resulting in a true belief situation where both the dog and the participant belief the ball to be absent. In the P−A+ condition, the dog already leaves the scene while the ball is still behind the occluder (8 s). During the absence of the dog, the ball rolls from behind the occluder, returns behind the occluder and finally leaves the scene (14 s). Then the dog returns. This results in a false belief situation where the participant believes the ball to be absent, while the dog thinks that the ball is present. In the P+A− condition, the ball rolls from behind the occluder out of the scene. Then the dog leaves the scene (11 s). During the absence of the dog, the ball returns behind the occluder (14 s) and then the dog reappears. This results in a false belief situation where the dog

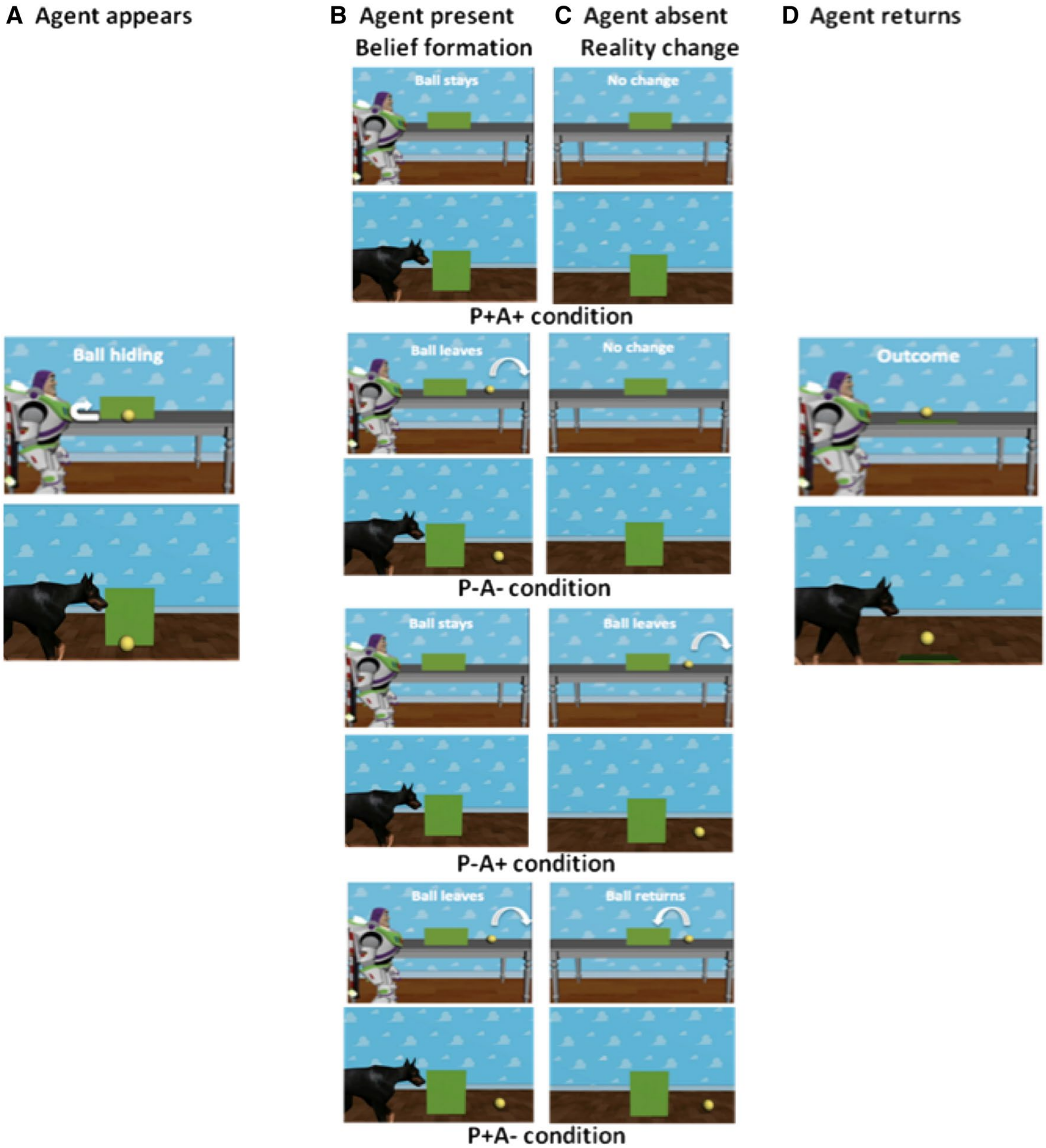

Fig. 1 Frames of the movies presented during the spontaneous ToM task. Frames are shown for the dog task (Experiment 1) and the Buzz task (Experiment 2). There were eight conditions, resulting from the combination of the expectation of the participant and the belief of the Agent (**b, c**), and outcome phase (**d**). In the first part of the movie (**a**), the ball rolls behind the screen. In the second part (**b**), in the presence of the agent, the ball can change location or stay behind the occluder. In the third part (**c**) the agent leaves the scene. At this point the ball can change its location or not. In the last part (**d**), the agent comes back to the scene and the occluder is lowered. The ball is present or not

believes the ball to be absent and the participant beliefs the ball to be present. The final phase of the video clip (outcome phase) lasted 3 s and was presented in all conditions. During this outcome phase the occluder was lowered. In 50% of the trials a ball was present, in the other half of the trials there was no ball. The presence/absence of the ball was randomized over the four possible belief conditions.

Results

Two participants were excluded from the further analyses (one dog owner and one participant who did not own a dog). One of those participants responded too slow (> 2,5 SD's from the mean). The other participant did not understand the instructions correctly. Only trials in which a ball appeared (50%) were included in the analyses. Further, we excluded trials in which participants responded before the occluder was lowered, or where the response was incorrect (no response or multiple responses when the agent left the scene or no response or multiple responses when the occluder was lowered). This resulted in the exclusion of 1% in the trials where a ball appeared.

