When imitation falls short

The case of complementary actions

Faber, T.W.

Publication date
2017

Document Version
Final published version

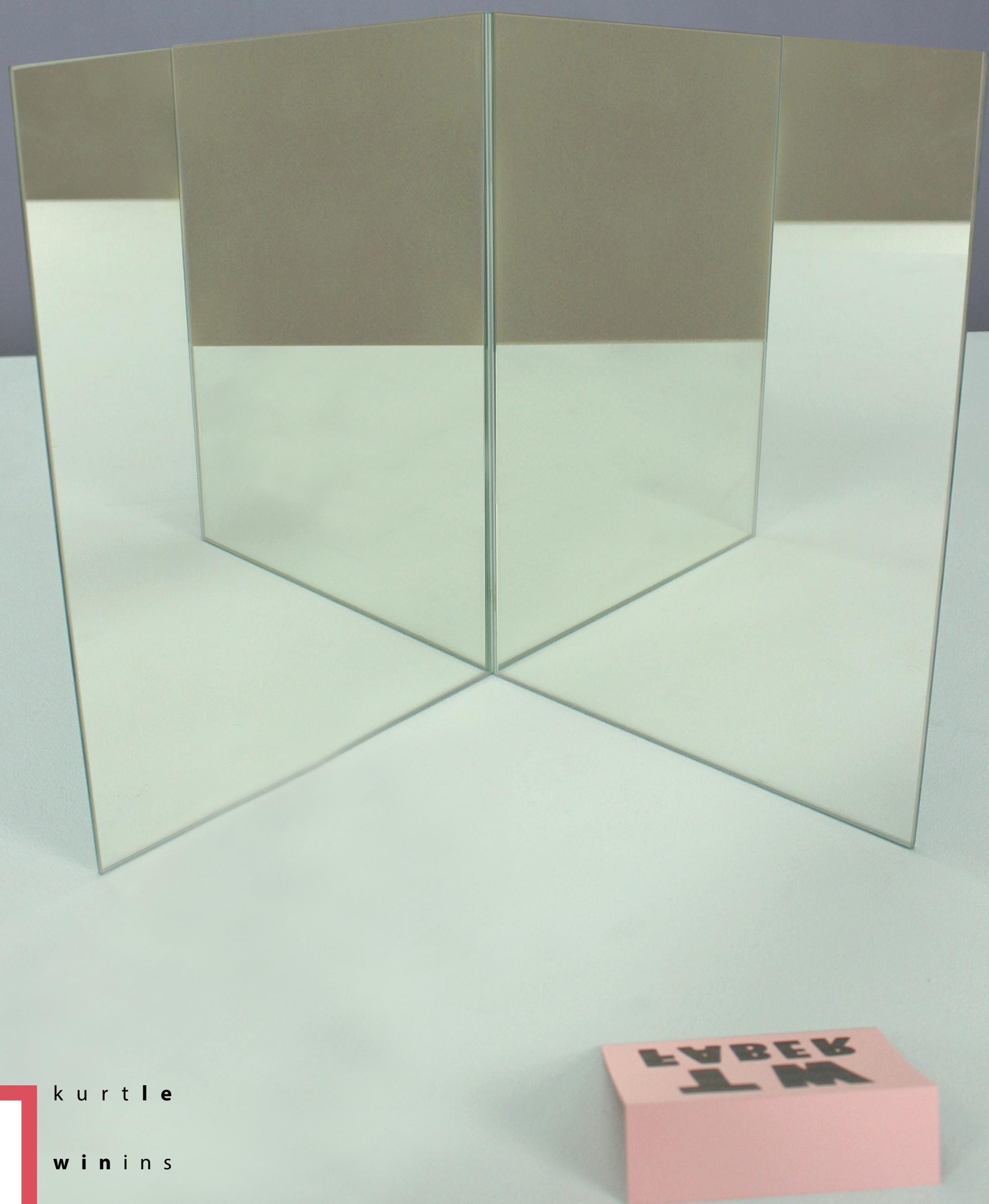
License
Other

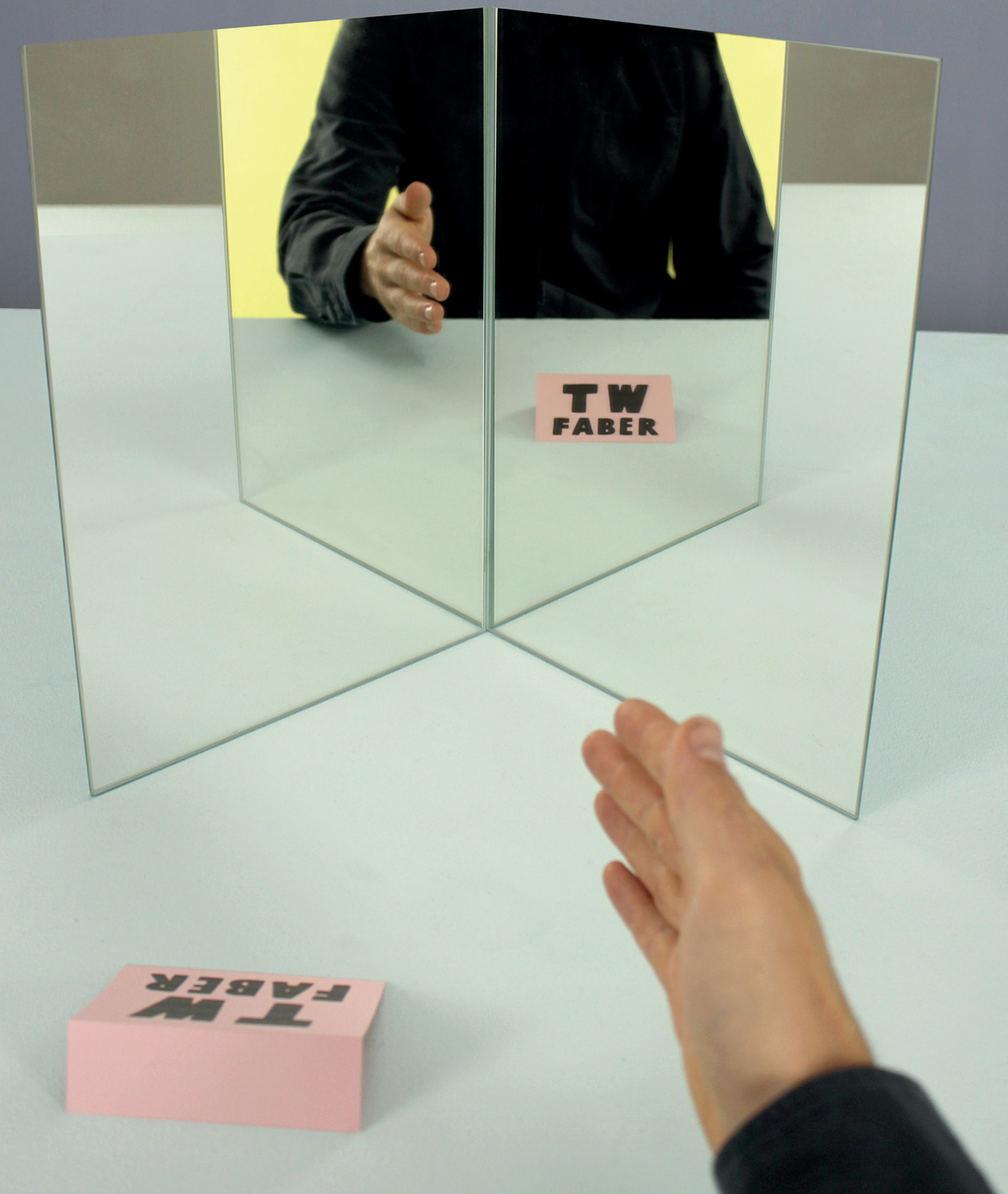
When Imitation Falls Short

~

The Case of Complementary Actions

Tim Willem Faber ~ When Imitation Falls Short





kurtle
winins
tituut

Dissertatiereeks Kurt Lewin Instituut 2016-19

When Imitation Falls Short: The Case of Complementary Actions

Tim Willem Faber

Colofon

ISBN: 978-94-028-049-8

Drukwerk

Ipskamp printing

Vormgeving

Sanne Sofia Broks

(sannesofia.nl)

When Imitation Falls Short: The Case of Complementary Actions

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Aula der Universiteit

op woensdag 25 januari 2017, te 13:00 uur

door Tim Willem Faber

geboren te Amsterdam

Promotiecommissie:

Promotor: Prof. dr. A. H. Fischer, Universiteit van Amsterdam

Copromotor: Dr. K. J. Jonas, Universiteit van Amsterdam

Overige leden:

Prof. dr. H. Aarts, Universiteit Utrecht

Prof. dr. D. Wigboldus, Radboud Universiteit

Prof. dr. K. R. Ridderinkhof, Universiteit van Amsterdam

Prof. dr. G. A. van Kleef, Universiteit van Amsterdam

Prof. dr. R. W. Holland, Universiteit van Amsterdam

Faculteit der Maatschappij- en Gedragswetenschappen

Contents

Chapter 1	General Introduction	7
Chapter 2	Perception in a social context: Attention for response-functional means	23
Chapter 3	Complementary hand responses occur in both peri- and extrapersonal space	41
Chapter 4	Perspective switching in social interactions: Does responding favor an egocentric perspective?	85
Chapter 5	Dissociating predictive from online motor activation: an EEG study	111
Chapter 6	Discussion	133

—

References	150
Dutch summary	176
English summary	182
Dankwoord	188

This project was supported by the Kurt Lewin Institute (KLI)

Chapter 1

General Introduction

General Introduction

Imitation has often been described as a defining characteristic of human behavior and evolution (Bandura, 1977; Carroll & Bandura, 1990; Chartrand & Bargh, 1999; Hess & Fischer, 2013; Preston & de Waal, 2002). The importance of imitation, or the copying of (goal directed) behavior, is evident from the writings of Darwin, who claimed that imitation serves as the motor for developing mental capacities such as language (Darwin, 1871/1989). Whereas basic forms of imitation are not exclusive to humans (Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009), imitation is argued to drive cultural evolution and separate humans from their predecessors. Numerous examples linking imitation to culture come from evolutionary biology (and psychology), such as the concept of the meme, or 'mimeme' translated as imitated thing (Dawkins, 1976). This refers to the production of cultural concepts through imitation such as culture specific styles or trends (e.g., fashion or music). Similarly, Legare and Nielsen (2015) proposed that cultural progress is a function of two primary processes: Imitation and innovation. Whereas imitation serves as a tool for acquiring instrumental skills as well as social norms, innovation serves as a tool to combine learned skills and applying them to novel situations. Accordingly, it is not imitation alone but the flexibility or selectivity in adopting imitative strategies that is essential for innovation and hence for cultural progress. Darwin already acknowledged that besides being a unique benchmark for higher functioning, 'blind' imitation is also what humans share with idiots, savages and monkeys (Willer, 2009) who "…unconsciously imitate every word which is uttered, whether in their own or in a foreign language, or every gesture or action which is performed near them" (Darwin, 1989). This provides a crucial distinction between automatic forms of imitation and more intentional forms driven by selection.

Evidence suggests that developmental as well as social-psychological factors, such as intentionality and group membership, contribute to selectivity in imitation (Uzgiris, 1981; Bourgeois & Hess, 2008; van der Schalk et al., 2011). For example, infants show selectivity in producing target actions when perceiving failed actions made by human actors but not machines, suggesting that they rely on the inferred intention of the actor when copying actions (Meltzoff, 1995). Also, infants seem to copy more complex actions and in a more exact manner at a later age (McCabe & Uzgiris, 1983) and infants and adults are affected by social dynamics such as identification with the actor or social group when imitating observed actions or facial gestures (Over & Carpenter, 2012; Bourgeois & Hess, 2008). These findings are often, but not exclusively, explained within a rational conception of imitation whereby imitation is facilitated when it provides the most rational path given the circumstances (Gergeley, Bekkering, Kiraly, 2001; Bekkering, Wohlschläger, & Gattis, 2000). This paints a picture of imitation as primarily an intentional form of behavior in purpose of instrumental learning and as a means to strengthen social connections to increase adaptability (e.g., through skill learning). Nonetheless, imitation has been popularized in cognitive and social psychological research largely as an automatic form of behavior. Whereas examples of intentionally driven imitation relate to distal causes of imitation, summarized in the *why* question (why do we imitate?), research on automatic forms of imitation often deal with the proximal causes of imitation, captured in the *how* question (how do we imitate?).

The *how* question is primarily aimed at solving the so-called correspondence problem (Brass & Heyes, 2005), which refers to the mismatch between the observation of an action and the subsequent reproduction of the same action. When an imitator observes an actor making a right hand movement the imitator only has information about the external (visual) consequences of the action (e.g., a right hand movement) observed from a third person perspective. Through some mechanism then the imitator has to use this visual input of the observed action and translate this into muscle commands in order to perform the same action, which is subsequently observed from a first person perspective. It is this translation that provides a problem: how does the imitator know what pattern of muscle activation mirrors or allows one to reproduce the observed movement?

A host of accounts have been put forward trying to tackle this problem, but I will highlight two prominent ones: Ideomotor theory (Prinz, 1990) and the theory of associative sequence learning (ASL; Heyes, 2001). Before explaining each theory in detail, it is important to first discuss the discovery of mirror neurons, which had a large impact on the development as well as reaffirmation of certain ideas central to both theories. Mirror neurons refer to a group of neurons that fire not only when performing certain actions, but also during the observation of the same actions (Blakemore & Frith, 2005; Di Pellegrino et al. 1992; Gallese et al. 1996; Rizzollati & Craighero, 2004). This effect was discovered initially in monkeys for which neurons in the inferior premotor cortex responded not only when a monkey grasped a piece of food but also when watching an experimenter perform the same grasping action (Di Pellegrino et al. 1992). Similar findings have been reported for humans where motor involvement during action observation was found in the ventral premotor cortex as well the inferior parietal lobule, homologous to the mirror effect previously reported in monkeys. Interestingly, mirror responses seemed to respond not only to goal-directed actions but also to actions that were not directed at an object or had any symbolic meaning (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995).

The discovery of mirror neurons was heralded as a critical scientific discovery. As stated by V.S. Ramachandran (1995), mirror neurons and imitation learning provided an understanding of the 'great leap forward' in human evolution. The reason or perhaps consequence of this popularity is the connections that were made between mirror neurons and a range of higher-order functions, such as action understanding and empathy, as well as the role of mirror neurons as presumed biological foundation for imitation behavior. At the same time, these higher-order attributions have inspired many opposing views (Hickok, 2009; Mahon & Caramazza, 2008). Nevertheless, mirror neurons provided a basis for solving the correspondence problem given their potential role in translating observed into performed actions and thereby solving the mismatch between action observation and performance. Although mirror neurons primarily reflect a covert rather than overt measure of imitation given that mirror activity is usually measured while a participant passively observes human actions, the link between mirror activity and overt imitation has been empirically demonstrated (Iacoboni et al., 1999; Iacoboni, 2009).

The ideomotor theory and the theory of ASL provide different proposals for the mechanisms behind automatic imitation, in part by linking mirror neuron functioning to the correspondence problem. I will briefly discuss each theory and discuss how they converge on the research question that is central to this dissertation. Subsequently, I will discuss the setup of each chapter in this dissertation and provide a short summary at the end.

Ideomotor theory

In short, ideomotor theory suggests that people perform actions upon perceiving them as a result of the shared representation of perceptual and motor features of the same action (Prinz, 1990). More specifically, actions are represented in terms of their perceptual consequences, which means that thinking about the outcome of the observed action in part triggers the motor commands used to produce it. The correspondence problem, as stated in Iacoboni (2009), is therefore not necessarily a problem for ideomotor theory since it is not necessary to translate the visual input of an action into a corresponding motor command given that both perceptual and motor features of the observed action are represented in the same way. Opinions diverge as to whether or not this representational overlap is an inborn quality in humans. For example, one nativist view of imitation has suggested that humans are endowed with an inborn mechanism, the so called active intermodal matching (AIM) mechanism, which provides a coupling between observed and performed actions (Meltzoff & Moore, 1997). A defining attribute of ideomotor theory in relation to imitation is similarity, which means that perception-action links are restricted to actions that are similar to observed actions (e.g., seeing a ball being kicked – kicking a ball). A simple demonstration of this idea is provided in Brass, Bekkering, Wohlschläger, and Prinz (2000). Participants in their study observed images of a left hand, depicted from a third-person (mirror) perspective. After a short delay, either the middle or ring finger made an upward movement while a number was presented next to the finger, instructing subjects to perform a pre-specified movement which was either similar (e.g., seeing an index finger moving – making an index finger movement) or dissimilar to the observed movement. It was shown that when the instructed movement was similar to the observed movement subjects responded faster compared to when it was dissimilar (a congruency effect). Although a congruency effect can be found even when using different effectors (e.g., mouth opening upon observing an open hand), the effect is strongest when observed and performed movements require a similar versus dissimilar effector (Leighton & Heyes, 2010).

Based on these and similar findings, a host of related theories and mechanisms have been proposed that build on ideomotor principles such as the perception-behavior expressway (Dijksterhuis & Bargh, 2001) and the direct-matching principle (Rizzolatti, Fogassi, & Gallese, 2001), summarized in Heyes (2001, 2011). These theories vary to the extent that imitation is defined as matching action kinematics (e.g., making a left finger movement), matching action goals (e.g., grabbing and manipulating a toy) or intentions (e.g., standing in line at the cash register). Furthermore, theories differ in the level of explanation they provide about the mechanics involved in facilitating automatic imitation, ranging from low-level biological theories involving the role of mirror neurons in coding observed actions and facilitating imitation to more social psychological theories about mental representations of behavior. Nonetheless, they share the idea of an automatic transformation of observed into performed behavior and provide similar examples in which imitation positively contributes to intergroup relations through, for example, social bonding or increased liking of those who imitate or mimic you (Chartrand & Bargh, 2001; Rizzolatti & Craighero, 2004)[1].

