**When Your Beliefs Guide My Hands:**

**Visibly Constraining an Agent Modulates Observers’ Automatic Belief Tracking**

**Jason Low1\*, Katheryn Edwards1, Stephen A. Butterfill2**

1School of Psychology, Victoria University of Wellington, New Zealand

2Department of Philosophy, University of Warwick, United Kingdom

**1. Introduction**

Tracking what is likely to happen in dynamic social situations is inherently challenging, and— because observing others activates our own motor systems—one view is that motor representations facilitate action and social understanding. Converging evidence revealed through brain imaging, transcranial magnetic stimulation (TMS), and reaction times suggests that observation of an agent’s action leads to activation of a corresponding motor representation in the observer (as reviewed by Rizzolatti & Sinigaglia, 2016). Neuroimaging studies support that the observation of actions done by others triggers activity in the motor execution-related brain areas that includes the ventral premotor and inferior parietal cortices (Cattaneo & Rizzolatti, 2009; Rizzolatti, Fogassi & Gallese, 2009). Consistent with imaging data, studies applying magnetic stimulation to the motor cortex during observation of reach-to-grasp actions reveal increased excitability in those parts of the motor cortex that correspond to the observed movement. For example, when adults watch another individual grasp an object, motor evoked potentials triggered by the stimulation are elevated in the very same muscles that would be active if the observers were making that grasping action (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Urgesi, Moro, Candidi, & Aglioti, 2006). Most importantly for the present study, behavioural studies profiling reaction-time benefits on motor priming tasks indicate that movement observation (a task-irrelevant stimulus showing an agent lifting her index finger, for instance) can accelerate our execution of the same action and decelerate our execution of an incompatible action (Brass, Bekkering, Wohlschlager, & Prinz, 2000; Cracco, Bardi, Desmet, Genschow, Rigoni, De Coster, et al., 2018).

Tracking others’ behaviour involves being sensitive not just to an action itself but also to the context of an action. There is evidence suggesting that our motor system can code the context in which specific actions are embedded. Iacoboni, Molnar-Szakacs, Gallese, Buccino, and Massiota (2005) monitored adults’ brain activations as participants watched three types of movies: clips of different background contexts (cup, teapot and plate of food arranged as if someone was about to have breakfast, or had finished); clips of a hand physically executing different grips (precision or whole-hand) to grasp a cup without context; and clips of a hand physically executing different grips to grasp a cup in the different contexts. The imaging data showed that the condition where participants witnessed grasping actions in embedded contexts, compared to the other conditions, elicited higher cortical motor activations. The results suggest motor representations can code more complex goal outcomes (grasp for drinking, grasp for putting away) of which the immediate action is a part.

Motor representations can carry information about the features of a situation even in the absence of any movement towards a target object. Costantini, Ambrosini, Tieri, Sinigaglia and Committeri (2010) required participants to execute a grip with either their left or right hand upon the presentation of a task-irrelevant go-signal (a virtual mug with its handle aligned to the left or right). The researchers found that elicitation of a spatial alignment effect—faster responding when the required hand to execute the motor act was compatible with the orientation of the mug handle—depended on the apparent possibility for participants to interact with the virtual object. The spatial alignment effect was elicited when the virtual mug appeared in near reachable space where it could be apparently acted upon by participants. In a follow-up study, Costantini, Committeri and Sinigaglia (2011) found that a bystander’s space constraints for potential action could also modulate participants’ own motor behaviour; the spatial alignment effect occurred also when the mug was presented outside the reaching space of the participants but within the reaching space of a computer avatar. The findings suggest that motor activations can be deployed even without watching movements being performed, such that anticipating someone else’s potential action can excite planning and control of actions in the observer to influence task performance. It appears that the range of some motor representations can go far beyond merely mapping the kinematic features of movement: they can even enable tracking others’ potential actions given a certain situation.