Two analyses were conducted. First, we tested whether the reality bias was present by comparing the condition where the participant expected the ball to be absent P− [(P−A−) + (P−A+)/2] with the condition where the participant expected the ball to be present P+ [(P+A−) + (P+A+)/2]. Results show a significant reality bias, $t(28) = 3.8$, $p = 0.001$, with detection times in the P− conditions being larger than in the P+ conditions (P− conditions: $M = 381$, SD = 82, P+ conditions: $M = 352$, SD = 80). This outcome confirms that participants formulated predictions about the presence of the ball and these predictions affect performance.

Second, we tested whether the dog's belief affected participants' performance and whether belief representation is influenced by dog ownership. A repeated-measures ANOVA was computed on detection times with the factors Belief (P−A− vs. P−A+) and the between variable Ownership (dog owner vs. no owner). The two conditions that were entered in the ANOVA are the conditions that represent ToM processes, namely the condition where the agent believes the ball to be present but where the participant does not expect the ball (P−A+) and the condition where both the agent and the participant do not expect the ball to be present (P−A−). Larger detection times in the P−A− condition compared to the P−A + condition indicate that participants take into account the beliefs of the agent. A positive index reflects a representation of the beliefs of the other agent. The factor Belief was not significant, $F(1,27) = 1.42$, $p = 0.24$, $\eta^2_p = 0.05$, indicating that detection times were not affected by the belief of the dog (P−A−: $M = 387$, SD = 92, P−A+: $M = 374$, SD = 84). There was no significant main effect of Ownership, $F(1,27) = 2.73$, $p = 0.11$, $\eta^2_p = 0.09$. Further, there was no significant interaction between Belief and Ownership, $F(1,27) = 0.81$, $p = 0.38$, $\eta^2_p = 0.03$, indicating that there was no difference between the P−A− and P−A+ conditions in neither groups. In other words, both dog owners and no owners show a non-significant ToM index (see Fig. 2).