[1] Whereas imitation is the copying of goal-directed movements (e.g., grasping cup to drink), mimicry is defined as copying overt movements that are not directly goal-driven (e.g., touching face during a conversation). Both imitation and mimicry fit within an ideomotor

There is one essential discrepancy between ideas concerning automatic imitation and everyday life that is problematic, namely: We do not imitate all the time. We normally don't throw ourselves from a flight of stairs if we see somebody do the same, nor do we grab a box handed over by somebody because we imitate the actor's movement, but we are more likely to receive it and move it to another location. There are two explanations for this. First, spontaneous copying of actions is not a necessary outcome of a shared representation of perception and action since the activation of the representation of a motor movement does not always lead to the actual performance of the action (Brass, Ruby, & Spengler, 2009). A crucial factor regulating performance is action control, which has been used to distinguish automatic from more intentional forms of imitation (Heyes, 2009). Another important explanation is that individuals have certain roles in social interactions that require them to perform complementary actions (actions that serve as a response to observed behavior). These responses are more often dissimilar rather than similar to actions performed by interacting partners (Sartori & Betti, 2015). Nonetheless a sufficient amount of findings has inspired the idea that the production of actions follows the observation of actions in an automatic fashion and is usually limited to copying similar rather than dissimilar actions (Dijksterhuis & Bargh, 2001; Heyes, 2001; Iacoboni, 2009). It has for example been shown that priming words that represent social categories triggers automatic associated behaviors, such as priming a professor stereotype leading to better trivia performance or an elderly stereotype leading to slowing walking speed (for a summary see Iacoboni, 2009).

While ideomotor theory is driven by the similarity between observed and performed actions, the theory of associative sequence learning (ASL), makes room for a more flexible approach to perception-action coupling that allows for the development of meaningful (interactive, complementary) actions as well. An essential element of ASL is the role of sensorimotor learning which I will explain in more detail next.

Associative sequence learning

In ASL, coupling between perception and action is a consequence of repeated learning. As for ideomotor theory, the correspondence problem for an infant copying a parent's smile

---

framework given that they both entail similarity in observed and produced actions. Despite the similarity in structural characteristics, imitation is more often coupled to its function in instrumental or observational learning while mimicry is more often linked to its social function in creating rapport or affiliation (Wang & Hamilton, 2012; Lakin & Chartrand, 2003).

suggests that infants need to solve the discrepancy between seeing a smile from a third-person perspective and contracting muscles to produce the same smile for which they do not receive visual feedback. ASL suggests that this problem can be explained by the repeated coupling of a parent smiling at approximately the same time that an infant smiles. The infant then learns to associate the visual properties of a smile with the motor command used to produce a smile (Heyes, 2009). Rather than encoding the similarity between observed and performed actions, infants learn to associate the performance of an action to the observation of this action as a result of repeated experience. As a result, imitation is initially driven by parents copying infant movements and expressions rather than infants copying parents' movements. Over time, an infant learns the same link through other forms of feedback by self-observation (watching your hands while you move them), mirror-learning (watching yourself in the mirror) or during social activities. This view counters the idea that humans are hardwired to detect and translate observed actions into the performance of the same actions. A re-analysis of studies on neonatal imitation supports this idea by showing that there is little evidence for neonatal imitation across a host of physical and vocal gestures (Ray & Heyes, 2011). Besides providing disconfirming evidence for the innateness of shared representations, ASL, contrary to ideomotor theory, is not restricted to coupling between similar observed and performed actions. Rather, it serves as a domain-general explanation that favors links between any observed and performed actions that are a result of repeated (associative) learning, which has also been shown to modulate the responsiveness of mirror neurons (Cook, Bird, Catmur, Press, & Heyes, 2014). Consequently, ASL explains the development of imitative behavior but also the development of dissimilar perception-action links including complementary actions.

The importance of repeated learning is demonstrated in Catmur, Walsh and Heyes (2007). Using a similar paradigm to the one used in Brass et al. (2000), Catmur et al. (2007) trained participants to perform finger movements while concurrently observing either congruent (e.g., observing an index finger movement – performing an index finger movement) or incongruent finger movements. Results showed that observing finger movements led to stronger motor evoked potentials (MEPs) for the muscle used in the performance of the observed action. However, when trained to perform incongruent actions (e.g., observing an index finger movement – performing a little finger movement) the pattern reversed: Stronger motor responses were recorded for the muscle used for movements incongruent to the observed action. A similar effect was shown in van Schie, Waterschoot, and Bekkering (2008) in which participants performed either faster incongruent or congruent object-directed actions depending on the block instruction in which either congruent or incongruent actions were performed on non-target trials. Finally, a response benefit was found for social gestures (e.g., open hand gesture) as well, if responses were incongruent (and complementary) to the observed hand gesture (Liepelt, Prinz, & Brass, 2010).

The flexibility of ASL (Cook et al., 2014) suggests that observing actions is not limited to the (covert or overt) performance of similar actions, but allows for the development of non-similar perception-action coupling as well. This has inspired research demonstrating the automatic display of complementary actions in contrast to imitative (similar) actions in action observation.

Complementary actions

Important work that has contributed to the role of complementary actions (i.e., actions used for object or person coordination that are often dissimilar to observed actions) was done by Jonas and Sassenberg (2006) and Cesario, Plaks, Hagiwara, Navarrette, and Higgins (2010). Using a semantic priming task Jonas and Sassenberg showed that priming participants with words representing social categories can prime target words or action tendencies that reflect complementary actions to these categories. For example, priming social category words (e.g., doctor) facilitated responses representing complementary (e.g., to trust) relative to control words. Cesario et al. (2010) extended these findings by showing that the direct physical surroundings in which an approach-avoidance task was performed determined response facilitation, which suggests that taking account of environmental and social constraints is essential in preparing complementary responses towards others (Cesario & Jonas, 2014).

Similar evidence in which complementary actions were directly compared to imitative actions comes from the research of Luisa Sartori and colleagues. Typically in their studies (Sartori, Cavallo, Bucchioni, & Castiello, 2011; Sartori, Cavallo, Bucchioni, & Castiello, 2012; Sartori, Betti, Chinellato, & Castiello, 2013), participants observe short movies depicting object-directed actions, with either a whole-hand or precision grip, that either results in a non-social (i.e., the object is placed at a target location) or in a social action (i.e., the object is handed in the direction of the observer). In social action trials, the action depicts a complementary request that requires a non-similar response from the observer, while no request is implied in the non-social action trials. Using a TMS-MEP setup, stronger MEPs were recorded during non-social action trials for the muscle used to perform actions similar to those observed (covert responses for the muscle used to perform a precision grip while observing precision grip actions). In contrast, during social action trials stronger MEPs were recorded for the muscle used to prepare complementary actions when a complementary request was observed. This research shows that while motor involvement during passive observation (i.e., where no response from the observer is required) might be used to code observed actions, observing social actions (i.e., where a response is required from the observer) requires the preparation of complementary responses that are often non-similar to observed actions.

These findings might be explained by the fact that actions observed in an interactive context are deemed more important and henceforth should elicit stronger effects of motor involvement than non-social actions. Support for this prediction is provided by Oberman, Pineda, and Ramachandran (2007) who found enhanced neural responses in motor regions when participants observed a video in which three actors performed a ball tossing task which in some trials was directed to the observer compared to the same actions being performed amongst the group of actors or by an actor individually. Similarly, when observing partners in a physical interaction task, enhanced (covert) motor activity was found when subjects perceived interacting partners making object-directed actions compared to actors that were not interacting with the subject in non-target trials (Kourtis, Sebanz, Knoblich, 2010). Importantly, a study by Ménoret et al. (2014) showed that stronger motor activation was found when observing object directed actions that subsequently required a response from the participant compared to no response. This pattern of enhanced motor activity might in part reflect the stronger neural responses to complementary responses compared to imitative responses (Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007).

In addition to the role of action relevance, another important element for complementary actions is the predictive nature of action observation. Instead of passively responding or imitating perceived actions, observers actively anticipate upcoming actions, which in turn facilitates imitation and complementary actions interchangeably, depending on the (social) environment. Although there is limited empirical evidence for the predictive nature of perception-action coupling, there are some theoretical accounts that specify how this might be implemented on a neural level (Kilner, Friston, & Frith, 2007; Keysers & Gazzola, 2014). As argued in Keyers and Gazzola (2014), similar to associative sequence learning, observing actions is coupled to the production of actions through repeated learning. This favors similar connections given that the production of self-produced actions (e.g., open hand gesture) is naturally paired with the observation of the same (open hand gestures) rather than different actions (e.g., closed hand gestures). However, because of the delay in neural firing involved in the production of an action and the subsequent (delayed) visual feedback of the same action, visual feedback serves as a prediction of upcoming events. More specifically, predictions are being sent from motor areas to visual areas, so that observation of actions is coupled to the performance of subsequent actions. For example, when initiating a reaching movement towards an object, individuals will process the visual consequences of this action at a specific delay, such that their hand is almost reaching the object, which triggers the subsequent grasping of the object. This can be nicely translated to social settings in which throwing a ball to a friend produces the visual feedback of a friend catching the ball, which in turn triggers the subsequent preparation to receive the ball in return. Accordingly, action observation not only triggers non-similar complementary actions, the translation between observing and performing actions is driven by a predictive mechanism in which action outcomes are anticipated prior to observation.

Evidence so far has shown that motor activity is not only recorded while participants observe an object directed hand movement, but also when the final part of a grasping sequence is occluded (Umilta et al., 2001), demonstrating that motor activity does not strictly depend on visual input. Additionally, when observing actions that were fully predictable by contextual cues (e.g., a color cue predicting a precision grip toward an object), stronger motor system activity was found prior to action onset, compared to trials in which no movement was predicted to take place (Kilner et al., 2004). Whereas most of these effects were found for adults who commonly have had abundant sensorimotor experience, the same predictive abilities have been found for 15-month infants (Southgate, Johnson, Osborne, & Csibra, 2009).

Before I present the central goal of this dissertation I will briefly summarize the research and theories presented so far. According to ideomotor theory, observing actions trigger the (covert) preparation of similar actions (e.g., mental representations of actions), possibly through the involvement of mirror neurons, thereby facilitating overt imitation. From the perspective of ASL, imitation is a result of sensorimotor experience that allows coupling between the observation of actions and the performance of similar as well as dissimilar actions. These dissimilar actions include complementary actions that are not only dissimilar, but also meaningful and therefore play an important role in interactive (social) situations. In this dissertation I examine which factors play a role in the development of imitative and complementary actions through associative learning and to what extent this learning process has predictive properties.

This dissertation

Building on previous research, I have focused on the role of perception-action coupling in social interactions, using the theory of associative sequence learning as well as the role of predictive associations as described in Keysers and Gazzola (2014). I will argue that imitation, which is often deemed as a crucial human capability, might be irrelevant in day-to-day life and suboptimal in social interactions specifically. Rather, in social interactions complementary actions are often more appropriate and more fitting. I will limit my argument to automatic forms of imitation given that I acknowledge the relevance of observational and imitative learning in infancy and the role of social imitation throughout life. These forms, however, reflect more intentional and higher-order goal driven imitation that surpass explanations based on simple sensorimotor transformations that rely on perception-action similarity (Wang & Hamilton, 2012; Over & Carpenter, 2012). Note that these explanations do not have to be mutually exclusive (Heyes, 2011).

In order to investigate the main proposition I have used the theory of associative sequence learning to come up with a range of factors that are important in motor learning. A crucial feature of imitation that applies to all proposed factors is that it requires matching the observation of a self-produced action, commonly seen from a first or egocentric perspective, to the same action seen from an allocentric perspective (as perceived by somebody else). As a result, copying an open hand gesture is a result of transforming the visual input of the actor's open hand (e.g., in mirror perspective) to the motor production of the same open hand. Motor production in semi-interactive tasks is then typically measured as a covert rather than overt response (Sartori et al., 2011). In contrast, complementary actions are often a result of repeated overt motor learning, particularly in social settings. We lift someone up who just fell down, we shake the hands of somebody extending theirs and we catch the ball that somebody throws at us. In this case, no translation between visual feedback from an egocentric to an allocentric perspective is necessary given that complementary actions are commonly produced in response to actions seen from an egocentric perspective.

Another specific factor important in motor learning is that of (social) context. As argued in Cesario et al. (2010) as well as Cesario and Jonas (2014), complementary actions are learned in relation to the physical context in which they provide a meaningful response. This fits with the idea that perception-action coupling is defined by the combined coding of perceptual and motor features structured by events in which these features are deemed relevant (Hommel, Müsseler, Aschersleben, & Prinz, 2001). The fact that some actions in response to others are often performed in one context but less so in another context, facilitates context-dependent learning. Heyes (2015) makes the point that ASL shows sensitivity to contextual cues in coupling observed to produced actions by means of contingency learning (Cooper, Cook, Dickinson, & Heyes, 2013; Keysers & Gazzola, 2014). Contingency learning means that actions can be triggered by a visual stimulus given their repeated co-occurrence in a specific context, but the absence of an action related to this stimulus reduces the stimulus-action association. For example, repeatedly producing a right finger movement upon perceiving somebody shake their head in context A, will constitute learning over time that seeing somebody shake their head will result in the automatic tendency to produce a right finger movement in the same context. However, when a finger movement does not follow the head shake in a number of instances this tendency will decrease for context A, but not necessarily for context B. In Chapter 2 of this dissertation I use the role of context in motor learning to argue that complementary actions are primed in response to social categories (e.g., athletes, criminals) specific to the context in which these actions are produced. For example, whereas defensive behavior in response to a threatening person (e.g., criminal) is fitting when you find yourself in an abandoned street, it may be less useful when observing the same person in a courtroom. To test this hypothesis I used gaze tracking in order to track subject's attention for objects used in complementary actions (e.g., defensive means), both when observing a social category member in a context that triggers the complementary action or not (e.g., criminal observed in a court room). I hypothesized that subjects would pay more attention to objects used in a complementary action when observing a social category preceded by a context that facilitates the complementary action rather than inhibits it. Two studies are reported using different social categories with amount of attention as dependent variable.

The third chapter uses a simpler paradigm to test the same principle, but here I explicitly compared complementary actions to imitative actions and focused on more low-level actions (i.e., hand gestures). Based on the role of context in motor learning, I argued that observing social hand gestures (i.e., an open hand) triggers the production of complementary hand gestures (e.g., handshake), but only when perceived within physical reach. Because performing a handshake depends on physical interpersonal distance, seeing an open hand gesture in peripersonal space (the space surrounding the body) should trigger complementary responses, whereas it should not trigger these responses when observed outside of peripersonal space. As in Chapter 2, space therefore provides a contextual cue that predicts complementary actions. At the same time, actions that trigger imitative actions (e.g., meaningless actions) are not necessarily learned in a specific context and mostly rely on internal simulation of observed actions irrespective of differences in space. Therefore, such actions are not expected to be dependent on spatial parameters. I predicted that participants would be faster in performing complementary actions in response to social hand gestures but only when perceived in peripersonal space. Five studies are reported testing this hypothesis using a manual response paradigm with response time estimates as dependent variable.

The fourth chapter taps into the role of perspective taking in producing complementary actions. As explained in Keysers and Gazzola (2014) and in more detail in Vogeley and Fink (2003) and Birch and Bloom (2004), coupling between observed and performed actions is shaped primarily by visual feedback from a first-person perspective. For example, left-handed actions are seen from an egocentric perspective where the fingers extend away from the palm in the left visual field. Although it is argued that over time action information is translated to multiple viewpoints (i.e., viewpoint invariance), it seems that motor priming effects seem to be strongest when observing actions from an egocentric rather than allocentric or third person perspective (Jackson, Meltzoff, Decety, 2006). This difference might stem from the fact that different types of perspective taking rely on different neural mechanisms (Vogeley et al., 2004; Ramnani & Miall, 2004). I adapted the idea of preferential processing of information from a first-person perspective to social interactions. While it has been argued that social situations increase allocentric perspective taking due to the increase in social relevance (Tversky & Hard, 2009; Zwickler & Muller, 2009), I proposed that social interactions that require complementary actions trigger egocentric rather than allocentric perspective taking, which should lead to a decrease in voluntary perspective taking in social interactions compared to social settings that do not require complementary responses. In addition, when instructed to take the perspective of others in a social interaction this should interfere with the tendency to adopt an egocentric perspective (i.e., egocentric interference; Surtees, Butterfill, & Apperley, 2012; Vogeley & Fink, 2003). To test these hypotheses I adapted a spatial perspective taking task from Tversky and Hard (2009) in which subjects had to judge, using pictures, the relative position of two objects on a table, while an actor was either holding or handing over one of the objects on the table. Using response time measurements I tested whether participants would be faster in making spatial judgments from an egocentric or allocentric perspective, if the situation depicted an interactive rather than non-interactive setting. I report three studies aimed at answering this question.

The fifth chapter looks specifically at the role of prediction in motor learning. As proposed earlier, repeated perception-action coupling shapes connections between the performance and observation of actions. In ASL the connection between observed and performed actions is fairly unrestrictive in that observing stimuli can trigger both the production of the observed action as well as actions commonly performed in response to the observed action. In addition, besides providing a direct means for imitative behavior (overt action performance), action observation is argued to trigger simulative and predictive processes used to anticipate others' future actions (Wilson & Knoblich, 2005; Kilner et al., 2007). One common way to test predictive properties of action observation is to use action sequences (e.g., goal-directed hand movements), in which prior information is provided about the action that is displayed in the next part of the sequence. Most studies to date have selectively focused on either the time window prior to action onset or have focused on the average response across the observation phase. Additionally, studies that have looked at the action anticipation phase (i.e., prior to action onset; Kilner et al., 2004; Southgate et al., 2009) have only used fully predictable action sequences, while studies that have only studied the action observation phase have typically used unpredictable action sequences as a result of trial randomization (Fadiga et al., 1995). To get a full picture of the role of prediction in action observation, I showed participants short videos depicting an object-directed hand movement (e.g., hand making a precision grip towards an object) and included both the action anticipation phase (prior to movement onset) and the action observation phase (during movement). Additionally, I varied the information about upcoming actions by using both predictable and non-predictable trial sequences. I hypothesized that motor activation prior to movement onset, measured using EEG-indices, would be strongest for predictable actions relative to unpredictable actions. In contrast, during action observation, motor activity should be stronger for unpredictable actions relative to predictable actions reflecting enhanced monitoring of unpredicted action outcomes (Stapel, Hunnius, van Elk, & Bekkering, 2010). I report a single study using both event related potentials and time frequency data to test my hypotheses.