Costantini and colleagues’ (2011) findings that graspable objects potentially ready to an agent’s hand prime the execution of participants’ own hand movements might suggest that some motor programmes can be relatively effector specific. That said, Costantini and colleagues remind readers that there can be different levels of motor coding during action observation (see also Grafton & Hamilton, 2007; Rizzolatti & Sinigaglia, 2010). Whilst some motor representations can map the low-level parameters that compose action (patterns of bodily configurations, joint displacements or muscle activations) (Fadiga, Craighero, & Oliver, 2005), others can selectively carry information relating to the goal of an agent’s action (i.e., the anticipated or observed outcomes to which another’s purposive behaviour may be directed). There are TMS data indicating that motor representations code the goal of an agent’s tool-mediated behaviour and not necessarily the muscular pattern required to perform the movement, and that virtual lesions to the ventral premotor cortex impair judgments about the outcomes that bodily actions are directed to but preserves judgements about which body parts are being observed (Cattaneo, Caruana, Jezzini, & Rizzolatti, 2009). Likewise, when we witness an agent perform an action goal using a *non*-typical effector (kicking a ball with her finger or grasping a pencil with her foot, for instance), there is motor facilitation in the effector muscles that we typically use to achieve that action goal (i.e., the leg or the hand, respectively) (Betti, Castiello, & Sartori, 2015; Senna, Bolognini, & Maravita, 2014). Similarly, Triberti, Repetto, Costantini, Riva and Sinigaglia (2016) found that adults, after being trained to use their feet to grasp a ball, started to respond more as if they would typically use their feet to operate with other objects such as hammers and cups, which typically involve hand movements. Overall, the context of an action may be coded at multiple levels, whereby some motor representations that support social understanding may be relatively abstract for tracking the goal—and perhaps even the belief-informed goal—that an agent’s potential or overt action is directed to (Rizzolatti & Sinigaglia, 2016; Rizzolatti, Fogassi & Gallese, 2001).

If tracking the outcome-directedness of potential actions is an extension of motor processing, then constraining an agent’s opportunity to act should significantly disrupt motor representations concerning another’s realm of agency. Liepelt, Ullsperger, Obst, Spengler, von Cramon and Brass (2009) instructed participants to lift their index or middle finger in response to a number stimulus presented between the index and middle finger of a photograph of an agent’s static hand. There was a marked slowing of participants’ reaction time in the condition where the observed agent’s corresponding index and middle fingers were tied to the table with metal clamps as compared to the conditions where the agent’s fingers were unrestrained or where the agent’s non-corresponding fingers (thumb and ring finger) were restrained. The results highlight that the physical body of an agent is constitutively relevant to certain kinds of cognition. The fact that another person’s restraint leads to a finger-specific slowing of reaction time in the observer, even when the restraint was response irrelevant, suggests that our motor system automatically takes on the task-relevant circumstances governing someone else’s potentially upcoming goal-directed action. Dovetailing with Liepelt et al.’s findings, Costantini et al. (2011) found that priming of participants’ own motor system to react more quickly (as demonstrated with the spatial alignment effect paradigm) was obliterated when the agent’s apparent possibility to interact with an object was temporarily restricted (a transparent panel was placed between the computer avatar and the handled mug).

The available evidence that motor representations enables us to track the goals of others’ actions, although broad, does not entail that motor representations underpin the primary ways in which human beings engage in social cognition to play a role in the tracking of beliefs which specify someone’s reasons for action (Rizzolatti & Sinigaglia, 2010, p. 271). To make this bridge, our research experiment aims to test whether, and to what extent, visibly constraining an agent from potentially interacting with an object modulates observers’ abilities to automatically track that agent’s belief. Our experiment bears on two broader issues.

The first issue concerns how observers could ever distinguish between motor representations that are self-triggered and other-triggered. This issue arises given that witnessing someone else’s action leads to a direct activation of corresponding motor processes in the observer. The challenge can be acute. For example, patients with frontal lobe injury overly imitate someone else’s interfering finger movements that are not response relevant, and patients with Parkinson’s disease find it difficult to control automatic imitation of others’ postural instability as induced by motor contagion (Brass, Derfuss, Matthes-von Cramon, & von Cramon, 2003; Pelosin, Bisio, Pozzo, Lagravinese, Crisafulli, Marchese, et al., 2018). Neurotypical individuals, however, can enforce their own internally planned movement against an externally perceived movement. Bardi and Brass (2016) suggest that a deeper understanding of control of concurrently activated goal-directed motor representations may be gained by investigating its embedding in theory-of-mind (ToM) abilities. Consistent with this suggestion, studies show that there are overlapping brain activations when adults perform mentalising and motor-priming tasks (Spengler, von Cramon, & Brass, 2009). Moreover, training on a motor task that makes salient differences between the one’s own motor action and someone else’s motor action can transfer to improvements in spontaneously tracking another’s visual perspective on a ToM task (Santiesteban et al., 2012). This raises the question of whether impairing abilities to represent actions motorically might impair performance on a ToM task, which our experiment is designed to answer.