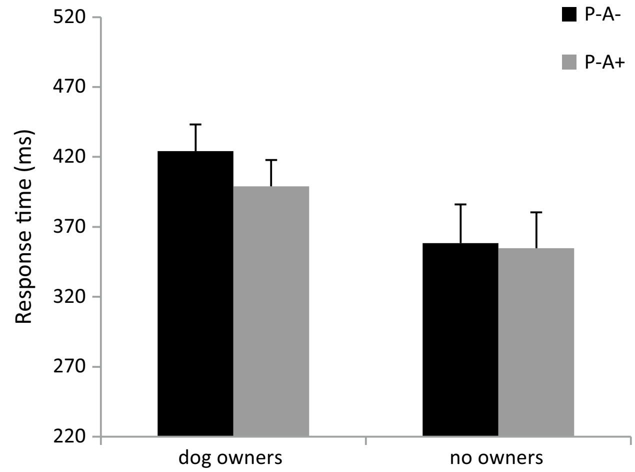

Fig. 2 Upper panel. Reaction times in experiment 1 for the conditions P−A− and P−A + for the group of dog owners and the group of no owners. Lower panel. RDI plots of the ToM index in the two groups

Experiment 2

The results of Experiment 1 show that while participants take into account their own expectations when carrying out the task, the beliefs of the dog do not affect their detection times. Second, there was no difference in ToM index for participants that own a dog and participants that do not own a dog. These data suggest that we do not spontaneously take into account the belief of a dog. However, this conclusion may be premature since it is based on a null finding. Therefore, we conducted a second experiment in which we presented a new group of participants with two different versions of the spontaneous ToM task. One task was identical to the one we used in Experiment 1 and showed a dog as an agent. The other task was identical to that used in previous studies (e.g., Nijhof et al., 2016; Bardi et al., 2017a) and showed Buzz Lightyear as the agent. If the conclusion from Experiment 1 is warranted, we should obtain a significant ToM index in the Buzz Lightyear task but no significant ToM index in the dog task. Since dog ownership did not influence the data pattern in Experiment 1, we only recruited participants that did not own a dog in Experiment 2.

Method

Participants

Forty-one participants (eight males) participated in this experiment. They were all students and were paid 10 euros or received course credits in turn for participation. The local ethical committee approved the study.

Design and stimuli

We doubled the amount of trials with respect to the previous experiment. Instead of 40 trials, every task (one with the dog as an agent and one with Buzz Lightyear as an agent) now counted 80 trials. The rest of the design was identical to the design used in Experiment 1. The stimuli for the task with the dog were identical as in Experiment 1. The stimuli for the other task were different. Instead of a dog, Buzz Lightyear appeared as an agent. Further the ball was placed on the table by Buzz before it began to roll. However, the different events in the movie and their timing were exactly the same as in the task with the dog as an agent. The order of the two tasks (dog task and Buzz task) was counterbalanced across participants. Participants performed the two tasks one after the other with a few minutes break.

Results

One participant was excluded from further analyses because the participant made a lot of premature responses (50% of the responses in the Buzz task occurred before the occluder was lowered). As in Experiment 1, we only included trials in which a ball appeared (50%). Further, we excluded trials in which participants responded before the occluder was lowered, or where the response was incorrect (no response or multiple responses when the agent left the scene or no response or multiple responses when the occluder was lowered). This resulted in the exclusion of 3% in the trials where a ball appeared.

As in Experiment 1 we first tested whether the participant's expectations about the presence of the ball affected their detection times. A repeated-measures ANOVA with the factors Task (dog vs. Buzz), Participant (P− vs. P+) and Order (dog task first vs. Buzz task first) was computed on detection times. A main effect of Participant was observed (reality bias), indicating that participants were slower when they believed the ball to be absent compared to when they believed the ball to be present, $F(1,38) = 27.55$, $p < 0.001$, $\eta^2_p = 0.42$ (P− conditions: $M = 771$, SD = 106, P+ conditions: $M = 734$, SD = 93). There was no main effect of Task, $F < 1$, and no significant interaction between Participant and Task, $F(1,38) = 2.73$, $p = 0.11$, $\eta^2_p = 0.07$. Explorative simple comparisons confirmed that the reality bias was significant for both the dog $t(39) = 4.03$, $p = 0.006$ (12 ms) and Buzz $t(39) = 4.03$, $p < 0.001$ (26 ms). A significant interaction between Order and Task revealed simple practice effects $F(1,38) = 5.16$, $p = 0.029$, $\eta^2_p = 0.12$. Participants who performed the dog task first, were slower in the dog than in the Buzz task and the reverse was true for the participants who started with the Buzz task.