Preliminary summary

The chapters in this dissertation are aimed to question the usefulness of imitation in day-to-day settings and social interactions in specific. To answer this question I have used the notion of motor learning as specified in ASL to come up with specific hypotheses. I have tried to argue that complementary actions are more common in social interactions and provide an alternative to imitative actions that rely on an internal simulation of observed actions in which repeated (overt) learning does not play an important role.

The overview presented in this chapter has some flaws which I feel are important to present (preemptively) here. The argument pertaining to the usefulness of automatic forms of imitation and the relation between imitation and ASL were not clear at the start of this project. Rather, these ideas are a product of extensive changes to the main question and theoretical background over the last four years. For example, the initial aim of the project was to look at the role of visual attention in terms of complementary actions. Later on I decided that attention would better fit as one of the dependent variables rather than the primary dependent variable. Furthermore, the second chapter was changed from a project not part of the main research line into a main project given that it fit with the broader research question. So even though the introduction reads as a 'fairly' consistent set of hypotheses with accompanying studies, reading through the chapters might make the reader miss the overview. While some projects seem to have a consistent plan of action from the start I have experienced this to be practically impossible and even undesirable when confronted with new ideas and insights.

Besides these comments, I want to take the opportunity to make a very brief statement about the current research climate in psychology (and outside). Although debates concerning open data, pre-registration and QRPs are prominent at the University of Amsterdam I have noticed this is not necessarily the case elsewhere. Extensive contributions from networks of psychological researchers (e.g., Manylabs 1; Manylabs 2) have nonetheless stressed the importance of this discussion and the problems that not adhering to standards of open science generate (Simmons, Nelson, & Simonsohn, 2011; Ioannidis, 2005). Specifically social psychology has dealt with a fair share of related problems such as the inability to replicate major findings in the field (e.g., https://osf.io/92dhr/). Similarly, the literature discussed here, in specific the early findings on mimicry and automatic imitation have dealt with issues of non-replications (e.g., Doyen, Klein, Pichon, & Cleeremans, 2012) which affects the theoretical basis of processes underlying imitation and partly inspired the current research effort. To this end, I want to stress that the only way forward is to change common practices by means of a few key principles: a) Pre-registration b) Data/material sharing c) Power analyses. At the same time, institutions that are integral to the research culture should be open to adapt their hiring and contract policies that are still driven mostly by publication criteria, which maintain the incentive structure that has created most problems in the first place. Only by making changes on both the personal and institutional level real progress can be made. This requires both old and new generations of scientists (and policy makers) that are willing to put in the effort. In this dissertation I have tried to contribute where possible to the principles addressed here (e.g., data sharing, pre-registration, power analyses) and aim to extent these practices in the future.

Chapter 2

Perception in a social context: Attention for response-functional means

This chapter is based on: Faber, T. W., & Jonas, K. J. (2013). Perception in a social context: Attention for response-functional means. *Social Cognition*, *31*, 301-314. doi: 10.1521/soco.2013.31.2.301

Abstract

Research on automatic behavior has shown how social category priming can activate unique responses towards such categories. Recently, the importance of the context in which social category members are perceived in has been demonstrated for the response selection process. While response selection has been investigated, other dependent variables, such as visual attention processes have yet not been tested. In line with top-down perception theories, visual attention plays an important role in response selection. The attention processes that precede response selection in social encounters remain unclear, namely which functional means are focused on to determine a response proper. We conducted two gaze-tracking studies in order to test how attention for context cues is influenced by changing environments. Our results show that context determined attention to functional means for the behavioral response, too. Our study provides first evidence that contextualized social category primes affect visual attention for response functional means.

Introduction

When approached by a threatening person one can, in an instant, think of a number of ways to respond. The behavioral response depends, however, on the context one finds oneself in when this happens. Trapped in a dead end street at night one might see neither the potential to flee, nor the use in calling for help and thus feel forced to lash out at the attacker. Encountering the same person in a lively street next to a police station one would probably adapt their response to the situation by feeling less forced to fight back.

The importance of the context, as illustrated above, is a recognized idea in studying selection of responses in general (Lewin, 1939; Prinz, Aschersleben, & Koch, 2009; Smith & Semin, 2007) and has been picked up in current automatic behavior research. Based on an extension to the classic behavioral priming paradigm (Bargh, Chen, & Burrows, 1996), Jonas and Sassenberg (2006), and Cesario, Plaks, and Higgins (2006) could show that social categories also activate response behaviors directed towards members of that category (i.e., behavior directed as a response to stereotypical behavior of the primed category). Recently, Cesario, Plaks, Hagiwara, Navarrette, and Higgins (2010) have extended this idea. They could show that the activation of response behaviors is dependent on the physical context in which the social category member is perceived. In another prominent line of research, the shooter paradigm, defensive responding (to shoot or not to shoot) has been shown to depend on the specific context in which a social category prime is placed in (Correll, Wittenbrink, Park, Judd, & Goyle, 2011).

A vital and precursory part of automatic response selection, in which the environment is perceived in order for a suitable action to be executed, has not yet been investigated. Response selection, as shown by Cesario et al. (2010) and Correll et al. (2011) is driven by context. Context entails information relevant for generating action, as well as facilitating or inhibitory information (e.g., objects, other social categories, and scenery) that determines which path of action is pursued. We argue that attention to response-functional means (e.g., weapon as a means of defense) can play a crucial role in the perception of one’s own ability to produce a response (e.g., resource-holding potential; Cesario et al., 2010; Parker, 1974).

In the current research, we will present two studies using visual attention measurement that show how priming social categories in different social contexts even affects attention for means functional for later response options. More specifically, visual attention is directed to means only when they are useful for response behaviors afforded by the context in which a social category member is being perceived. This way we can investigate a crucial link between the placement of a social category in a context and the behavioral response.

Context-dependent response behavior

Context dependency moved into focus in automatic behavior research within two paradigms. First of all, using the response-priming paradigm (Jonas & Sassenberg, 2006), Cesario et al. (2010) demonstrated how changing the direct physical context of participants constrained the number of possible response options. Participants were seated outside in an open field or in an enclosed booth in the lab while doing the same priming task. In the task, participants were shown either fight-related, escape-related or non-words and were asked to categorize these words by pressing “fight” or “escape” as labeled on a two-button box. During this task, pictures of African American or Caucasian males were primed simultaneously with the target words. Results showed that participants in the open field were faster at categorizing escape-related words whereas participants in the enclosed booth were faster at categorizing fight-related words following pictures of African American males. The findings demonstrate that showing a defensive response is moderated not only by the association between danger and social category membership (African American vs. Caucasian) but also by the context, (booth vs. field) in which the social category member is perceived.

Secondly, experiments within the shooter paradigm (Correll, Park, Judd & Wittenbrink, 2002; Payne, 2001) have looked at response selection using a first-person-shooter game in which pictures of either black or white males were presented carrying either a gun or a gun-like object. Consistent over studies a racial bias was found: Black males are shot at faster when holding a gun compared to white males, and more mistakes are made shooting a black male holding a gun-shaped object relative to white males. Correll et al. (2011) demonstrated that changing the context in which the social category member is perceived in from a safe to a threatening one, can eliminate the racial bias. More specifically, participants had a tendency to shoot more at targets situated in threatening environments irrespective of the race of the targets. In safe environments, this tendency to shoot was consistent with the racial bias pattern found earlier. The authors argued that people are predisposed to shoot as a reaction to dangerous cues that are part of the context, suggesting that the context in which others are perceived is capable of overriding a response based only on the association between danger and the target person.

Both approaches tap into response selection, which can be interpreted as an outcome of situation perception. Yet, certain responses are linked to the use of a mean, for example to shoot a Black or White person requires a gun, too. The purpose of the current research was to unveil this process by testing how response selection leads people to perceive their environment in a top-down fashion (Bruner, 1957). The question we pose here is: Does the change in social settings modulate attention for context cues that are only useful for later responses?

Visual attention for social and contextual cues

Attention for one's direct environment is driven by two forces, one in which attention is being grabbed, bottom-up, by visual features in the perceived environment and one in which attention is directed, top-down, by cognitive factors towards relevant visual features (Henderson, 2003; 2007). The idea that the social category-response behavior link affects attention for functional cues in the environment is in line with a top-down control of visual processing. Perceiving others in context can prime behavioral responses, which drives attention towards functional cues in the environment.

We rest our reasoning on research looking at attention for attractive others. There, Leder et al. (2010) also found attention holding for attractive faces to be modulated by the context in which they are presented. In their first study, Leder et al. (2010) found that pictures of attractive individuals received more attention than non-attractive counterparts. In the second study the situational context in which the target individuals were perceived (social approach vs threat) was manipulated in a writing task prior to the attention task. The results showed that male participants in the social approach condition looked longer at attractive than non-attractive faces. In the threat condition participants no longer looked at attractive males more than unattractive faces. The authors suggested that an approach tendency towards attractive stimuli could be either strengthened through social approach while discounted by threat. More specifically this means that, in the threat condition, the attention for only the attractive male, now as potential aggressor, was moderated by situational demands that determine one's adaptive response in the environment. For our research question, simply using attractive faces as stimuli (as in Leder et al. 2010) is not sufficient since we are interested in means functional to the response towards the social category. These means are adaptive to the situational demands since they function as means to perform behavioral responses primed by the perception of others (Jonas & Sassenberg, 2006). So the value of functional means can lead to attention holding for these means while less attention is paid to non-functional, irrelevant means. Also, our experimental paradigm is more complex than the context manipulation and attention holding measure employed by Leder et al. (2010). Instead of manipulating the context in which response means are perceived in a task separate from the attention task (Leder et al., 2010), here we implemented the context manipulation in the attention task itself.

Based on Jonas and Sassenberg (2006) and Cesario et al. (2006) we now know that representations of social categories can activate specific response behaviors. The studies by Logan and Bundesen (2004) and the reasoning in Prinz, Aschersleben and Koch (2009) suggest that separate representations of contexts and specific social categories can be integrated into new compounds of information. These compounds, or events (Hommel et al., 2001; Zwaan and Radvansky, 1998), activate associated behavioral responses as demonstrated in Correll et al. (2011) and Cesario et al. (2010). Thus, perceiving events in a top-down fashion should be driven by the representation of a response towards a category member as shaped by the context that person is perceived in. For example, giving money to a cashier in a store. In this case, the typical response towards a cashier in this setting is paying for a purchase. This information will subsequently determine which response means available in the environment are deemed useful in order to respond. The functional means in this case are any means used to pay for an item (e.g., cash).

Overview of the Current Research

To sum up, evidence from behavioral priming studies suggests that perceiving social category members can activate specific response behaviors modulated by the context in which they are perceived. In order to better understand how this response selection process depends on the direct visual context we investigated attention for response functional means.

We hypothesized that the perception of social category members in a specific context would lead people to pay more attention to means functional to perform a response behavior determined by the context, relative to control means. We conducted two studies in order to test our hypothesis. In Study 1, we used pictures representing different contextual information (safe vs unsafe) in which different social categories (criminal vs. control) were perceived. The goal was to show how a change from a safe to an unsafe context would affect the attention for means used in self-defense against criminals. In the second study we aimed to structurally replicate the effect of Study 1, but then by using neutral stimuli instead of negatively valenced stimuli, thus resting our claim on mean attention not being driven by valence.

In the design of both studies the focus is on the time frame of attention from the perception of others in their direct context until the behavioral response is performed. Therefore, attention persists over time as a process preceding action. A measure of attention holding, or fixation duration, is therefore most suitable for this purpose.

Study 1

In the first study, our aim was to manipulate and provide the context in which a social category member was perceived. A 3 factorial within-subjects design was used with a context picture (safe vs. unsafe), a social category word prime (criminal vs. control) and means pictures (functional for self-defense vs. control means). Our main hypothesis, expressed in a 3-way interaction, was that critical category members (here: criminals) perceived in an unsafe context would lead to more attention allocation to means of self-defense, versus control means, compared to the same category members perceived in a safe context (H1). Alternatively, if only stereotypic associations or semantic accessibility would determine attention then, irrespective of the presented context, people would pay more attention to self-defense means, versus control means, following matching social category primes (here: criminals) (H2). For both the main and alternative hypothesis, no difference in mean attention to functional vs. control means was expected in the control trials. This is because neither self-defense nor control means provide a functional role for possible responses to the control category (here: athletes) with respect to the context (safe vs unsafe) they are perceived in. Mean attention holding was measured using gaze tracking.

Participants

In total, 40 participants (of which 36 females) took part in the experiment in turn for either course credit or monetary compensation (7 Euros). The mean age was 22 years (range = 17-36).

Materials

Visual Attention Task

Participants started with a visual attention task as last of a series of studies all executed on a desktop computer. The preceding studies were unrelated to the target study and included, among others, an attention task looking at product attention. Participants were told that they would take part in a task looking at object perception. This task consisted of a number of visual trials. Each trial in the task started with the display of a fixation marker in the center of the screen ("+") which lasted for 500 ms. Following the fixation marker a context picture was shown for 500 ms in the same position as the marker (each 357 by 500 px). Two categories of context pictures were displayed representing either safe (including a police car), or unsafe (including a deserted metro station) social contexts. After the context picture a word was primed in the center of the screen for 100 ms, depicting either "CRIMINEEL" [criminal] or "SPORTER" [athlete] representing the social categories. Subsequent to this social category prime two pictures in color of means (each 200 by 200 px) were presented simultaneously, one on each side of the screen with a distance of 300 px from the center of the screen. For each trial on either the left or the right side the picture displayed a mean functional to a behavioral response towards the primed social category of criminals with on the other side a control means. Among the response related means for the categories were different weapons (e.g., knife, gun) and a TV or a couch as control means (see Fig. 1). The position of the pictures in both categories of means was counterbalanced so that they would appear an equal amount on the right and left side of the screen. The picture set was presented for 1000 milliseconds. As a cover task, participants had to do a simple arrow indication task following each picture set. In this task people where shown arrows (25 by 50 px), randomly and evenly displayed on either side of the screen in the same position as the pictures in the picture set, and were asked to press either the right or left arrow on the keyboard that matched the direction of the displayed arrow. The task lasted for 64 trials. Before the main task a few practice trials were done with stimuli different from the ones used in the main task. Between each trial a blank screen was presented for 100 ms. As a dependent measure we looked at the gaze duration for the target pictures. After the task participants had to fill out an exit-questionnaire followed by a debriefing after which they received their compensation.

Gaze tracking Device

Attention holding, that is fixation duration, was recorded using faceLAB 5 gaze tracking hardware and software (Seeing Machines, 2009). The setup of the hardware, consisting of a set of two cameras, was mounted underneath the 19 inch wide computer screen (1680 x 1050 px resolution) with a refresh rate of 60 Hz. Eye-movements were recorded and written to a Presentation (Neurobs, Inc., Albany, California, USA; www.neurobs.com) extension of faceLAB (faceLAB_ET2, version 1.1, Molenkamp, 2010). The data was used to calculate focal gaze duration from the onset of the picture set presentation averaged over consecutive trials for the conditions of the manipulation (e.g., safe context and criminal prime).

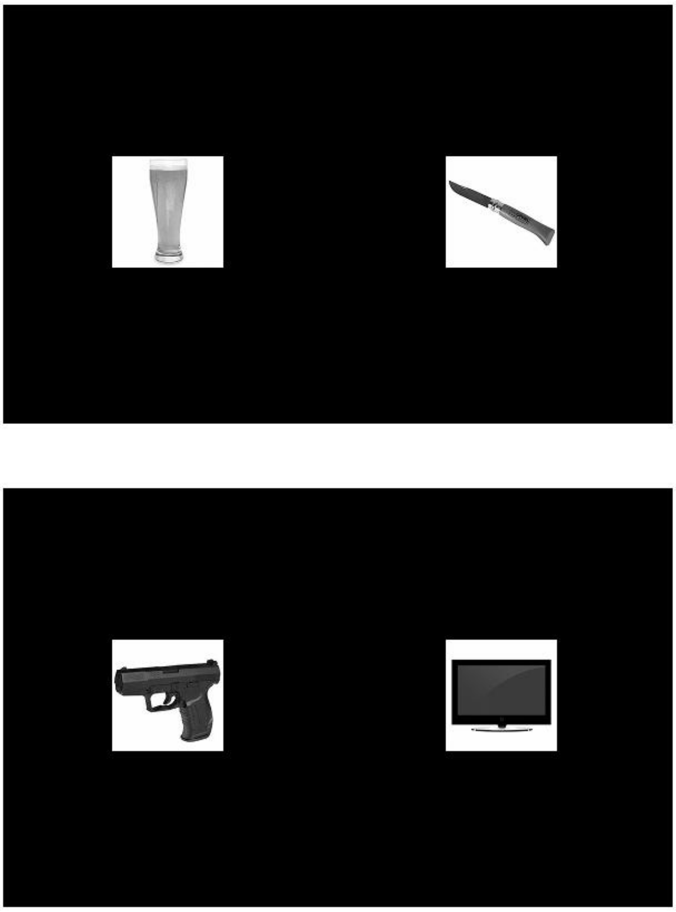


Fig 1. Sample pictures of means in an array as presented on the computer screen.

Manipulation check

No sex or age differences were found for differences in mean attention between both categories of context pictures or social category primes, *F*s<1. As a manipulation check we calculated ratings of safety of the safe and unsafe contextual primes. The data, taken from the exit-questionnaire, showed that predefined pictures of safe environments were rated as safer (*M* = 4.43, *SD* = 1.14) than predefined pictures of unsafe environments (*M* = 2.88, *SD* = 1.03) on a 7-point Likert scale, $F(1,40) = 59.12$, $p < 0.001$, $\eta^2_p = 0.596$. As part of the exit-questionnaire we asked participants to predict the rationale behind the study. Our results indicated that none of the participants seemed to be aware of the true purpose of the study.

Results

We first calculated the length of gaze duration during the presentation of the pictures of means. Subsequently, a frame was created around both of the pictures during the means presentation and only the attention directed inside this frame was analyzed. In case of too much information loss, for example when individual gaze tracking configurations failed, we removed the data for these participants.