Second, in many social situations, successfully tracking an agent’s motor goal that specifies a target object also involves tracking the correctness of the agent’s belief about that object. Consider, for example, that Maxine is going to perform an action the goal of which is to grasp her ball. She falsely believes that her ball is in the right-side box, whereas actually it is inside the left-side box. If we ignore the Maxine’s false belief, then fixing only upon grasping the ball as the goal of her action would generate the wrong expectation of how it would unfold – incorrectly predicting that she would reach to the left-side box to grasp her ball (Butterfill & Apperly, 2016). We need to track Maxine's belief in order to correctly identify the potential motor outcome of her action, which is that she would reach to the empty right-side box to grasp her ball. This creates a challenge for action observation: if we are to track how an agent’s action will potentially unfold, we cannot always rely on how things are but must also take her belief or belief-like state into account. Could the motor representations critical for tracking how the agent’s actions will potentially unfold also play a role in tracking her beliefs? No study, until now, has demonstrated that motor representations also matter for successfully tracking of an agent’s belief.

Bardi and Brass (2016) conjectured that a functional connection between control of motor representations and belief tracking may be studied “in implicit ToM tasks where participants’ performance depends on concurrent activations of different representations of the environment” (p. 162). We explored Bardi and Brass’ conjecture by taking advantage of Kovacs and colleagues’ (2010) object-detection task, which yields a critical effect showing that adults’ rapid and automatic tracking of a bystander agent’s belief has an impact on observers’ own actions. The object-detection task involves adults observing a ball rolling behind a wall on a table and then rolling from behind the wall to off the table. The video sequence also includes an agent who was present for some of the ball’s movements but not others and, consequently, could have a true or a false belief about the ball’s location. In the outcome phase, the agent returns, the wall is lowered, and participants must use their hand to press a button as fast as possible if the ball is present behind the wall. The task is considered to target automatic belief tracking because no reference is made to the agent’s belief about the ball’s location, and the agent does not perform any overt action when the wall is lowered. The critical finding, which has been well-documented, is that, compared to a baseline situation in which neither the participant nor agent expected the ball to be present (P-A-), participants are faster to respond when only the agent expected the ball to be present (P-A+), implying that the agent’s belief regarding the ball’s location is automatically encoded (Bardi, Six, & Brass, 2017; Bardi, Desmet, & Brass, 2018; Deschrijver, Bardi, Wiersema, & Brass, 2016; Edwards & Low, 2019; El Kaddouri, Bardi, De Bremaeker, Brass, & Wiersema, 2019; Kovács, Teglas, & Endress, 2010; Nijhof, Brass, Bardi, & Wiersema, 2016; Nijhof, Brass, & Wiersema, 2017). Constrained

**1.1. Current Study**

The primary aim of our study was to determine whether, and to what extent, visibly constraining an agent from potentially interacting with an object modulates observers’ abilities to automatically track that agent’s belief. We measured elicitation of susceptibility to an agent’s belief (as indicated by the P-A+ < P-A- effect) using the ball-detection task. We adapated this task to create three distinct versions which differed in the constraints upon the agent’s potential movement in the outcome phase (see Fig. 1). In the Free-Agent version of the ball-detection task, the agent returned at the outcome phase wearing the same clothing that he wore during the first phase of the video; the agent was visibly free move and, potentially, to act on the ball. In the Constrained-Agent version of the ball-detection task, the agent returned at the outcome phase with his upper body and limbs bound up tightly in a white linen sheet (“mummified”); the agent was visibly constrained from using his upper body and limbs to potentially act on the target object. Extrapolating from research evidence showing that physical constraints on others’ action possibilities are mapped onto our own representations of the environment (Costantini et al., 2011; Liepelt et al., 2009), one possibility may be that the critical effect of the agent’s belief speeding up participants’ own action performances is a consequence of the way that information about beliefs feeds into motor predictions concerning the agent’s potential actions, and the way that those motor predictions would then facilitate response times. With respect to our primary research aim, we predicted that the critical P-A+ < P-A- effect would be elicited in the free-agent context but obstructed in the constrained-agent context; in other words, there should be a main effect of constraint on agent’s movement modulating automatic belief tracking.