Second, to test for the effect of the agent's belief on participants' performance, data were analyzed by means of a repeated-measures ANOVA with the factors Task (dog vs. Buzz), Belief (P−A− vs P−A+) and Order (dog task first vs. Buzz task first). We found a significant main effect of Belief, $F(1,38) = 13.39$, $p = 0.001$, $\eta^2_p = 0.26$, indicating that participants responded faster when the agent believed the ball would be present compared to when the agent believed that the ball would be absent. In line with our hypothesis, we observed a significant interaction of Task × Belief, $F(1,38) = 7.23$, $p = 0.01$, $\eta^2_p = 0.16$. Post hoc *t* tests, revealed that the difference between the P−A− and the P−A+ condition (i.e., the ToM index) was significant in the Buzz task, $t(39) = 4.49$, $p < 0.001$ (P−A−: $M = 405$, SD = 71, P−A+: $M = 369$, SD = 54), but not in the dog task, $t < 1$ (P−A−: $M = 386$, SD = 68, P−A+: $M = 380$, SD = 58). A significant interaction between Order and Task revealed simple practice effects $F(1,38) = 7.25$, $p = 0.01$, $\eta^2_p = 0.16$. Participants who performed the dog task first, were slower in the dog than in the Buzz task and the reverse was true for the participants who started with the Buzz task. There was no significant main effect of Task, $F < 1$. See Fig. 3.

Experiment 3

The results of Experiment 1 and Experiment 2 indicate that we do not take into account the beliefs of a dog. Here, we thus show that spontaneous belief representation is absent when confronted with an agent that is not human-like such as a dog. In line with research in other domains, such as

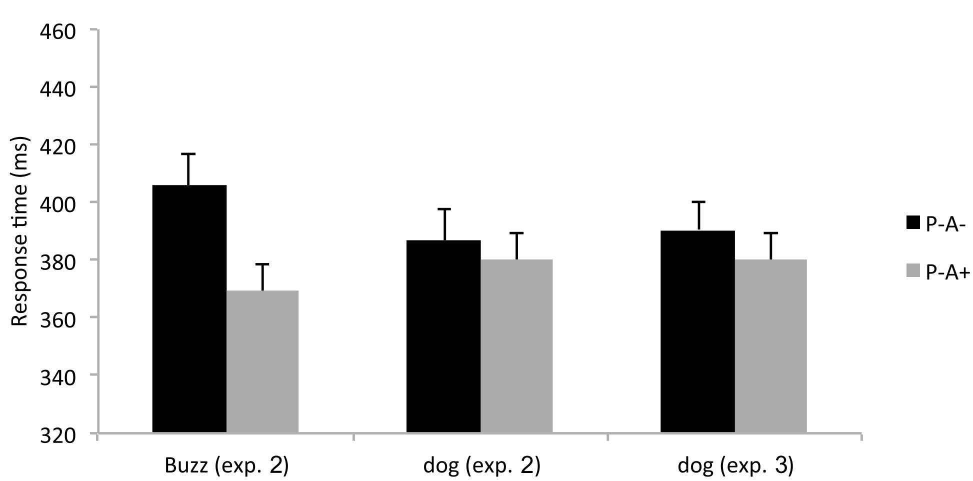

Fig. 3 Detection times in experiment 2 and 3 for conditions P−A− and P−A+ for the spontaneous ToM task displaying Buzz as an agent and for the spontaneous ToM task displaying a dog as an agent (experiments 2 and 3)

automatic imitation, we thus show that different agents have a different influence on our behavior. However, this raises the question whether it is possible to introduce a manipulation that induces spontaneous belief representation for a dog. Studies on imitation have shown that the amount of attention directed to the specific agent can enhance the imitation processes (Heyes, 2011). Moreover, it has been recently shown that directing participant's attention towards the location of the agent in an implicit perspective-taking task increases the effect of the agent's perspective on performance (Bokowski, Hietanen, Samson, 2015). This idea that has its roots in dual-route models of automatic processes (Kornblum, Hasbroucq, Osman, 1990) will be tested here in the context of spontaneous belief representation. We included two manipulations to increase attention to the nonhuman agent. First of all, before the experiment started, we let participants play an interactive ball game on the computer with the dog. In this way, we aimed to increase the bond/familiarity between the agent and the participant and as such generally increase participants' focus on the dog. Second, we colored the collar of the dog. This collar could be blue or red and occasionally subjects were asked what color the collar in the previous trial was. In this way, they should focus their attention more to the dog in each video clip. If attention indeed alters the spontaneous representation of beliefs, we should obtain a significant ToM index in this third experiment.