To test our main hypothesis we performed a repeated measures analysis with factors of Context (safe vs. unsafe), Category (criminal vs. athlete) and Means (functional vs. control). The first analysis showed a significant 3-way interaction between Category, Context and Means as factors in the design, $F(1,39) = 9.65$, $p = 0.004$, $\eta^2_p = 0.198$, qualifying all lower order effects[2]. In order to further break down the relevant 3-way interaction and to simplify the investigation of our hypotheses we decided to calculate difference scores for the factor Means. By doing this we were able to account for a difference between mean attention time for functional and control means and subsequently compare these estimates between different contexts.

The difference scores were computed by subtracting mean gaze duration for control means from mean gaze duration to functional means. A positive value represents a higher amount of gaze duration for means of self-defense versus control means, measured in number of frames. We used the difference scores to look at differences in mean attention between levels of the category and context.

To test our hypothesis we conducted a repeated measures analysis using the difference scores as DV. No main effects were found for context or category, *F*s<1. We did however find the expected interaction between context and category, $F(1,39) = 7.13$, $p = .011$, $\eta^2_p = .155$. Using this estimate we analyzed simple main effects by looking at mean gaze duration

[2] Breaking down the significant 3-way interaction of Study 1 further led to a significant 2-way interaction between Context and Category, $F(1,39) = 42.12$, $p = .0001$, $\eta^2_p = .519$ not germane to our analysis. The other two 2-way interactions were not significant, Means by Category, $F < 1$, and Means by Context, $F(1,39) = 2.84$, $p = .1$, $\eta^2_p = .068$. In addition, we found a significant main effect of Category, $F(1,39) = 29.15$, $p = .0001$, $\eta^2_p = .428$ of Context, $F(1,39) = 48.74$, $p = .0001$, $\eta^2_p = .540$, while the main effect of Means was not significant, $F < 1$.

for functional means relative to control means for both levels of the social category. Within the criminal category, we compared trials in which a safe versus an unsafe context prime preceded the social category prime. The results of this analysis showed that when a safe context preceded the social category prime, gaze duration for means of self-defense was less (M = -1.48, SD = 5.63) than when an unsafe context was presented before priming the word criminal (M = 1.06, SD = 3.84, range = -20 - 20), $F(1,40) = 5.03$, $p = .031$, $\eta^2_p = .112$ (see Fig. 2). The same analysis was repeated for the conditions in which the control social category (athlete) was preceded by safe or unsafe contextual pictures. These results showed no difference with respect to mean gaze duration for self-defense means versus control means in a safe (M = 0.32, SD = 8.45) compared to an unsafe context (M = -0.92, SD = 7.04, range = -19 - 15), $F(1, 40) = 1.37$, $p = .250$, $\eta^2_p = .034$, (See Fig. 2). Means and standard deviations are given in Table 1[3].

The results of the first study were in line with our hypothesis that people attend more to means functional to a response behavior afforded by the context in which a social category member is perceived. When the category member is perceived in a context he is normally not perceived in, no difference was found in mean attention between functional and control means. This result therefore surpasses an effect based on the stereotypic association of the presented means with the perceived social category in which attention to these means would be equal across contexts.

Discussion

Even though the results are in line with our predictions, potentially, the use of negatively valenced stimuli could have differently impacted on attention. One option is that approach or avoidance reactions towards negative stimuli (e.g., gun, knife) or the valence of the pictures in the unsafe condition have affected attention above the predicted category-response link. This is in line with the idea that people attend more to negative than positive information or stimuli even when controlling for salience effects (Fiske, 1980; Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Similarly, the ‘weapon focus’ effect demonstrates how people more strongly attend to weapons versus control, non-weapon objects even when the context in which these objects are perceived is the same (Loftus, Loftus, & Messo, 1987). To rule out this alternative explanation we sought to replicate our general effect in the second

[3] The means represent the number of frames and since the frame rate was 60 Hz the difference in time (ms) can be calculated by multiplying the mean difference scores by 1 sec / 60.

study by using neutral instead of negatively valenced context and category stimuli.

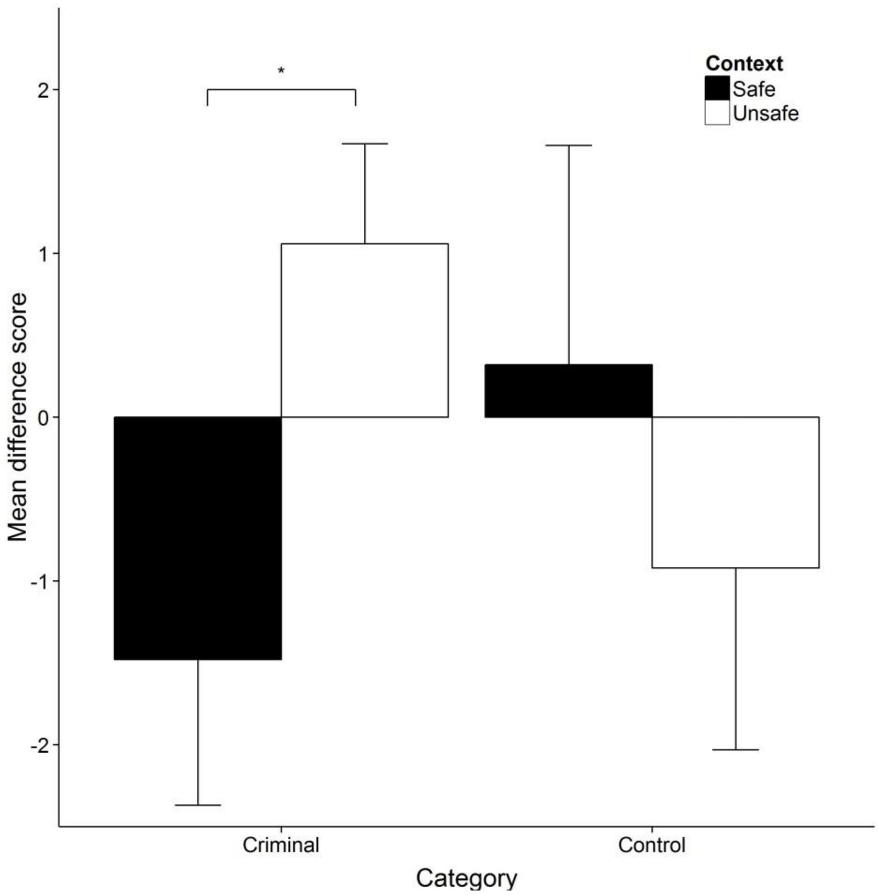


Figure 2. Mean attention for self-defense vs. control means in safe and unsafe contexts for both social category primes.

	Unsafe context	*Safe context*
Category		
Criminal	1.06 (3.84)	-1.48 (5.63)
Athlete	-0.92 (7.04)	0.32 (8.45)

Table 1. Mean difference score values for both social category primes in both contexts (Standard deviations are given within brackets.)

Furthermore, in the second study we departed from using difference scores as used in the first study. These scores provided a way of giving a mean estimate of attention for functional means in different contexts while taking attention for control means into account, too. However, one potential problem with difference scores is that based on relative differences in mean attention you ignore direct differences between attention for functional and control means in each context separately. To correct this in the second study, here we used a simpler design and analyzed the difference in attention between functional and control means within different levels of the context factor. This would further support our hypothesis that people attend more to functional means if a social category member is perceived in a context he or she is typically perceived in.

Study 2

In the second study we used a within-subjects design with a context picture (here: stadium vs. store) and athlete as social category word prime followed by pictures of potential means for cheering (the dominant response) combined with control means. Our hypothesis was that an athlete perceived in a stadium context would lead to more attention to be directed to means functional for cheering as behavioral response relative to control means (H1). This difference in attention was not expected when an athlete was perceived in a store context. Alternatively, if perceiving an athlete would activate cheering as behavioral response irrespective of the context, no difference in mean attention was expected between functional and control means (H2). As in the first study we measured attention holding using gaze tracking.

Participants

In total, 35 participants (of which 26 females) took part in the experiment in turn for either course credit or monetary compensation (7 Euros). The mean age was 23 years (range = 18-37).

Materials

The visual attention task in the second study was for a large part the same as the one used in the first study. Similarly, participants were told that they would participate in a task looking at object perception. As stimuli, here we used a (sports) stadium and a store as context stimuli with "SPORTER" [athlete] as social category word prime. The picture set of functional means of cheering included items typically used at sporting events (e.g., horn, flag). Both the context and pictures of means were randomly taken from a set of four different pictures of the same type. The set of control stimuli included pictures of means such as a plastic bag and a pair of coins. The position of functional and control means was counterbalanced so that they would appear an equal amount of times on either side of the screen. The arrow-indication task was the same as the one used in the first study. Also, the same gaze tracking materials and configurations were used in the second study. The visual attention task included 64 trials.

Results

No main effects of sex and age were found for mean attention across conditions, Fs<1. Similar to the first study, results from the exit-questionnaire indicated that participants did not seem to be aware of the purpose of the task. For the main analysis we performed a repeated measures analysis with mean attention as DV as a function of context (stadium vs. store) and mean type (functional vs. control). Based on the analysis we found no main effects for context or category, Fs<1. We did however find an interaction between context and mean type, $F(1,34) = 13.17$, $p = .001$, $\eta^2_p = .279$. Breaking down the interaction, we analyzed simple main effects separate for both contexts (stadium vs store). The results showed, in accordance with our main hypothesis (H1), that when an athlete was presented in a stadium context participants paid more attention to means of cheering (M = 304.38 ms, SD = 126.89) relative to control means (M = 263.69 ms, SD = 132.87), $F(1,34) = 9.36$, $p = 0.004$, $\eta^2_p = 0.216$.

Similarly in line with our main hypothesis we found that when a control context (store) was presented with the same categorical prime, there was no difference in mean attention for response means for cheering (*M* = 283.13 ms, *SD* = 117.66) relative to control means (*M* = 301.55 ms, *S*D= 119.88), $F(1,34) = 1.86$, $p = 0.182$, $\eta^2_p = 0.05$ (see Fig. 3).

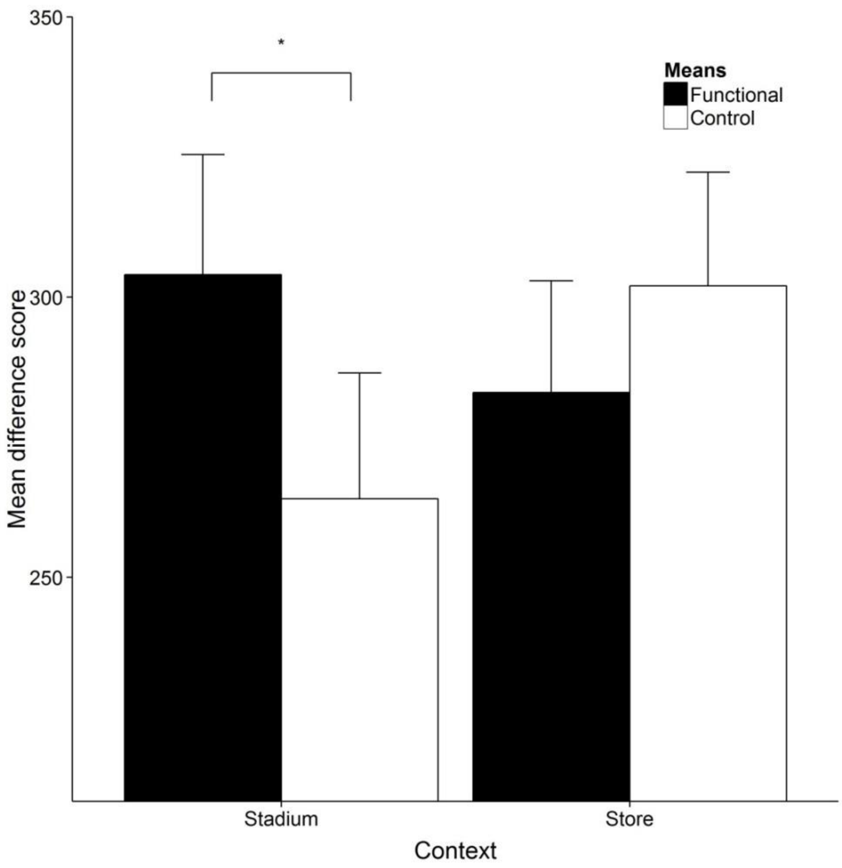


Figure 3. Mean attention for functional and control means when perceiving the social category "athlete" in both context conditions (stadium vs store).

Discussion

The results of the second study further support our general hypothesis. When category members are perceived in a context they are typically perceived in people attend more to means functional to perform associated response behaviors. Here, we found that people pay more attention to means of cheering when perceiving an athlete in a sports stadium context. This preference for functional means over controls was not found when the same social category member was perceived in the control context. Most importantly, changing the type of stimuli from negatively valenced to neutral did not alter the attention effect. Furthermore, the use of absolute measures has made it possible to better assess differences between attention for response means and control means. This gives us even more support for the predicted attention effect.

General Discussion

Overall, the results taken from our two studies fully supported our hypotheses. We found that the activation of response behaviors associated with social category members can even affect attention for means functional to perform these actions. In the first study we found that people attended more to means functional to response behaviors that are determined by the social context in which social category members were perceived. In Study 2, we replicated the effect with neutral valenced stimuli as well as using a different analysis that directly compared attention for functional versus control means in both levels of the context manipulation. This implies that the attention effect is found irrespective of the type of context displayed. Since the participants were not aware of the true purpose of the study, this suggests that the response selection process unconsciously affected attention for the presented means.

The results further support the notion that cognition, or cognitive representations are not stable but dependent on the direct context they are used in (Clark, 1997; Smith & Semin, 2007). Studies on stereotype representations (Blair, 2002) as well as behavioral tendencies (Cesario et al., 2010) have shown how the context can constrain or provide more opportunities to perform different kinds of behavior. In this way the context provides a framework in which others and their action ecology are perceived. Since others are always perceived in a visual context, attention for means in this context can play an important role when approaching others.

One may object that we did not present a naturalistic scene as in most studies using eye- or gaze tracking for scene perception (Duchowski, 2002; Henderson, 2007; Rayner, Smith, Malcolm, & Henderson, 2009), but chunked the relevant information parts and presented them sequentially. We did so for a reason. As we are interested in attention towards functional means following both a context and a social category, the stimulus material is naturally richer. Yet, by presenting all information (context, social category and means) together, attention for means would have to compete against attention for all other information. Of course it is interesting to test which elements of a scene perception grab most attention, but then the social category would win, most likely, followed by the context and finally by functional means. Such a paradigm would have not allowed to obtain the robust differences on the selective attention for functional means that we found. Therefore, we did not use pictures in which all variables, here context, category and means, were implemented but manipulated the context and category separately in order to show how the context could modulate the perception of others. More specifically, in line with Cesario et al. (2010) the context provides a framework in which someone is perceived. Moreover, this method made visual cues more easily discernable which in turn made it easier to assess the preference for preselected cues over others by controlling for distracting and non-relevant stimuli. This sequential method has also proven to be effective in research looking at the weapon bias where prime and target stimuli are presented in a sequence (Payne, 2001).

The absence of a measurement of response activation can be seen as a further possible limitation in this study. In previous research, this was explicitly measured using a lexical decision task (Jonas & Sassenberg, 2006; Cesario et al., 2006). We left this part out for two reasons. Besides making the task too complicated, response behaviors could direct attention to means that match this behavior irrespective of the preceding context and/or social category primes. Even if people would respond faster to response behaviors following appropriate context-social category combinations, the effect on attention could be determined by the response behavior words alone, while we sought to test the impact of the contextualized social category.

This study adds to the existing literature in that it provides first-time evidence that the combined context-social category compounds of information (Prinz, Aschersleben, & Koch, 2009) affect attention patterns preceding behavioral responses (Cesario et al., 2010; Correll et al., 2011). This response activation is linked to the attention for functional means, as determined by the appropriate response for the context-category compound of information. The results are in line with a top-down view on attention (Bruner, 1957; Henderson, 2007) in which perception is driven by expectancies, motives and goals. With respect to the current setting, as part of the attention selection process people perceive the whole event (Barsalou, 1999; Zwaan, 1999) in which means, functional to the actual responses, are part of. These means hold people's visual attention over other non-relevant means. Aarts, Dijksterhuis and de Vries (2001) demonstrated how thirsty participants better remembered drinking-related items they saw in an office during an incidental recall task. Our results take a step further in that more attention is devoted online to means because of their relevance for the upcoming response.

Overall, we have demonstrated here the importance of top-down perception on the processing of potential response means determined by context. Our novel results provide first evidence that means functional to potential responses will be attended to only when they are functional in the actual context. By making an analogy to real environments, our results suggest that the perception of the direct physical environment plays an important role in preparing subsequent actions by both constraining and increasing the number of means related response options. Further research should be done to see whether these findings hold for the perception of scenes in which multiple stimuli compete for attention and what factors determine the functionality of some of these stimuli over others.

Besides the complexities of social perception a fundamental link is missing between attention for functional means and the subsequent performance of associated actions. Even when accepting that attention is in part driven by action selection, the translation between attention and overt action performance requires a host of intermediate decision processes that are far beyond the current paradigm. Nonetheless, the link between processing contextual information and behavior is central to the current thesis. Therefore, in the next chapter we use a similar social perception task but use behavioral responses as a dependent measure rather than visual attention. Although this means translating the current paradigm to a different setting and therefore making it more difficult to compare findings, we tried to repeat the context manipulation in a different way in order to keep the same predictions in a behavioral task paradigm.

Chapter 3

Complementary Hand Responses Occur in Both Peri- and Extrapersonal Space

This chapter is based on: Faber, T. W., van Elk, M., & Jonas, K. J. (2016). Complementary Hand Responses Occur in Both Peri-and Extrapersonal Space. *PloS one*, *11*, e0154457. doi:http://dx.doi.org/10.1371/journal.pone.0154457

Abstract

Human beings have a strong tendency to imitate. Evidence from motor priming paradigms suggests that people automatically tend to imitate observed actions such as hand gestures by performing mirror-congruent movements (e.g., lifting one's right finger upon observing a left finger movement; from a mirror perspective). Many observed actions however, do not require mirror-congruent responses but afford complementary (fitting) responses instead (e.g., handing over a cup; shaking hands). Crucially, whereas mirror-congruent responses don't require physical interaction with another person, complementary actions often do. Given that most experiments studying motor priming have used stimuli devoid of contextual information, this space or interaction-dependency of complementary responses has not yet been assessed. To address this issue, we let participants perform a task in which they had to mirror or complement a hand gesture (fist or open hand) performed by an actor depicted either within or outside of reach. In three studies, we observed faster reaction times and less response errors for complementary relative to mirrored hand movements in response to open hand gestures (i.e., 'hand-shaking') irrespective of the perceived interpersonal distance of the actor. This complementary effect could not be accounted for by a low-level spatial cueing effect. These results demonstrate that humans have a strong and automatic tendency to respond by performing complementary actions. In addition, our findings underline the limitations of manipulations of space in modulating effects of motor priming and the perception of affordances.