A screenshot of a cell phone

Description automatically generated

***Fig. 1.*** *Logical structure of the study. There were three versions of the ball-detection task, presenting distinct constraints upon the agent’s potential movement when he returned to the scene (between-subjects manipulation): individuals experienced the Free-Agent version (no sheet), or the Constrained-Agent version (sheet restricted potential upper body movement) or the Loose-Sheet version (sheet allowed potential upper body movement). Each group of participants completed two forms of the ball-detection task (within-subjects manipulation): in the Hand-Placement form of the task, the agent placed the ball on the table with his hand; in the Foot-Placement form of the task the agent placed the ball on the table with his foot.*

One might argue that the elicitation of the critical effect in the Constrained-Agent version as compared to the effect’s obstruction in the Free-Agent version stems from the perceptual novelty of the agent’s (“mummified”) appearance in the former version’s outcome phase. To rule out this low-level possibility, we also studied elicitation of automatic belief tracking in a third version of the ball-detection task. In the Loose-Sheet version, the agent returned at the outcome phase with a white linen sheet (same yardage as in the Constrained-Agent version) placed loosely (like a tunic) over the clothing he wore during the first phase of the video (see Fig. 1). Consequently, in the Loose-Sheet version, the agent also returned with a novel appearance, but he remained visibly able to move his upper body and limbs, and, potentially, to act upon the ball. If perceptual novelty modulates belief tracking, then the P-A+ < P-A- effect should be obstructed in the Loose-Sheet and the Constrained-Agent versions. But if motor representations of action context modulate belief tracking, as per our primary prediction, then the P-A+ < P-A- effect should be elicited in the Free-Agent and Loose-Sheet versions but obstructed in the Agent-Constrained version.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Is the P-A+ < P-A- effect predicted in this version of the ball-detection task? | | | |
|  | Free-Agent version | Constrained-Agent version | Loose-Sheet version |  |
| Our prediction | Yes | No | Yes |  |
| Low-level perceptual novelty prediction | Yes | No | No |  |

A secondary aim of our study was to explore whether the P-A+ < P-A- effect would depend on there being a match between the effector used by the agent to act and the effector used by participants to respond. Considering the hierarchical structure of motor representations wherein information can be instantiated from effector-specific to effector-general levels (Rizzolatti & Sinigaglia, 2010; Sartori et al., 2015), any indication on whether the P-A+ < P-A- effect depends on matching agent’s and participants’ effectors would bear on whether, if motor representations of action do indeed drive the P-A+ < P-A- effect, those motor representations specify which effector the agent uses to act. Would the effect of belief tracking facilitating participants’ own representations of the environment be elicited only when the agent uses his hand to grasp and position the ball? Or might the P-A+ < P-A- effect also be elicited in were the agent to use his foot to interact with the ball? Each of the three versions of the ball-detection task (Free-Agent, Constrained-Agent, and Loose-Sheet versions; between-subjects manipulation) came in two forms. The Hand-Placement form of the task began with the agent grasping and placing a ball on the table using his right hand; the Foot-Placement form of the task begin with the agent grasping and placing the ball on the table using his right foot (within-subjects manipulation and counterbalanced). If motor representations modulating belief tracking are relatively effector-specific, then the P-A+ < P-A- effect should be elicited only, or more strongly, in the Free and Loose-Sheet agent versions where the agent had used his hand (rather than foot) to place the ball on the table; in other words, there should be an 2 (form: Hand-Placement or Foot-Placements) by 3 (version: Free-Agent, Constrained-Agent, or Loose-Sheet) interaction. However, if the motor representations specify goals but are neutral on the effector used by the agent, then the P-A+ < P-A- effect should be elicited in the Free- and Loose-Sheet versions and elicited regardless of whether the agent had used a typical (hand) or an atypical (foot) effector to position the ball; in other words, there should be a main effect of version that is not qualified by a form by version interaction.