Method

Participants

Forty-two participants (24 females) participated in this experiment. They were all students from Ghent University and were paid 10 euros or gained course credits in turn for participation. The local ethic committee approved the study.

Design and stimuli

Participants carried out the spontaneous ToM task with the dog as an agent. There were two differences compared to the task used in Experiment 2. First, the dog in the video clips wore a coloured collar. This collar could be blue (50% of the movie clips) or red (50%). Second, to make sure that participants paid attention to the agent, we inserted catch questions in 20% of the trials. The question was always the same namely 'Does the dog have a blue collar?'. The catch questions were randomly inserted and appeared after the end of the movie clip. Questions were presented in black text on a light grey background for 1000 ms. The words 'Yes' and 'No' were presented on the left or right of the screen. 50% of the catch questions had 'Yes' printed left and 'No' right, 50% vice versa. In this way, responses could not be planned in advance. Participants had to respond to the answer on the left with their left middle finger (key 'W') and to the answer on the right with their left index finger (key 'X').

Before participants started the spontaneous ToM task, they played an interactive ball game with the dog. Participants were told that they would have to perform a visual detection task later but that they first would play a small game with the dog. During this ball game, participants saw (and heard) a barking dog with a ball placed in front of him. Participants were told that when the dog barked, they could throw the ball away by pressing the space bar. When participants pressed the space bar, the ball rolled away, the dog caught the ball and brought it back to the participant. This was repeated five times.

Results

Two participants were excluded from further analyses. One participant responded too slow (> 2,5 SD's from the mean), the other participant often (> 2,5 SD's from the mean) reacted with more than one button press to the detection of the dog and to the detection of the ball. As in the previous experiments we only included trials in which a ball appeared (50%). Further, we excluded trials in which participants responded before the occluder was lowered, or where the response was incorrect (no response or multiple responses when the agent left the scene or no response or multiple responses when the occluder was lowered). This resulted in an extra exclusion of 4% in the trials where a ball appeared.

As in the previous two experiments, we analyzed the reality bias and the ToM index. A positive reality bias was observed, $t(39) = 2.42$, $p = 0.02$, showing that participants were slower when they expected the ball to be absent compared to when they expected the ball to be present (P− conditions: $M = 770$, SD = 116, P+ conditions: $M = 741$, SD = 111). Interestingly, the comparison between the P−A− condition and the P−A+ condition (ToM index) reached significance, $t(39) = 2.05$, $p = 0.047$. In contrast to the other experiments, we here observe that the dog's belief affect participants' detection times. More specifically, when the dog believed that the ball would be present behind the occluder, responses were faster as compared to when neither the agent nor the participants expected the ball to be present (P−A−: $M = 390$, SD = 60; P−A+: $M = 379$, SD = 60) (Fig. 3). Although significant, it should be noted that this effect was still significantly smaller with respect to the effect obtained with Buzz in experiment 2, $t(78) = 8.04$, $p = 0.009$.

Discussion

Previous research indicated that we spontaneously take into account the beliefs of other agents (e.g., Kovács et al., 2010; Bardi et al., 2017a; Schneider et al., 2011; Schneider et al., 2017), even when these beliefs are irrelevant for or current tasks. In the present study, we examined the boundaries of this spontaneous belief representation system. First of all, we investigated if we represent the beliefs of nonhuman-like agents, such as dogs. Second, we wanted to see whether experience with the type of agent modulates the representation of the beliefs of this agent. Finally, we examined if drawing attention to the agent influences belief representation. To this aim, we used a spontaneous ToM task and instead of a human-like character as agent, we presented a dog as agent. Since a dog is clearly nonhuman we could investigate if we represent the beliefs of other agents that are not human-like. Further, we considered the experience people have with such agents by recruiting both dog owners and participants that did not own a dog. Finally, in our last experiment we investigated the effect of increased attention (and increased familiarity with the agent through the initial interaction game) to the nonhuman agent by letting participants play an interactive ball game with the dog and by asking questions about the color of the collar of the dog. The results extracted from these experiments are threefold. First, we show that detection times in the spontaneous ToM task are not reliably influenced by the beliefs of a dog (Experiment 1 and 2) while they were significantly influenced when observing a human-like character (Experiment 2). Second, also dog owners did not spontaneously represent the beliefs of dogs (Experiment 1). It thus seems that having a dog, or having experience with the other agent, does not lead to a spontaneous representation of the beliefs of this agent. Third, when drawing attention to the dog, and increasing the bond between the participant and the dog, beliefs of the dog were spontaneously taken into account (Experiment 3). In the following sections, we further discuss each outcome, underlying the limitations of the current investigation and future directions.