Introduction

Imitation is a key characteristic of human beings (Tomasello, Kruger, & Ratner, 1993). Copying actions performed by others either intentionally or automatically is however not restricted to copying similar actions (moving your right finger when perceiving a right finger movement) but often requires complementing actions as well (e.g., shaking an extended hand; handing over a cup). While both mirror and complementary actions can be driven by the same social information (e.g., hand gestures), mirror actions do not involve direct interaction while complementary actions often do. In the present study we investigated whether the tendency to imitate or complement others' actions depends on the opportunity to directly interact with them. In a world in which we are increasingly surrounded by possibilities for virtual interactions (e.g., through virtual reality gaming, Skype, 3D movies etc.) insight in this topic is of high importance. One theoretical possibility is that mirror and complementary actions are a product of automatically detecting the potential for action (affordances) as dispositional properties of the observed person or object (Tomasello, Kruger, & Ratner, 1993). More specifically, information about a person or object properties such as knowing that an extended hand can be complemented with an opposite hand (i.e., handshake) might be sufficient to trigger a complementary response, irrespective of an individual's opportunity to actually perform this response. This reasoning fits well with evidence from research on object perception, which has shown how people make complementary actions towards objects even when motor properties associated with these objects are task-irrelevant and regardless of an object being physically manipulable or not (Parsons et al., 1995; Jeannerod & Frak, 1999; Jeannerod & Decety, 1995; Tucker & Ellis, 1998; Sartori, Cavallo, Bucchioni, & Castiello, 2011; Sartori, Cavallo, Bucchioni, & Castiello, 2012).

Recently however, neurophysiological and behavioral studies have demonstrated how perceived distance of objects in space modulates both observation and interaction with objects (Caggiano, Fogassi, Rizzolatti, Thier, & Casile, 2009; Costantini, Ambrosini, Tieri, Sinigaglia & Committeri, 2010; Valdés-Conroy, Sebastián, Hinojosa, Román, & Santaniello, 2014). Whereas this distance modulation has been shown for objects, it has not been shown for social actions (e.g., hand gestures), likely because these cues are primarily used to study imitation for which physical contact is not a prerequisite (Brass, Bekkering, Wholschläger, & Prinz, 2000). Social actions differ from objects given that they can trigger making either a congruent (from a mirror perspective) or incongruent (complementary) response such as imitating finger movements or shaking a person's hand. We note that this distinction might be confusing given that complementary actions are denoted here as incongruent in terms of mirror perspective but can at the same time be congruent in terms of outcome (i.e., handshaking as a social response outcome) or be anatomically congruent to the observed gesture (e.g., making a right open hand movement upon seeing an actor's right open hand movement). However, we decided to use this terminology for the sake of consistency with the discussed literature, for a more inclusive definition see Sartori and Betti (2015).

The question we pose is whether the perceived opportunity to interact with another person affects motor priming in response to social actions (i.e., making a fist or an open hand gesture). In five studies we test whether and how responding to hand gestures performed by an actor onscreen, by either making a congruent (mirror) or incongruent (complementary) response is modulated by the depicted distance in space between the actor and observer. Before describing the current research effort, we will summarize research on the effects of motor priming and the modulating role of peripersonal space on these effects.

Motor priming and space

A large number of studies have demonstrated how observing actions (goal-directed or 'meaningless') can activate certain brain areas that are also involved in the performance of these actions (i.e., the so-called mirror neuron regions; see Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti & Craighero, 2004; di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992). Besides neurophysiological and brain imaging evidence, much support for this visuomotor interaction where perceived actions trigger associated motor programs comes from behavioral response paradigms (Tucker & Ellis, 1998; Brass, et al., 2000; Brass, Bekkering, & Prinz, 2001). Importantly, a number of studies have shown how an observed action maps onto the representation of the identical or same action in the perceiver (i.e., direct matching). For example, people are quicker to perform a finger movement (upwards or downwards) in response to a symbolic cue when concurrently observing a congruent (mirror) finger movement relative to an incongruent movement (Brass et al., 2000; Stürmer, Aschersleben, & Prinz, 2000).

Besides examples of direct matching, recent findings have suggested how motor observation can also trigger more functional or complementary responses (Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007; Ocampo & Kritikos, 2010). For example in Sartori et al. (2012), participants showed automatic (covert) congruent responses when observing a hand grip towards an object but showed complementary, non-identical responses when observing grip postures that signaled a complementary request. One framework accounting for identical as well as non-identical motor responses to observed actions is the theory of associative sequence learning (ASL; Heyes, 2001; Brass & Heyes, 2005; Heyes, 2011). This theory proposes that motor priming is a product of domain general learning mechanisms that account for congruent as well as incongruent priming effects, given that both are a result of the same sensorimotor learning process. Some findings supportive of ASL have shown a facilitation for making incongruent relative to congruent gestures after only a short reverse training for both meaningless gestures (e.g., making an open hand gesture in response to a closing hand) as well as for object directed actions (Heyes, Bird, Johnson, & Haggard, 2005; van Schie, van Waterschoot, & Bekkering, 2008). Besides creating novel associations (Heyes et al., 2005; Catmur, Walsh & Heyes, 2007; Cook, Press, Dickinson, & Heyes, 2010) other studies have demonstrated how existing contingent stimulus-response associations may underlie motor priming. In Liepelt, Prinz, and Brass (2010) for example, participants were faster in making congruent (mirror) hand movements towards intransitive hand gestures (i.e., a fist) while for communicative gestures (i.e., an open hand) participants were faster in responding with incongruent relative to congruent hand movements signaling a complementary hand response (i.e., handshaking). These findings corroborate an earlier study by Flach, Press, Badets, and Heyes (2010) in which the same complementary effect was found for open hand gestures but not for arrows (see also Sartori et al., 2011).

Taken together it seems that differences between direct matching and more functional examples of motor priming are determined by the (minimal) social meaning of the cue and the task context. This fits with the idea that even though direct matching might suffice for low-level motor cues, in interactive settings matching behavior displayed by others is often suboptimal (Sartori, Betti, Chinellato, & Castiello, 2015). For example, when somebody throws you a ball or falls down the stairs a complementary response (e.g., catching; helping) is more fitting than copying the perceived behavior (e.g., throwing; falling; Jonas & Sassenberg, 2006). Diverse studies on joint action have underlined the importance of interactive settings in producing task-relevant complementary actions (Obhi, & Sebanz, 2011; Sebanz, Bekkering, & Knoblich, 2006; Schilbach, 2014; Richardson, Hoover, & Ghane, 2008). Still, the majority of motor priming tasks do not include information that signals the possibility to interact with social stimuli (real or imagined) and has commonly focused on mirror or complementary actions separately (at least for different types of gestures). One type of information that resolves this is (interpersonal) space, or the opportunity to interact with objects or others if they are within reach (although not all interactive settings require this such as throwing and catching a ball). By manipulating space in a motor priming task we hypothesize that the perceived opportunity to interact should not affect motor imitation, given that motor imitation does not require direct (physical) interaction, but should affect complementary responses to hand gestures (e.g., open hand) which can only be performed when the other person is within direct reach.

Early support for the role of space in motor processing comes from Rizzolatti, Fadiga, Fogassi, and Gallese (1997) who stated that visual input is discriminated in terms of manipulability as a function of the space around the body within (peripersonal space) or outside of reach (extrapersonal space). For example, Caggiano et al. (2009) found that in monkeys some mirror neurons selectively responded to the observation of object-directed actions performed by an actor in either the peri- or extrapersonal space of the monkey. These space-selective neurons therefore seemed to dissociate between object-directed actions that could be performed immediately (when the object was in the monkey's peripersonal space) and actions that could not be performed (i.e., when the object was outside of peripersonal space). This seems to suggest that the space around the body is defined by the function it provides for an individual to perceive and manipulate objects that lie within this space (Tversky, 2005).

Research using behavioral paradigms has similarly demonstrated how motor priming is modulated by peripersonal space (Costantini et al., 2010; Cesario, Plaks, Hagiwara, Navarette, & Higgins, 2010). For example in Costantini et al. (2010), participants were faster to pantomime a correct relative to incorrect grip response towards an object (mug) but only when the object was presented in peri- and not in extrapersonal space. Importantly, the effect of distance was replicated when instead of manipulating distance through onscreen pictures, objects were placed in a physically reachable location (Ambrosini & Costantini, 2013). Here, participants were faster to pantomime congruent object-directed actions only when the mug was positioned within physical reaching distance and not when outside of reach.

Taken together, motor system activity in response to perceived objects (social or non-social) seems to be affected by the (peripersonal) space or the perceived opportunity to interact with them. For our current study we chose a simplified behavioral response paradigm to test whether social cues selectively facilitate responding as a function of the perceived opportunity to interact with the other person. We first propose that congruent (mirror) motor responses to intransitive gestures do not require interpersonal contact whereas complementary responses to open hand gestures do. Therefore, presenting gestures in close distance (i.e., within peripersonal space) should only affect motor priming for open hand and not to intransitive gestures. The gesture types and design are primarily based on Liepelt et al. (2010) in that we presented both fist and open hand gestures and looked at the response time between perceiving the gesture and performing a congruent or incongruent hand gesture. We expected that participants would show faster incongruent (complementary) responses to an open hand in close distance relative to performing a congruent response. When the same open hand was depicted in far distance this complementary advantage was predicted to disappear.

Study 1

In line with Liepelt et al. (2010) we instructed participants to mirror hand gestures made by a visual actor onscreen either with the same specular hand (congruent response; actor's right hand – participant's left hand) or using the same anatomical hand as the hand displayed (incongruent response; actor's right hand – participant's right hand). The actor either made a fist or an open hand gesture with his right or left hand. Similar to the handshaking effect found in Flach et al. (2010) and Liepelt et al. (2010) we predicted a complementary effect (faster incongruent relative to congruent responses) for open hand trials but not for trials involving a fist as hand gesture. To investigate whether the complementary effect would change as a function of space we showed the visual co-actor sitting either opposite on the short end (close) of a table or on the long end (far) of a table. With respect to fist gestures we predicted, in line with Liepelt et al. (2010), an advantage for performing congruent relative to incongruent hand responses irrespective of interpersonal distance whereas for open hand gestures we predicted that participants would be faster at making incongruent responses compared to congruent responses only in close distance.

To control for our experimental distance manipulation, we asked participants to rate the reachability of the actor in the task in both close and far distance trials (see Valdés-Conroy et al., 2014). After the task participants indicated whether it would be possible to touch the hand of the actor sitting across from the short and long end of the table (separately), if they would be positioned opposite to the actor.

Participants

Forty one participants (mean age = 22; range = 18-34), including 34 females, participated in the experiment in exchange for either course credit or monetary compensation. In total, 36 participants were right-handed as assessed through self-report. All participants signed an informed consent form before participating in the study. The study was approved by the Psychology Department of University of Amsterdam ethics committee (2014-SP-3731).

Materials

The stimulus material in the study consisted of photographs made in a lab room with an actor sitting behind a table. All possible hand gestures were photographed so that the actor was shown sitting either behind the short or far end of the table. The exit-questionnaire was used to assess if touching hands with the actor sitting across either side of the table was perceived to be possible. In addition, a number of demographic questions were included.

Procedure and design

The experiment used a 2 x 2 x 2 design with Distance (close vs far), Hand gesture (open hand vs fist) and Hand response (congruent vs incongruent) as within-subject factors. Participants took part in the experiment as the last of a series of three studies. They were seated at a table with a response box placed in front of them at the center of the table. Behind the response box a computer screen was positioned (resolution: 1680 x 1050 pixels; 22-inch diagonal; 60 Hz refresh rate) on which the visual stimuli were displayed. The custom modified response box (Psychology Research Tools Inc., 2012) had three horizontally aligned response keys, and the most right and left key was used as functional response key, respectively, in the experiment.

Participants were instructed to mirror the specific hand gesture (fist vs open hand) displayed by an actor onscreen, by responding with the same specular hand as the hand appearing onscreen (i.e., congruent response). In half of the trials the hand was given a color, which meant participants had to mirror the hand gesture but then using their hand opposite to the displayed hand on the screen (same anatomical hand; incongruent response).

After the instruction participants received an additional vocal instruction to make sure the task was clear. Each trial in the main task (see Fig 1) started with a message displayed in the center of the screen for 500 milliseconds (ms) instructing participants to hold the left and right key of the response box pressed with their left and right index finger. When the message disappeared a picture was presented at the center of the screen (1400 x 798 px; consistent with a 32° horizontal visual angle and a viewing distance of 70 cm) depicting a man seated at a table with his hands alongside his body under the table. The man was displayed from the lower part of the neck down, either sitting across the short side (width) of a table (from the perspective of the participant) or the long end of a table for 1000 ms. Following this rest picture a second picture (gesture) was shown in which the actor had lifted his right or left arm and made either a fist or an open hand gesture while keeping the other hand/ arm at resting position. The gesture picture always displayed the same distance position (i.e., short or long end of the table) as the rest picture. In half of the trials a non-colored hand was shown indicating that a congruent response had to be made (see Fig 2). In the other half of the trials the hand was colored green (medium opacity) instructing participants to perform an incongruent hand gesture, mirroring the observed hand gesture using the opposite hand. Response time was calculated from the onset of the gesture picture until participants released the key on the response box in order to perform the hand gesture. In total, 128 trials were presented, in which the presentation of specific trials was fully randomized with respect to distance, hand gesture and hand response. Before the main task participants went through a training session for 28 trials, randomized similar to the main task, to make sure participants were sufficiently acquainted with the task rules. Individual performance was monitored shortly during the training session by the experimenter. During the onscreen instruction it was only shortly mentioned that the depicted actor was sitting at either the short or long end of the table. No specific instructions were provided with regard to the relevance of the spatial distance for the hand response. When finished, participants filled in an exit-questionnaire and were debriefed.

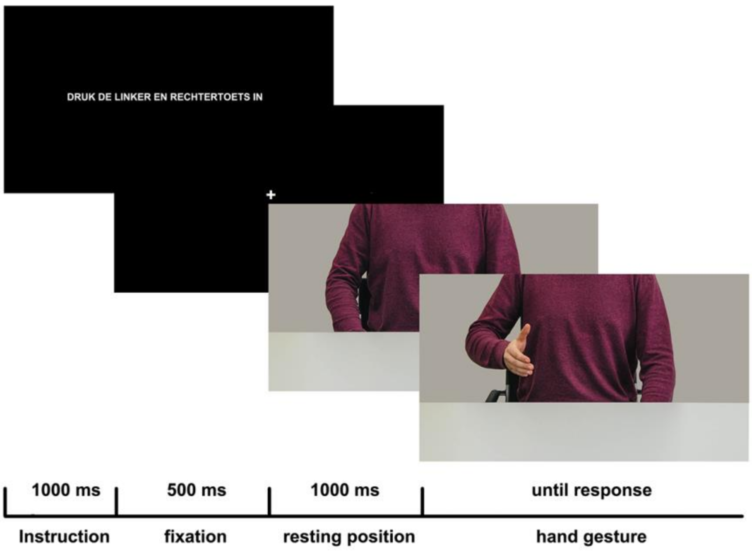


Fig 1. Experimental setup. The general experimental design for the first four studies. The first screen includes the text: “Press and keep pressing the right and left key”. Only non-colored hands are shown here

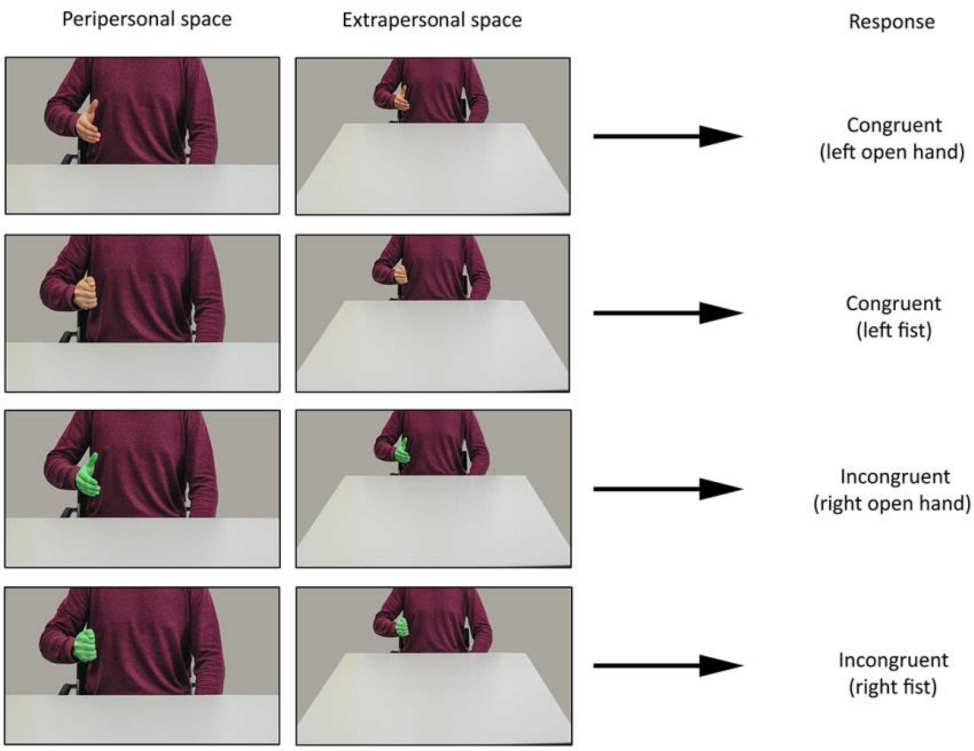


Fig 2. Full design. Full design for the first study. Note: Hands were presented on either side of the screen (left and right), here only hands on the left side are shown. Also, the bottom two rows display the hand gestures including a green (medium opacity) colored layer on top that makes the hand appear green but still sufficiently visible in terms of texture

The color cue instructing participants to mirror the perceived hand gesture using their opposite (incongruent) hand was not counterbalanced across participants or conditions (i.e., responding to a color cue with alternating congruent and incongruent hand responses). Although this decision might have affected the results this would have only strengthened rather than weakened the main findings. Given that incongruent responses to open hand or fist gestures require a secondary step (mirror the gesture – use non-corresponding hand) this would have potentially slowed reaction times rather than made it faster (see van Schie et al., 2008).

Data analysis

Response errors were first deleted from the task but used separately to analyze differences in response errors between different trial types. For all trials (for participants combined), 6 %

(305 trials) of response trials with a wrong response (a mirror response to a green colored hand and a complementary response to a normal colored hand) were deleted. In addition we excluded 1.9 % (99 trials) of the total sum of trials in which response times were above or below 2.5 x *SD* from the mean of each individual participant.

Exit-Questionnaire

In the exit-questionnaire participants were asked whether they could imagine making physical contact (touching hands) with the actor opposite of the table either in the short or far distance setting in order to assess the perceived opportunity to interact. One participant did not complete the exit-questionnaire and was removed before analyzing the data. For the remaining participants, in total 45 % (18) imagined it was possible to touch the hand of the person in the short setting while for the far setting only 12.5 % (5) imagined this to be possible. Besides ratings of reachability the exit-questionnaire included demographic questions (sex and age) and a question concerning handedness ("Are you left or right-handed?").

Results

To analyze response time data we ran a 2 x 2 x 2 repeated measures analysis with Distance (close vs far), Hand gesture (fist vs open hand) and Hand response (congruent vs incongruent) as within-subject factors (for RT and error data, see Fig 3). The analysis yielded a non-significant main effect for Distance as well as a non-significant main effect for both Hand gesture and Hand response (all *F*s < 1). Also, no interaction between Distance and Hand gesture was found ($F < 1$). The interaction between Distance and Hand response did not reach statistical significance either, $F(1, 39) = 3.75$, $p = .060$, $\eta^2_p = .09$. We did however, find a significant interaction between Hand gesture and Hand response, $F(1, 39) = 16.78$, $p < .001$, $\eta^2_p = .30$. Breaking down this interaction yielded a non-significant difference in response latencies between congruent and incongruent responses to fist gesture trials, $t(39) = -1.41$, $p = .167$. Response latencies were faster for incongruent ($M = 706$, $SD = 117$) compared to congruent ($M = 738$, $SD = 140$) responses to open hand trials irrespective of distance, $t(39) = 2.63$, $p = .012$, $d = 0.48$ (d_z throughout). The three-way interaction between Distance, Hand gesture and Hand response was not significant ($F(1, 39) = 0.02$, $p = .691$), contradicting our main hypothesis.

One option is that the specific features of the cue (e.g., type of gesture; type of distance) require different processing times so that the potential modulation of distance on the complementary effect is found only in slower responses rather than faster. To investigate this, we segmented response time latencies in two time-bins focusing separately on the fastest and slowest responses. Running a repeated measures analysis with Distance, Hand Gesture and Hand response revealed a Hand gesture by Hand response interaction for the fastest responses (time-bin 1), $F(1, 39) = 16.74$, $p<.001$, $\eta^2_p = .30$, as well as for the slowest responses (time-bin 2), $F(1, 39) = 14.20$, $p = .001$, $\eta^2_p = .27$. For both time-bins no interaction with Distance, Hand gesture and Hand response was found (both *F*s < 1).

For the response errors (e.g., a congruent response to a green colored hand) we performed the same 2 x 2 x 2 repeated measures analysis as for the RT data based on the number of response errors per trial type. We found no main effects for Distance and Hand gesture (both *F*s < 1) but we did find a main effect for Hand response, $F(1, 39) = 7.96$, $p = .007$, $\eta^2_p = .17$. Also, a two-way interaction was found between Distance and Hand response, $F(1, 39) = 4.12$, $p = .049$, $\eta^2_p = .10$ as well as between Hand gesture and Hand response, $F(1, 39) = 11.83$, $p = .001$, $\eta^2_p = .23$. However, these effects were all qualified by a three-way interaction between Distance, Hand gesture and Hand response, $F(1, 39) = 7.75$, $p = .008$, $\eta^2_p = .17$. We performed follow-up tests for the three-way interaction (corrected for multiple comparisons using a Bonferroni correction). A higher error rate was found for congruent ($M = 1.48$, $SD = 1.34$) relative to incongruent responses ($M = 0.53$, $SD = 0.88$) in open hand trials, $t(39) = 4.05$, $p < .001$, $d = 0.64$, whereas this difference was not found for fist trials ($p = .068$). In far distance trials we found the same pattern as indicated by a higher error rate for congruent ($M = 1.23$, $SD = 1.21$) compared to incongruent responses ($M = 0.45$, $SD = 0.96$) in open hand trials, $t(39) = 3.54$, $p = .001$, $d = 0.57$, while no difference was found for fist trials.

Discussion

In line with Liepelt et al. (2010) we found a complementary effect (i.e., faster incongruent relative to congruent responses) for open hand gestures. Contrary to our hypothesis the results indicate that this effect is found irrespective of spatial distance. Furthermore, the error pattern suggests that the tendency to make a complementary response to open hand gestures, as indicative of a higher error rate for congruent responses to open hand gestures, was evident in both close and far distance trials. Furthermore, no congruency effect (faster responses in congruent compared to incongruent trials) was found for fist gestures contrary to Liepelt et al.

(2010). We discuss further issues with respect to this finding as well as the unexpected variability of reachability ratings in the re-analysis paragraph following Study 5.

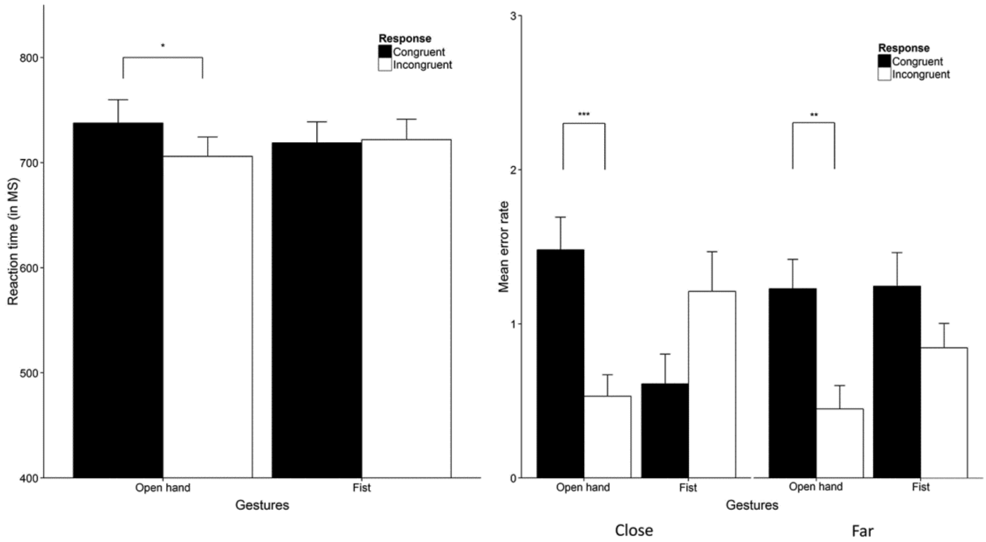


Fig 3. Response data Study 1. The RT results (Left) show reaction time in milliseconds averaged over distance (close and far) trial. The right graph shows error responses for both close (left) and far distance (right) trials. Error bars are 1 x SE. * = $p<.05$, ** = $p<.01$, *** = $p<.001$. All figures were created using ggplot2 (Wickham, 2009)

One possibility for the absence of an effect of distance is that the close versus far distance images were not well designed to convey visual differences in peripersonal space given that both close and far distance images literally appeared at an equal physical distance from the participant. Also, perhaps the smaller spatial distance between the gesture position and the center of the screen in the far distance trials may have differentially affected hand responses. To control for visual differences (of the actor and response hand) and to replicate the task using a stronger manipulation of peripersonal space we ran a second study in which we used the same stimuli but with a glass screen placed on the table in front of the actor, providing a measure of functional rather than physical space (see Costantini et al., 2010). Given that peripersonal space requires that people or objects are able to directly interact, the placement of a screen would obstruct such an interaction. Somewhat in line with Heed, Habets, Sebanz, and Knoblich (2010) if an actor is seated in peripersonal space but not engaged in all aspects of the task this is not sufficient to foster interpersonal coordination. Therefore, if an actor is seated in peripersonal space but is unable to interact with the participant this should interfere with producing complementary motor responses.

Study 2

Study 2 only differed from Study 1 in terms of the stimulus set. As in the previous study there was a 2 x 2 x 2 within subjects design but here with Frame (glass frame vs no frame), Hand gesture (open hand vs fist) and Hand response (congruent vs incongruent) as within-subject factors. We used the same stimuli from the previous study depicting hand movements from the close distance trials as no frame trials. In the glass frame trials a glass frame was shown on the table in front of the actor onscreen (see Fig 4). This was done for both the gesture and the rest trials.

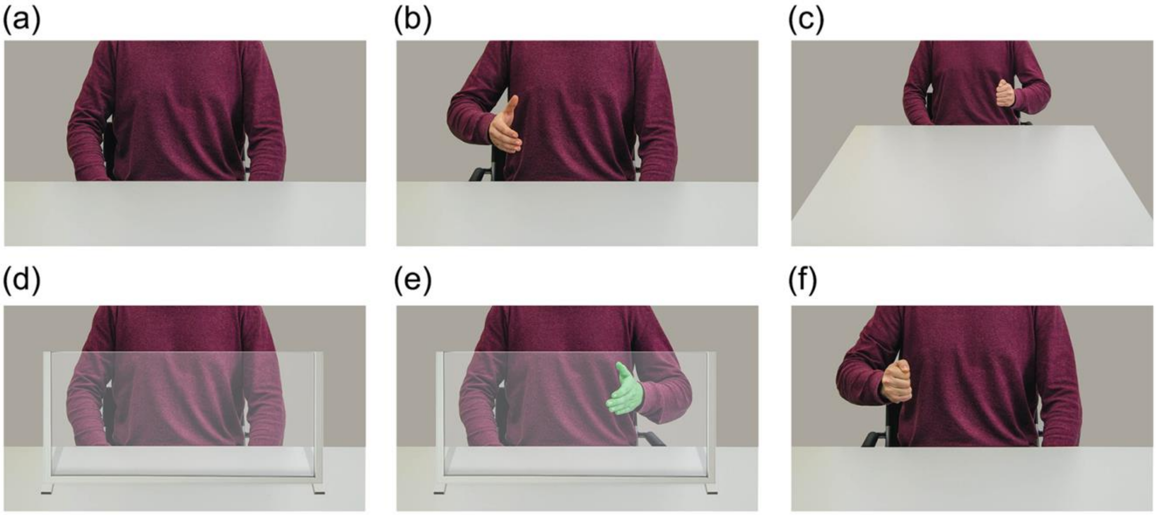


Fig 4. Sample images from Study 1, 2 and 4. (a) Resting position close (b) Open hand left close (c) Fist right far (d) Resting position frame (e) Open hand green right frame (f) Fist left close

Participants

In this study, 30 participants took part including 21 Females with a mean age of 23 (range 18-30). Twenty-seven participants were right-handed. All participants took part in the study in exchange for course credit or monetary compensation. All participants signed an informed consent form before participating in the study. The study was approved by the Psychology Department of the University of Amsterdam ethics committee (2014-SP-3731).

Materials and Procedure

For the stimuli, we took the no frame trials and edited, using Photoshop (version CS6; Adobe Systems, 2012) a glass frame on the table in front of the actor. The specific glass frame was chosen from a pretest in order to select a screen that most people would rate as clearly obstructing interpersonal contact. The instructions, setup and task presentation was identical to the first experiment. In total 128 trials were presented additional to a practice run including 28 trials, both fully randomized with respect to frame, hand gesture and hand response.

Data analysis

In the total number of trials, 6.9 % (266) incorrect responses were made which were removed before performing the analysis. Furthermore, we excluded 2 % (76) of response trials where the reaction time was above or below 2.5 x *SD* each participant's mean. Additionally, we excluded response data from one participant who was not a native speaker and had trouble understanding the task instructions.

Exit-questionnaire

As in the previous study, we asked participants to indicate whether they thought it was possible to make physical contact with the depicted actor on the screen in the no frame and glass frame trials. Of all participants, 51.7 % (15) indicated that they thought it would be possible to make physical contact in the no frame trials while all participants thought this would not be possible in the glass frame trials.

Results

For the main analysis we ran a 2 x 2 x 2 repeated measures analysis with Frame (glass frame vs no frame), Hand gesture (open hand vs fist) and Hand response (congruent vs incongruent) as within-subject factors (for RT and error data see Fig 5). First, the analysis showed no significant main effects for Frame, Hand gesture or Hand response (all *F*s < 1). Also, no hypothesized three-way interaction effect was found between Frame, Hand gesture and Hand response ($F(1, 28) = 1.13$, $p = .297$) nor a significant interaction between Frame and Hand response ($F < 1$). We did find an interaction between Frame and Hand gesture, $F(1, 28) = 6.88$, $p = .014$, $\eta^2_p = .20$ as well as an interaction between Hand gesture and Hand response, $F(1, 28) = 17.65$, $p < .001$, $\eta^2_p = .39$. For the Frame by Hand gesture interaction, we found that participants responded faster to fist gestures in the no frame trials ($M = 676$, $SD = 108$)

relative to the glass frame trials (*M* = 691, *SD* = 115) irrespective of the type of hand response made, *t*(28) = -2.51, *p* = .018, *d* = -0.47. No difference was found in response times for open hand trials averaged over hand response type between glass frame and no frame trials. To obtain simple main effects for the Hand gesture by Hand response interaction, we looked at mean response times averaged over levels of Frame. Results showed that participants made faster incongruent responses to an open hand gesture (*M* = 656, *SD* = 92) compared to congruent responses (*M* = 693, *SD* = 131), *t*(28) = 2.45, *p* = .021, *d* = 0.45. For trials where the actor showed a fist gesture, there was no significant difference in response times between incongruent and congruent hand responses, *t*(28) = -1.73, *p* = .095.

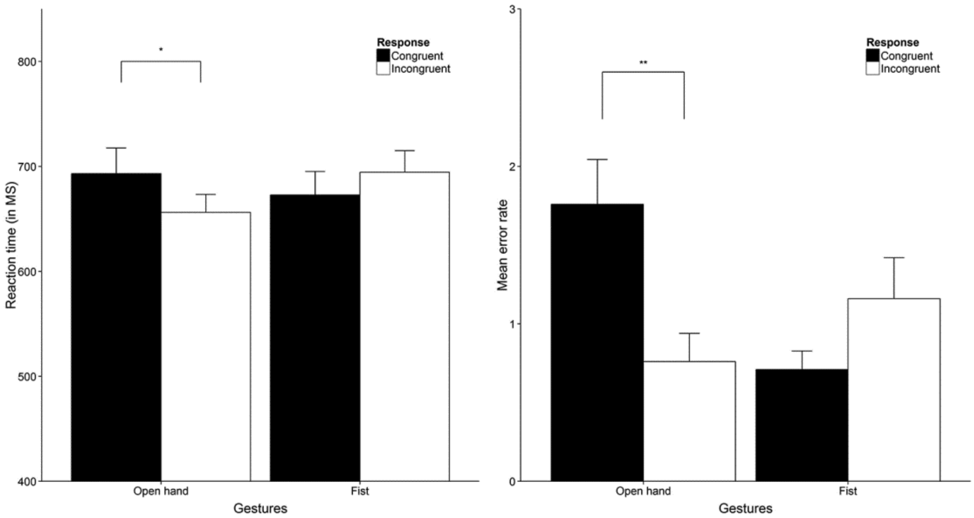


Fig 5. Response data Study 2. The RT results (Left) show reaction time in milliseconds averaged over distance (close and far) trials. The right graph shows mean error responses averaged over distance trials (close and far). Error bars are 1 x SE. * = *p*<.05, ** = *p*<.01, *** = *p*<.001

For the response error data we found a main effect for Hand gesture, *F*(1, 28) = 5.05, *p* = .033, η^2_p = 0.15 which was qualified by a significant interaction between Hand gesture and Hand response, *F*(1, 28) = 15.25, *p* = .001, η^2_p = .35. Simple main effects indicated that more errors were made in congruent (*M* = 1.76, *SD* = 1.53) compared to incongruent trials (*M* = 0.76, *SD* = 0.97) in response to open hand gestures irrespective of the frame being present or not, *t*(28) = 2.92, *p* = .007, *d* = 0.54. No difference was found in the errors made in response to trials including a fist as hand gesture. No effects of distance on the error rates were observed.

Discussion

The results were in line with the first study in that incongruent responses to open hand gestures were faster than congruent responses. Moreover, the findings did not show an effect of space, here manipulated with a frame positioned in front of the actor. The results further support the idea that peripersonal space, by either increasing distance or obstructing physical contact, does not affect the performance of complementary hand gestures.

Furthermore, in line with Study 1, the response error data suggest that participants were more inclined to perform incongruent responses in response to open hand trials when congruent responses were required. Although the results provide additional support for the space independence of the complementary effect it might still be the case that the type of stimulus environment does not accurately reflect the setting it is designed to imply (i.e., participants always responded to a 2D stimulus with which no 'real' interaction was possible in the first place, thereby rendering our space manipulation mute). Increasing the ecological validity by increasing the realism of the stimuli might be one way to improve the current design.

Study 3

As suggested, it could be that the type of images used so far was not fit to convey differences in perceived physical or functional space. Therefore, in the third study we used the exact same task as in the first study but then presented the images on a pair of 3D glasses to enhance the experience of interpersonal space. A similar manipulation was used in Costantini, Ambrosini, Scorolli, and Borghi (2011) who let participants categorize words that specified motor responses to objects that were presented at different points in space while wearing 3D glasses.