Belief representation of nonhuman characters

It thus seems that we do not represent the beliefs of all agents that surround us. When observing a dissimilar other, such as a dog, the beliefs of this agent do not influence our behavior. This null finding cannot be attributed to a lack of sensitivity of the task for two reasons. First of all, in all our experiments we replicate the reality bias. That is, participants were slower in the ball detection when they believed the ball to be absent compared to when they believe the ball to be present. This replicates earlier findings and shows that the task is sensitive to the participant's expectation. Second, when Buzz Lightyear was presented as an agent, we observed a significant ToM index. That is, participants were slower to detect the ball when Buzz believed the ball to be absent compared to the situations where Buzz believed the ball to be present. As already outlined in the introduction, one can argue that we perceive Buzz Lightyear as human-like and that, therefore, we take into account his beliefs. This is likely because, although Buzz is a robot, he has the body shape of a human and he is a cartoon character. In line with this finding, previous research in the action observation domain has shown that the more we perceive another agent as human-like, the more we represent his/her actions. For example, observing a wooden hand after watching a video of Pinocchio enhanced motor simulation compared to a situation where a control video was shown (Müller et al., 2011). Here, we show parallel findings in the ToM domain. When we observe a nonhuman agent that has human-like physical characteristics and can thus be perceived as human, we represent the beliefs of this agent. In contrast, when we observe a nonhuman agent that is clearly dissimilar from us, the beliefs of this nonhuman agent are not spontaneously taken into account, or at least to a lesser extent. This result may appear surprising because we commonly refer to a dog's mental state using human-like descriptors (e.g., that dog loves me, my dog misses me). Moreover, surely dogs have beliefs (e.g., their knowledge about objects' location), while robots have not. Here, we can conclude that the physical properties of the agent might to be the critical factor in determining whether, and to which extent, other's beliefs influence our performance in a spontaneous ToM task. Future studies should broaden the present results by investigating the specific properties that determine the boundaries of belief representation.

Belief representation and experience towards the agent

Interestingly, dog ownership did not alter this result. This was not in line with our expectations. We hypothesized, based on the above reasoning, that the more you observe a dog as similar, the more you will take into account his beliefs. Previous research has indicated that dog owners anthropomorphize their dogs more than no owners (Horowitz & Bekoff, 2007), and thus perceive their dogs as more similar. Therefore, we expected to obtain an interaction between ownership and belief computation, which we did not find. This result suggests that even dog owners do not represent the beliefs of the dog spontaneously. However, it could still be that they represent the beliefs of their own dog but that this belief representation does not generalize to the category of the agent in general (i.e., dogs in general). More importantly, it should be noted that the sample size in the two groups was rather limited and, therefore, the outcome of experiment 1, with respect to the effect of familiarity, should be taken with caution. To test the effect of the experience with the agent on spontaneous ToM, future studies should test larger samples.

Belief representation and the type of agent

It must be noted that the dog we used in the video clips, is a protection dog (Doberman) and is in general perceived as rather aggressive. Therefore, it could be that participants had a negative feeling towards this dog and, therefore, did not represent the beliefs of the dog. It has been demonstrated that anxiety increases egocentrism and diminishes altercentrism in a perspective-taking paradigm that captures implicit visual perspective taking (Todd, Forstmann, Burgmer, Brooks, Galinsky, 2015; Todd & Simpsons 2016). It is, therefore, plausible that participants did not track the dog belief because of the anxiety generated by the fact that the dog was a Doberman. Furthermore, dogs have a sense of smell allowing them to track invisible objects, based on their smell. In the context of our task, if participants did attribute such ability to the dog agent, they could have inferred that the dog has tracked the correct location of the ball. In other words, the dog never had a false belief, which could explain the null finding. This observation calls for a replication of the previous results with the adoption of a different nonhuman-like agent.