Participants

Twenty-one participants took part in the third study including 13 Females with a mean age of 22 years old (range 19-27). In total, 19 out of 21 participants were right-handed. All participants received either course credit or monetary compensation and signed an informed consent form before participating in the study. The study was approved by the Psychology Department of the University of Amsterdam ethics committee (2014-SP-3880).

Materials and procedure

A new set of stereoscopic stimuli was created that matched the stimulus set used in the first study. These pictures were presented using an Oculus Rift head-mounted display (version DKII; Oculus VR, 2014). The same response box and procedure was used as in the previous two studies. Additionally, we added to the onscreen instruction that the actor seated across from the short end of the table was within reach whereas the actor seated across from the far side of the table was outside of reach. This was done in order to stress the differences in interpersonal distance in the images. Also, for the practice part participants went through 20 practice trials without the 3D device followed by 20 trials with the device. The main task included 128 trials, randomized with respect to distance, hand gesture and hand response. Finally, we removed the onscreen instruction in the task that indicated to press and keep pressing the right and left key at the start of each trial. The reason to do this was that accommodating to a stereoscopic display of instructions was quite unpleasant for the eyes (Shibata, Kim, Hoffman, & Banks, 2011). Based on the response error data we saw that this change did not increase the total number of response errors (3.3 % versus 6 % in Study 1 and 6.9 % in Study 2).

Data analysis

For all recorded trials, 3.3 % of the trials (89) were removed due to response errors. Also, 2.5 % of the trials (55) were removed for which the response latency was above or below 2.5 x *SD* the mean of the participant.

Exit-questionnaire

In line with the earlier studies we found that still 42.9 % (9) of the participants thought it would be impossible to make physical contact with the depicted actor in close distance. All participants thought it would be impossible to do this in the far distance trials.

Results

For the main analysis we conducted a 2 x 2 x 2 repeated measures analysis with Distance (close vs far), Hand gesture (fist vs open hand) and Hand response (congruent vs incongruent) as within-subject factors (see Fig 6). We only found a significant Distance by Hand gesture interaction irrespective of the type of hand response made, $F(1, 20) = 5.14$, $p = .035$, $\eta^2_p = .20$. All other main effects and interaction effects including the three-way interaction between

Distance, Hand gesture and Hand response yielded non-significant results. Looking at the mean response times per level of hand gesture we found that responding to open hand gestures, irrespective of the type of response, was faster in close (M = 676, SD = 97) relative to far distance trials (M = 698, SD = 121), $t(20) = -2.30$, $p = .032$, $d = -0.50$. No difference was found for trials where a fist gesture was shown.

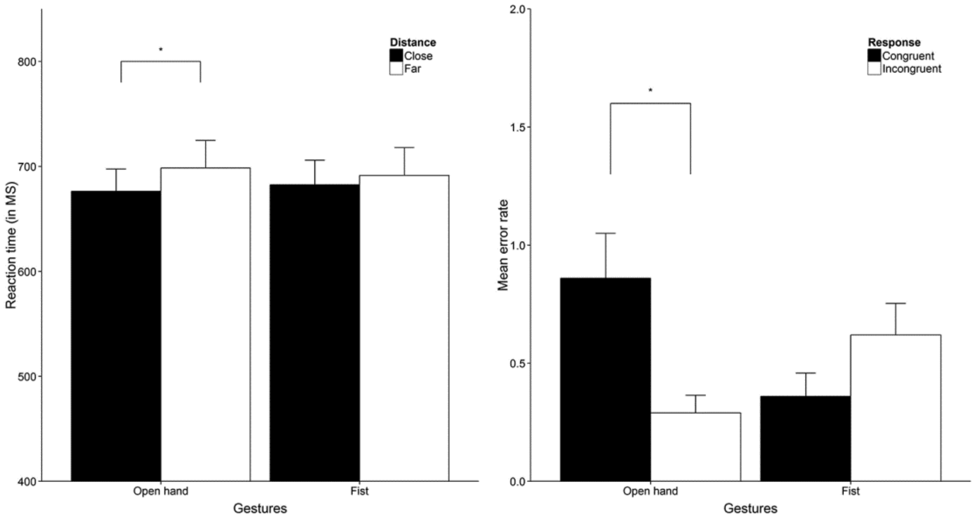


Fig 6. Response data Study 3. The RT results (Left) show reaction time in milliseconds averaged over Hand response (congruent and incongruent). The right graph shows mean error responses averaged over distance (close and far). Error bars are 1 x SE. * = p<.05, ** = p<.01, *** = p<.001

Results from the response error data showed a significant Hand gesture by Hand response interaction, $F(1, 20) = 9.83$, $p = .005$, $\eta^2_p = .33$ while no other main or interaction effects reached statistical significance. Simple main effects indicated that more errors were made in congruent (M = 0.86, SD = 0.87) relative to incongruent response trials (M = 0.29, SD = 0.34) for open hand gestures irrespective of distance, $t(20) = 2.61$, $p = .017$, $d = 0.57$. No difference was found in the error rate for fist trials and no effects of distance were observed.

Discussion

The results of the third study showed an effect of distance as indicated by faster responses to open hand gestures in the close relative to far distance trials irrespective of the type of hand response. Although participants' own hand responses were not visible during the task, thereby eliminating visuomotor feedback, the pattern of response errors matched that of the previous studies suggesting that even though participants could not see their hands complementary responses were facilitated for open hand gestures. The results indicate that motor priming affected hand responses irrespective of the specific outcome (mirror or complementary) as a function of space. Potentially, the effect of space on reaction times might therefore be related to differences in the stimulus set given that images depicted in the 3D glasses appear larger in close distance relative to the onscreen images.

In all studies presented so far, peripersonal space, manipulated by either changing the visual distance or physical obstruction between the participant and actor (Study 2), was always response-irrelevant. In terms of showing an automatic effect of distance on motor priming this is a strength of the design. However when responding to cues in their environment, people might not automatically take social distance information into account. Previous studies regarding the activation of motor affordances have shown that motor representations in response to objects or observed actions for instance, are only activated when these features are made task-relevant (van Elk, van Schie, & Bekkering, 2009; van Dam, van Dijk, Bekkering, & Rueschemeyer, 2012). By explicitly instructing participants to attend to differences in interpersonal space, it can therefore be investigated if motor priming in response to hand gestures is modulated as a function of distance when perceived distance is made task-relevant. This was the main aim of the fourth study.

Study 4

In the fourth study we used the same design and setup as in Study 1 but a blocked design was introduced where participants had to respond to hand gestures only when the actor was seated either in close or far distance depending on the block type. The study was approved by the Psychology Department of the University of Amsterdam ethics committee (2014-SP-3880).

Participants

In the fourth study, 20 participants took part including 12 Females with a mean age of 20 years old (range = 18-26). Thirteen participants were right-handed and three left-handed. Data for handedness for 4 remaining participants was mistakenly not recorded. All participants received course credit or monetary compensation for participating in the study and signed an informed consent form prior to participation.

Materials and procedure

The same stimuli were used as in Study 1. After a practice round of 20 trials, similar to the first study, participants were told that the main task consisted of 4 blocks, each including 40 trials, where they had to respond to hand gestures only when the actor was positioned across from the short or far end of the table depending on the block instruction. Two block orders where used where far and close distance instructions were alternated starting with either a far or close distance block (i.e., ABAB or BABA). During go-trials (80 % of the trials in each block) participants had to respond as instructed in the practice trials (i.e., mirror the hand gesture with either a congruent or incongruent response), during no-go trials (20 % of the trials in each block) participants were asked to keep their fingers on the buttons. For example, in a far distance block participants responded to 32 trials depicting an actor seated in far distance while withholding responses in 8 trials where the actor was seated in close distance. Wrong responses on no-go trials were not analyzed. In total, 160 trials were presented in 4 blocks, in which trials were randomly presented with respect to distance, hand gesture and hand response. In addition we changed part of a sentence in the introduction to explicitly state that participants had to imagine sitting at a table with the actor sitting across from them on the other side. Also, as in the third study, we stated that the actor would be sitting within reach for the short distance trials and outside reach in the far distance trials.

Data analysis

From the total sum of trials, 3.7 % (118) of trials were removed due to response errors. Additionally, 0.7 % (21) of all trials were removed for which response latencies were either above or below 2.5 x *SD* the individual participant mean.

Exit-questionnaire

Out of 20 participants, 65 % (13) thought it was possible to touch the hand of the actor in the close distance trials while only 15 % (3) thought it was possible in the far distance trials.

Results

First of all we checked if block order had any influence on the full within-subjects analysis. Neither a main effect for block order nor an interaction effect between block order and within factors turned out significant (all ps>.05). We first collapsed reaction time data over blocks and ran a 2 x 2 x 2 repeated measures analysis with Distance (close vs far), Hand gesture (fist vs open hand) and Hand response (congruent vs incongruent) as within-subject factors (see Fig 7). None of the main effects were significant (all *p*s > .05) nor the hypothesized three-way interaction ($F(1, 19) = 1.03$, $p = .322$), only the interaction between Hand gesture and Hand response yielded a significant effect, $F(1, 19) = 16.20$, $p = .001$ $\eta^2_p = .46$. Simple effects indicated that reaction times for congruent responses were slower (M = 770, SD = 117) compared to incongruent responses to open hand gestures (M = 714, SD = 108) irrespective of distance, $t(19) = 3.57$, $p = .002$, $d = .80$. There was no difference between congruent and incongruent responses to fist gestures ($p = .543$).

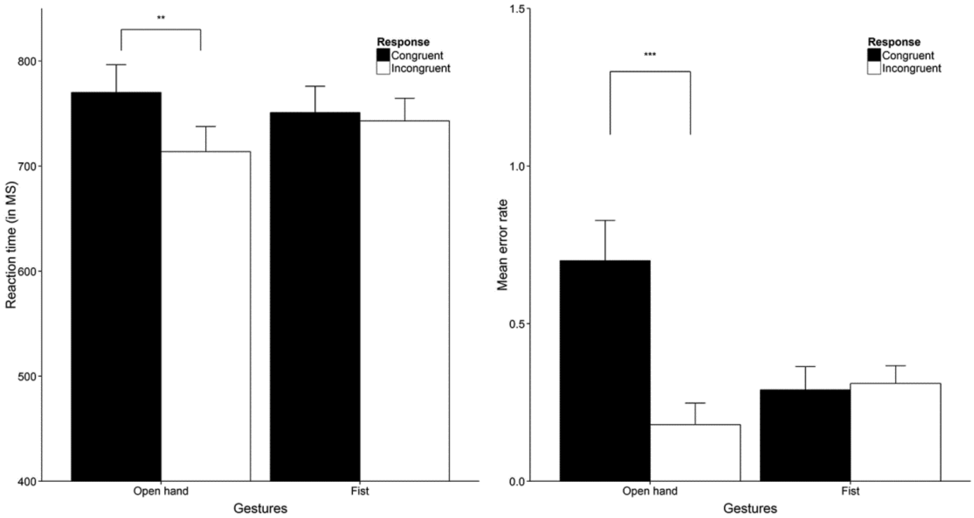


Fig 7. Response data Study 4. The RT results (Left) show reaction time in milliseconds averaged over response type (congruent and incongruent). The right graph shows mean error responses averaged over distance trials (close and far). Error bars are 1 x SE. * = p<.05, ** = p<.01, *** = p<.001

With respect to response errors we found a main effect for Hand response, $F(1, 19) = 8.44$, $p = .009$, $\eta^2_p = .31$, indicating that more errors were made in congruent (M = 0.49, SD = 0.45) relative to incongruent response trials (M = 0.24, SD = 0.28). More importantly, a two-way interaction was found between Hand gesture and Hand response similar to the previous studies, $F(1, 19) = 21.09$, $p < .001$, $\eta^2_p = .53$. Looking at the simple effects we found that participants made more errors when cued to make a congruent (M = 0.70, SD = 0.57) relative to incongruent response (M = 0.18, SD = 0.30) in open hand trials, $t(19) = 4.47$, $p < .001$, d = 1.00. No difference was found between fist trials where a congruent or incongruent gesture was the correct response ($p = .785$).

Discussion

The results from the fourth study corroborate the findings of Study 1 and 2 both in terms of reaction time latencies and response errors. Interestingly, making distance task-relevant did not reveal an effect of distance on complementary effects for open hand gestures.

Although the present experimental paradigm consisting of congruent and incongruent hand responses to open hand gestures has been used previously, it could well be that the so-called social complementary hand shaking effect is actually driven by low-level features of the stimulus. For example, one option is that the open hand was interpreted as a directional cue pointing towards the participant's incongruent relative to congruent response hand, rather than as a social action facilitating a complementary hand response. On this account, the priming effects observed would be a mere consequence of spatial stimulus features rather than reflecting social affordance effects. In favor of this notion, hand responses in a spatial cueing task have been found to be slower if a hand is presented in the center of the screen with a finger pointing away rather than towards the target position of the hand response (Crostella, Cadrucci, & Aglioti, 2009). This spatial or directional cueing account would fit well with the absence of an effect of distance on motor priming across the four studies that we conducted – as spatial cues are expected to affect motor responses irrespective of the distance at which they are presented (as long as they cue the same relative target location; see Lamberts, Tavernier, & d'Ydewalle, 1992). Previous studies have aimed to control for the potential confound that spatial stimulus features actually underlie the effects of observed hand gestures on behavior (see Heyes, 2011; Bertenthal, Longo, & Kosobud, 2006), but these studies have been primarily controlled for spatial confounds in imitative rather than complementary responses or, in the case of Flach et al. (2010), have replaced hand gestures by arrows rather than hand gestures that form spatial (or directional) cues. The fifth study was done to control for a general cueing effect produced by open hand gestures, which would provide an alternative and more low-level account of the effects observed in the first four studies.

Study 5

In this study a new set of images was produced in which an actor was displayed using his hand to point with an index finger along the same direction as the open hand pictures used in the first two studies. Given that a pointing finger does not represent a communicative gesture that requires interpersonal contact, this gesture allowed us to control for directional cueing effects produced by open hand trials in the previous studies. Given that our main goal was to see whether the pointing hand would show a directional effect that could replace the complementary effect and not the effect of space on processing the pointing hand, we only looked at hand gestures depicted in short distance.

Participants

For this study, 25 participants were recruited including 14 Females with a mean age of 23 years (range = 19-47). Nineteen participants were right-handed. All participants received course credit or monetary compensation in return for their participation in the experiment and signed an informed consent form before participating in the study. The study was approved by the Psychology Department of the University of Amsterdam ethics committee (2014-SP-3880).

Materials and procedure

New photographs were made of an actor making either a fist or pointing hand gesture using both the left and right hand. The body and hand gestures were cut out and pasted in the original pictures created in Study 1 (see Fig 8). Only the short distance trials were used for this study, so the introduction was adapted to remove information about the distance at which the actor was positioned. The procedure was the same as in the first study, participants were instructed to always mirror the perceived hand gesture (pointing for point hand gestures – fist for fist gestures) but use their opposite (same anatomical) hand when a green colored hand was observed.

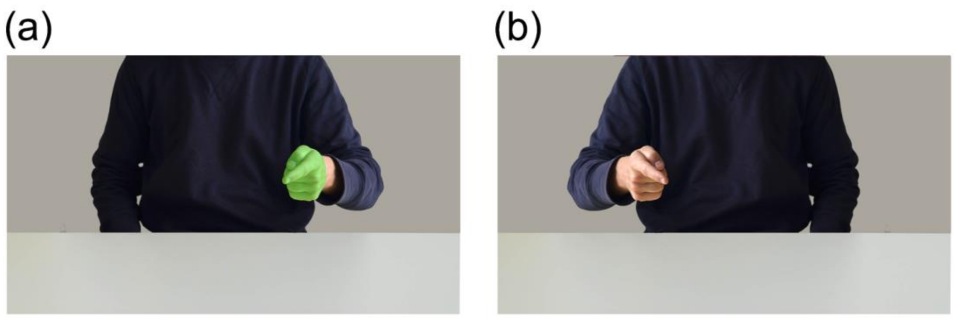


Fig 8. Sample images Study 5.. Sample images include (a) Pointing hand green right (b) pointing hand left

Data analysis

Two participants were excluded for which no correct responses were recorded for at least one response category. Based on the remaining data, we first removed 2.2 % (41) of the total sum of trials for which the reaction time was 2.5 x *SD* above and below the individual participant mean. Furthermore, 10 % (184) of trials were removed due to response errors.

Exit-questionnaire

Given that we only used close distance trials in this study we had participants only rate the close distance trials in terms of potential reachability. In total, only 21.7 % (5) of participants indicated it would be possible to touch the hand (pointing hand gesture) of the actor opposite from the table. In addition we asked whether participants felt they wanted to touch the hand of the actor. Only 10 % (2) of participants felt they wanted to do this.

Results

For the analysis we included data for 23 participants and ran a 2 x 2 repeated measures design with Hand gesture (fist vs pointing hand) and Hand response (congruent vs incongruent) as within-subject factors (see Fig 9). The analysis revealed a main effect for Hand response in which responses were faster for congruent (M = 928, SD = 693) relative to incongruent responses (M = 963, SD = 719) irrespective of the type of Hand gesture, $F(1, 22) = 4.33$, $p = .049$, $\eta^2_p = .16$. Given the skewness values results were analyzed using transformed (log) RT values as well. The interpretation of the ANOVA results do not differ between running the 2 x 2 analysis using transformed or untransformed estimates. Untransformed analyses are displayed here. No main effect for Hand gesture or an interaction effect between Hand gesture and Hand response was found. Additionally, no main effects or an interaction between Hand gesture and Hand response was found in the error pattern.