Belief representation and attention

In our last experiment, we increased attention towards the dog in the video clips by letting participants play an interactive ball game with the dog and by asking question about the collar of the dog. The results of this last experiment show that, indeed, after increasing attention to the agent, and familiarity with it, a significant ToM index was observed in the spontaneous ToM task. In line with research on the automatic imitation effect, this shows that spontaneous belief representation can be enhanced when more attention is directed to the agent. This result suggests that the main difference between the human-like and nonhuman-like agent in our task is not our ability to attribute mental states to the agents, but rather the degree of attention directed to them. Once, we attend to the dog, belief representation mechanisms start operating similarly to the human-like agent. Because we implemented two manipulations, one increasing the bond between the participant and the dog (the interaction game), and the other increasing attention to the agent throughout the task trials (the color task), we cannot know whether both added to the effect or if only the interaction game or the color task would let the effect emerge. Further research should investigate this issue. However, one might argue whether belief representation can be really expressed in terms of a dichotomy, i.e., either we do represent the other's belief or we do not, or it is rather gradually modulated. The fact that attention can make the ToM effect to emerge suggests that the latter may be true. Here we combined the data from the three experiments to reach a much larger sample ($N = 109$) and see whether we would obtain a significant ToM index for the dog with that sample. We performed a repeated measures ANOVA with the factors Task (exp. 1, 2, 3) as between-subject factor and Belief (P−D− versus P−D+) as within-subject factor. Results revealed a significant main effect of Belief $F(1,106) = 4.15$, $p = 0.03$, with a small effect size, $\eta^2_p = 0.04$. No interaction between Task and Belief was found. The representation of the other's belief for the dog seems to be weak, but consistent.

Spontaneous ToM task

In addition, we further validate the spontaneous ToM task. As already stated above, we show a reality bias in all experiments and when using Buzz Lightyear as a character, we replicate the previously observed data pattern (Kovács et al., 2010; Deschrijver et al., 2016). Recently, Philips et al. (2015) argued that the observed data pattern in the spontaneous ToM task, could be explained by a confound. They argued that the data pattern did not reflect ToM processes per se but were the result of the attention check. More precisely, the authors argued that the button press participants had to make when the agent left the scene influenced their reaction times. Because the timing of this button press was different for the different belief conditions this could have interfered with the data. However, here the timing of the button in the different conditions was exactly the same for the experiment were the dog was shown as an agent as for the experiment where Buzz Lightyear was shown as an agent. The only difference between these two experiments, was the observed agent. If the behavioral effect observed in the spontaneous ToM task adopted here would not measure ToM processes but would be entirely due to a artifact of the event timing in the movies, we would not have observed the current data pattern.

Conclusion

In the current experiments, we show that we do not spontaneously represent the beliefs of all other agents. The beliefs of nonhuman-like agents, such as dogs, do not strongly influence our behavior. Further, experience with the agent (as measured by dog ownership), is not sufficient to represent the beliefs of this agent. However, when attention was drawn to the dog shown in the video clips, the beliefs of the dog significantly affected participants' performance. Moreover, when combining data from three experiments to reach a much larger sample, we observed a significant, although small, ToM effects for the dog as well, suggesting that whether we represent or not the other's beliefs is not a dichotomy. Rather, in line with what has been previously observed for other social phenomena, such as joint actions and imitation, spontaneous ToM is strongly reduced, but not completely abolished for nonhuman-like agents. Much more research is needed to confirm and extend our results, with different kinds of agent and different tasks.

Funding This study was funded by FP7 Marie Sklodowska-Curie Individual Fellowship to Lara Bardi (GRANT 331323-Mirroring and ToM).

Compliance with ethical standards

Conflict of interest Author LB declares that she has no conflict of interest. Author CD declares that she has no conflict of interest. Author MB declares that he has no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Bardi, L., Desmet, C., Nijhof, A., Wiersema, J. R., & Brass, M. (2017a). Brain activation for spontaneous and explicit false belief tasks overlaps: new fMRI evidence on belief processing and violation of expectation. *Social Cognitive and Affective Neuroscience, 12*, 391–400.

Bardi, L., Six, P., & Brass, M. (2017b). Repetitive TMS of the temporoparietal junction disrupts participant's expectations in a spontaneous theory of mind task. *Social Cognitive and Affective Neuroscience, 12*, 1775–1782.

Bukowski, H., Hietanen, J. K., & Samson, D. (2015). From gaze cueing to perspective taking: Revisiting the claim that we automatically compute where or what other people are looking at. *Visual Cognition, 23*(8), 1020–1042.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12*(3), 314–325.

Deschrijver, E., Bardi, L., Wiersema, J. R., & Brass, M. (2016). Behavioral measures of implicit theory of mind in adults with high functioning autism. *Cognitive Neuroscience, 7*, 1–4.

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating social connection through inferential reproduction—Loneliness and perceived agency in gadgets, gods, and greyhounds. *Psychological Science, 19*, 114–120.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 243–259.

Heyes, C. (2011). Automatic Imitation. *Psychological Bulletin, 137*, 463–483.

Horowitz, A. C., & Bekoff, M. (2007). Naturalizing anthropomorphism: Behavioral prompts to our humanizing of animals. *Anthrozoos, 20*, 23–35.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*, 522–525.

Klapper, A., Ramsey, R., Wigboldus, D., & Cross, E. S. (2014). The control of automatic imitation based on bottom-up and top-down cues to animacy: insights from brain and behavior. *Journal of Cognitive Neuroscience, 26*(11), 2503–2513.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap—Cognitive basis for stimulus-response compatibility—A model and taxonomy. *Psychological Review, 97*, 253–270.

Kovács, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*, 1830–1834.

Kovács, Á. M., Kühn, S., Gergely, G., Csibra, G., & Brass, M. (2014). Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PLoS One, 9*(9), e106558.

Liepelt, R., & Brass, M. (2010b). Top-down modulation of motor priming by belief about animacy. *Experimental Psychology, 57*(3), 221–227.

Longo, M. R., & Bertenthal, B. I. (2009). Attention modulates the specificity of automatic imitation to human actors. *Experimental Brain Research, 192*(4), 739–744.

Muller, B. C. N., Brass, M., Kuhn, S., Tsai, C. C., Nieuwboer, W., Dijksterhuis, A., & van Baaren, R. B. (2011). When Pinocchio acts like a human, a wooden hand becomes embodied. Action co-representation for non-biological agents. *Neuropsychologia, 49*, 1373–1377.

Nijhof, A. D., Brass, M., Bardi, L., & Wiersema, J. R. (2016). Measuring mentalizing ability: A within-subject comparison between an explicit and implicit version of a ball detection task. *PLoS One, 11*(10), 1–15.

Nijhof, A. D., Brass, M., & Wiersema, J. R. (2017). Spontaneous mentalizing in neurotypicals scoring high versus low on symptomatology of autism spectrum disorder. *Psychiatry Research, 258*, 15–20.

Phillips, J., Ong, D. C., Surtees, A. D. R., Xin, Y. J., Williams, S., Saxe, R., & Frank, M. C. (2015). A second look at automatic theory of mind: Reconsidering Kovács, Teglas, and Endress (2010). *Psychological Science, 26*, 1353–1367.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.

Press, C., Gillmeister, H., & Heyes, C. (2006). Bottom-up, not top-down, modulation of imitation by human and robotic models. *European Journal of Neuroscience, 24*(8), 2415–2419.

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2011). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General, 141*, 433–438.

Schneider, D., Slaughter, V. P., & Dux, P. E. (2017). Current evidence for automatic Theory of Mind processing in adults. *Cognition, 162*, 27–31.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science, 325*, 883–885.

Todd, A. R., Forstmann, M., Burgmer, P., Brooks, A. W., & Galinsky, A. D. (2015). Anxious and egocentric: How specific emotions influence perspective taking. *Journal of Experimental Psychology: General, 144*(2), 374–391.

Todd, A. R., & Simpson, A. J. (2016). Anxiety impairs spontaneous perspective calculation: Evidence from a level-1 visual perspective-taking task. *Cognition, 156*, 88–94.

Tsai, C. C., & Brass, M. (2007). Does the human motor system simulate Pinocchio's actions? Coacting with a human hand versus a wooden hand in a dyadic interaction. *Psychological Science, 18*, 1058–1062.

Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: How attentional selection is shaped by ascribing intentions to others. *PloS One, 7*, e45391.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.