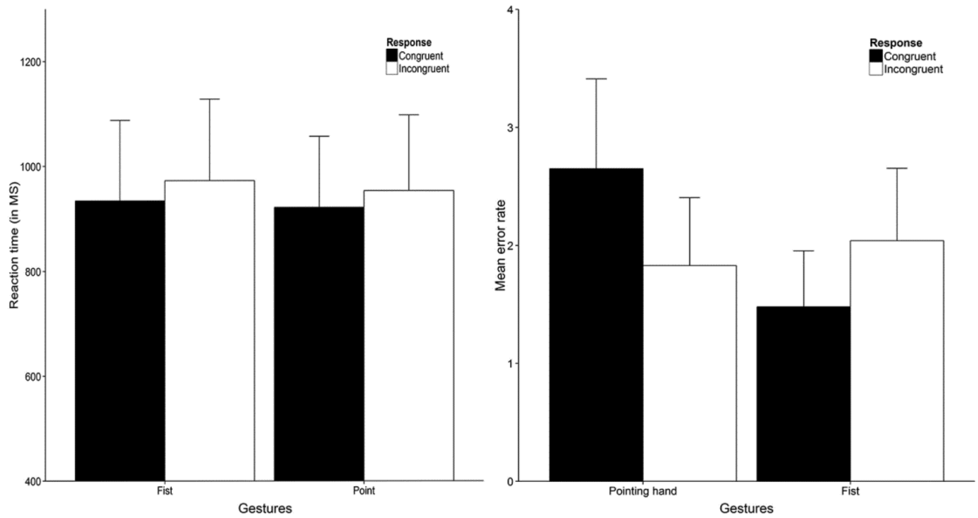


Fig 9. Response data Study 5. Reaction time (left) and response error data (right) for Study 5. Note that a main effect for response was found which is not shown in the figure. Error bars are 1 x SE. * = $p<.05$, ** = $p<.01$, *** = $p<.001$

Discussion

The results of the fifth study suggest that the complementary effect found in Study 1-4 cannot be explained by a more general cueing effect as alternative explanation to the social complementary effect. The results are in fact more in favor of a classical spatial congruency effect given the faster response times for spatially congruent relative to incongruent responses – irrespective of the specific hand gesture type that was presented.

Re-analysis across Study 1-4

Some of the effects observed appeared to be quite stable across the studies described here (i.e., the complementary effect for open hands) but other effects were less consistent (i.e., the effect of distance). Therefore, to get an overview of the combined effects over the first four studies (see Fig 10 and Fig 11) a combined analysis was conducted, with study as a grouping factor. The primary goal of this analysis was to see whether the complementary effect would be stronger in some studies compared to others, if the variability in subjective ratings would affect a sufficiently powered combined sample and whether differences in handedness would modulate the complementary effect. The design for this analysis was a 4 x 2 x 2 x 2 mixed design with Study (Study 1 to 4) as a between participants factor and Distance (close vs far), Hand gesture (fist vs open hand) and Hand response (congruent vs incongruent) as within-subject factors.

A significant main effect for Distance was found, demonstrating faster responses for close (*M* = 706, *SD* = 123) relative to far distance trials (*M* = 712, *SD* = 124), $F(1, 106) = 3.96$, $p = .049$, $\eta^2_p = .04$. However, this effect is weak at best considering the large sample size. Additionally, a significant interaction was found between Hand gesture and Hand response, $F(1, 106) = 45.30$, $p < .001$, $\eta^2_p = .30$. Interestingly, averaged over studies we found a congruency effect for fist gestures in that faster responses were made in congruent *(M* = 704, *SD* = 122) relative to incongruent response trials (*M* = 719, *SD* = 116), $t(109) = -2.34$, $p = .021$, $d = -0.22$. For open hand gestures, faster responses were made in incongruent (*M* = 690, *SD* = 107) relative to congruent response trials (*M* = 723, *SD* = 132), $t(109) = 4.76$, $p < .001$, $d = 0.46$. No interaction effect between Study and any of the other within-subject factors was found (range: $p = .082$ – $p = .802$). In addition we re-analyzed the raw data using a 3 x SD cutoff for outlier removal given that studies seem to differ in setting the criterion for removal (e.g., Liepelt et al., 2010) which might affect the robustness of our effects. Across studies 1-4 we found a similar interaction effect between Hand gesture and Hand response, $F(1, 109) = 42.99$, $p < .001$, $\eta^2_p = .28$ as well as a main effect for gesture, $F(1, 109) = 4.44$, $p = .037$, $\eta^2_p = .04$. None of the other factors reached significance (all *p*s > .05).

Another issue to resolve here is the variability in reachability ratings. Given that the images used in the first four studies were made to reflect settings in which physical contact was either possible or not the subjective ratings indicated that not all participants perceived this to be the case, which might be due to the specific stimulus set we used. This is problematic given that our hypothesis requires participants to correctly identify differences in close and far distance. The subjective estimates however resemble those used in other judgment tasks (e.g., Valdés-Conroy et al., 2014) even though participants in our studies had to picture themselves in a situation rather than being physically part of it. Additionally, in a separate post-test rating we found that subjects did (more) accurately interpret the pictures if asked to indicate whether the actor was within their personal space or whether extending one's hand made it possible to touch an extended hand of the actor (see the S1 appendix).

Effect	*Df*	*F*	*MSE*	*p*
Study 1				
Distance	(1, 39)	0.05	2,569.7	.83
Gesture	(1, 39)	0.69	2,636.5	.41
Response	(1, 39)	0.65	8,112.4	.43
Distance x Gesture	(1, 39)	0.98	2,561.5	.33
Distance x Response	(1, 39)	3.75	2,907.0	.06
Gesture x Response	(1, 39)	16.78	2,679.1	<.001
Distance x Gesture x Response	(1, 39)	0.16	1,454.9	.69
Study 2				
Distance	(1, 28)	0.52	948.2	.48
Gesture	(1, 28)	2.52	1,837.7	.12
Response	(1, 28)	0.40	8,368.9	.53
Distance x Gesture	(1, 28)	6.89	1,195.9	<.05
Distance x Response	(1, 28)	0.25	1,705.3	.62
Gesture x Response	(1, 28)	17.65	2,828.2	<.001
Distance x Gesture x Response	(1, 28)	1.13	1,134.2	.30
Study 3				
Distance	(1, 20)	3.35	3,012.5	.08
Gesture	(1, 20)	0.00	2,268.5	.95
Response	(1, 20)	0.00	3,813.9	.95
Distance x Gesture	(1, 20)	5.14	352.3	<.05
Distance x Response	(1, 20)	0.43	1,398.0	.52
Gesture x Response	(1, 20)	1.96	1,239.6	.18
Distance x Gesture x Response	(1, 20)	3.74	647.5	.07
Study 4				
Distance	(1, 19)	0.70	6,992.9	.41
Gesture	(1, 19)	0.68	1,534.2	.42
Response	(1, 19)	4.04	5,824.0	.06
Distance x Gesture	(1, 19)	0.00	2,764.6	.99
Distance x Response	(1, 19)	1.33	2,110.5	.26
Gesture x Response	(1, 19)	16.20	2,579.6	<.05
Distance x Gesture x Response	(1, 19)	1.03	1,391.5	.32
Study 5				
Gesture	(1, 22)	0.63	8,980.6	.44
Response	(1, 22)	4.33	6,644.6	<.05
Gesture x Response	(1, 22)	0.09	2,956.3	.77

Fig 10. ANOVA results for reaction time data (studies 1 through 5). Note: The factor Hand gesture is termed Gesture and Hand response termed Response

Study 1

	Close		*Far*	
	Congruent	*Incongruent*	*Congruent*	*Incongruent*
Fist	716.02 (20.1)	741.58 (22.3)	721.59 (21.0)	727.22 (18.6)
Open hand	727.66 (21.3)	709.23 (21.6)	747.84 (24.7)	702.65 (16.81)

Study 2

	No frame		*Frame*	
	Congruent	*Incongruent*	*Congruent*	*Incongruent*
Fist	669.07 (22.1)	683.38 (20.5)	676.50 (23.3)	705.59 (21.4)
Open hand	696.68 (25.8)	661.73 (18.1)	689.69 (24.1)	650.71 (17.3)

Study 3

	Close		*Far*	
	Congruent	*Incongruent*	*Congruent*	*Incongruent*
Fist	684.72 (29.0)	680.31 (19.5)	682.27 (28.2)	700.61 (25.9)
Open hand	678.63 (24.2)	674.18 (19.5)	704.50 (28.7)	692.43 (25.2)

Study 4

	Close		*Far*	
	Congruent	*Incongruent*	*Congruent*	*Incongruent*
Fist	738.64 (23.5)	744.33 (25.0)	747.26 (23.3)	757.71 (27.6)
Open hand	771.79 (28.8)	700.85 (24.2)	768.53 (28.2)	726.33 (26.3)

Study 5

	Congruent	*Incongruent*
Fist	934.11 (153.6)	972.81 (155.6)
Pointing hand	921.77 (135.4)	953.80 (144.3)

Fig 11. Mean reaction time results for Study 1 through 5 as a function of distance (or frame) and congruency (columns) for both fist and open hand gestures (+*SE*s within brackets)

This suggests that the variability in ratings might be due to the specific question we used. In terms of the current re-analysis one solution we used was to look at the role of Reachability in each study separately by splitting participants in groups that did or not did not accurately perceive the distance estimates. However, this might be affected by power issues (insufficient sample sizes) in each individual study. In order to circumvent this problem we added Reachability (averaged across studies) to the across-studies analysis, by putting all participants who successfully perceived our distance/ reachability manipulation (actor in close distance is reachable – actor in far distance is unreachable) into one group (53 participants) and the remaining participants in a second group (57 participants). There was no main effect for Reachability ($F < 1$) nor did inclusion of Reachability affect the interaction between Hand gesture and Hand response, $F(1, 108) = 48.84$, $p < .001$, $\eta^2_p = .31$. For more detailed analyses see the S2 appendix.

Besides Reachability, we added Handedness (left vs right-handed) to the within-participants analysis (averaged over Study 1 to 4). There was no main effect for Handedness ($F < 1$) and although adding handedness did decrease the strength of the Hand gesture by Hand response interaction, it was still significant, $F(1, 98) = 18.12$, $p < .001$, $\eta^2_p = .16$. However, the sample sizes (right handed = 113; left handed = 16) were strongly imbalanced so we analyzed data for right and left-handed participants separately. For these analyses we added a within-participants level of Response side (left vs right) to look at differences for the responding hand (e.g., right and left responses to left and right open hand gestures). This resulting analysis was a 2 x 2 x 2 x 2 with Distance (close vs far), Hand gesture (fist vs open hand), Hand response (congruent vs incongruent) and Response side (left vs right). Results showed a strong Hand gesture by Hand response interaction for right-handed participants irrespective of the responding hand, $F(1, 93) = 40.72$, $p < .001$, $\eta^2_p = .31$ while for left-handed participants this interaction was around the critical alpha-level of $p = .05$, $F(1, 11) = 4.83$, $p = .050$, $\eta^2_p = .31$. However, the sample size does not permit us to make strong conclusions about the group of left-handed participants. Interestingly, for both right and left-handed participants there was no main effect of Response side nor an interaction between Hand gesture, Hand response and Response side (all *F*s < 1), suggesting no modulation of the complementary effect by hand-dominance.

The response error data revealed a main effect for both Hand response, $F(1, 106) = 8.47$. $p = .004$, $\eta^2_p = .07$, as well as Hand gesture, $F(1, 106) = 4.05$, $p = .047$, $\eta^2_p = .04$ qualified by a two-way interaction between Hand gesture and Hand response, $F(1, 106) = 36.11$, $p < .001$, $\eta^2_p = .25$. No difference in the mean error rate for fist gesture trials was found, $t(109) = -1.90$, $p = .060$ but for open hand trials more errors were made across studies in congruent ($M = 1.25$, $SD = 1.17$) relative to incongruent trials ($M = 0.46$, $SD = 0.73$), $t(109) = 6.46$, $p < .001$, $d = 0.62$. When instead of Study, we added Reachability to the within-participants design the Hand gesture by Hand response interaction was still significant ($F(1, 108) = 43.06$, $p < .001$, $\eta^2_p = .29$) while Reachability did not interact with any of the remaining within-subject factors.

Bayesian analysis

Finally, we performed a Bayesian analysis of the full sample (Study 1 to 4) in order to compute the relative evidence in the data in support of the model including the Hand gesture by Hand response interaction compared to the model including the three-way interaction between Distance by Hand gesture by Hand response. Furthermore, we wanted to gauge the relative strength of the simple main effects for fists and open hand gesture trials across distance in order to see how they would fare in comparison to the frequentist analyses. Bayesian analyses provide an opportunity for model comparison (e.g., between two-way and three-way interactions), which allows us to find relative evidence in favor of the two way interaction rather than only refuting the three-way interaction. All analyses were performed using JASP, which is an open source statistical software tool that can be used for both frequentist and Bayesian analyses (Love et al., 2015). First off, we compared the relative fit of an ANOVA model including (Model 1) or excluding (Model 2) the critical 3-way interaction effect between Distance, Hand Gesture and Hand Response (in addition to modeling the main effects and two-way interactions in both models). The relative fit was calculated by dividing the Bayes Factor (BF) of Model 1 by the BF of Model 2 (including the three-way interaction) which produced a BF of ± 5.84 in favor of the simpler model (Model 1; without the three-way interaction). This factor therefore represents the unique contribution of the three-way interaction. Taking these steps was necessary to find the unique variance contribution of a single interaction effect (e.g., the three-way interaction) given that JASP produces models for interaction effects that include all lower level (main) effects.

The same procedure was performed for the two-way interaction between gesture and response. For this interaction we compared the relative fit of a model without the two-way interaction term between gesture and response (Model 3) compared to a model including the two-way interaction term (Model 4; in both models the main effects for gesture and response were included). A BF of ± 97.3 million in favor of Model 4 was found. Subsequently we could infer what the relative fit was of the unique contribution of the two-way interaction term (Hand gesture vs Hand response) compared to the unique contribution of the three-way interaction (Distance vs Hand gesture vs Hand response) by dividing the appropriate Bayes factors by each other. The result suggested that the data was ± 568.3 million times more in favor of the model including the two-way interaction compared to the model including three-way interaction, which counts as sufficiently strong evidence according to interpretation criteria for Bayes Factors (Jeffreys, 1961; Kass & Raftery, 1995). Thus, this analysis suggests that our data is best explained by a model describing the interaction between observed gesture and performed response, rather than a model that takes into account the role of distance. This suggests that the effects of observed on performed gestures is quite strong and automatic and occurs irrespective of the distance at which an action is observed.

Additionally, a Bayesian t-test was done comparing the mean RT difference for congruent and incongruent responses to fist gestures and open hand gestures. For fist gestures a BF10 of 1.43 was computed which suggests that support for the alternative hypothesis, as specified by a Cauchy prior distribution d~N(0,.707), compared to the null hypothesis of no effect was weak or inconclusive. This confirms the idea that in larger samples, increasingly larger effects (e.g., t-values) are necessary to produce Bayes factors that favor the alternative hypothesis over the null hypothesis (Rouder, Speckman, Sun, Morey, & Iverson, 2009). In contrast, for open hand trials a BF10 of 2504.86 was found suggesting that the data support the alternative hypothesis that there was an effect by a factor of ± 2504 under this specific alternative hypothesis relative to the null hypothesis (which doubled when specifying a directional effect for faster response times in incongruent relative to congruent trials). However, it is important to note that the specified prior distribution is not fully informative (i.e., unspecific) and the results for the fist gesture trials should therefore be interpreted with caution. One way to resolve this is to perform a robustness analysis in which different scaling parameters are used for the prior distribution (instead of 0.707) in order to see how being more or less specific about the predicted distribution affects the Bayes factor estimate. For fist gesture trials, varying the scaling parameter (from 0.5 to 1.5) produced a maximum Bayes factor of 1.85 in favor of the alternative hypothesis (a difference in response times), which falls within the bounds of anecdotal evidence. Even when estimating a directional effect (i.e., faster response times for congruent relative to incongruent trials) with the most narrow distribution here (d~N(0,.5)) the Bayes factor does not exceed 3.66 in favor of the alternative. This suggests weak evidence at best for a congruency effect in fist gesture trials. Given that the chosen ‘default’ prior produces a conservative estimate (biased with respect to small effects) we can be confident with respect to the results for open hand gestures. These results corroborate and extend the findings in the original ANOVA analysis.

General Discussion

The current paper demonstrates a facilitation effect for performing incongruent responses to observed open hand gestures relative to making congruent responses (i.e., a complementary effect). This effect was only present for open hand gestures but not for intransitive (fist) gestures supporting the idea that the communicative meaning of the hand gesture was driving the facilitation effect. The fifth study supports this conclusion by showing that the complementary effect is not found when replacing the open hand with a pointing hand. The current set of studies therefore replicate the findings in Liepelt et al. (2010) and Flach et al. (2010) and extend them by using more environmentally rich stimuli. However, against our predictions, the complementary effect for open hand gestures was found irrespective of the depicted distance in space, either in terms of physical (Study 1) or functional distance (Study 2) between the participant and the person making the gesture. Making space task-relevant, by letting participants directly respond to the spatial context of the perceived gesture, also did not affect the complementary effect. When taking into account ratings of reachability this did not alter the effect suggesting that motor priming (as operationalized here) seems to be driven by the known contingent relationship between a gesture and a response rather than the perceived possibility to perform the response in interaction with another person. However, given the variability in reachability ratings we cannot fully exclude that this unexpected finding might have been related to the use of sub-optimal stimuli. Nevertheless, the results (partly) support the role of sensorimotor learning, in which specific stimuli are associated to actions that do not only reflect spatially or anatomically matching information but learned (social) contingent relations as well (Liepelt et al., 2010). In contrast to our findings, it has been argued that environmental cues, including interpersonal space or contact, modulate motor priming given the predictive function of these cues in stimulus-response learning (Cavallo, Heyes, Becchio, Bird, & Catmur, 2014). For example, interpersonal space between individuals could be encoded during handshaking so that complementary actions are triggered by open hand gestures specifically when perceived within peripersonal space. Our results do not show any influence of interpersonal space and rather suggest that the affordance provided by perceiving an open hand drives the motor priming effect in an automatic fashion. Interestingly, this seems to be due to the meaning of the gesture rather than being the product of individual experience given that handedness did not modulate the effect (i.e., no preference was detected for right-handed participants to respond with their right vs left hand to open right hands). This seems to be in line with evidence for bimanual responses to complementary actions in right-handers as seen in Sartori, Begliomini, Penozzo, Garolla, and Castiello (2014).

We did find effects of distance in single studies such as the distance by gesture effect in Study 2 and the distance by gesture effect in Study 3. However, when looking at the Bayesian analysis, only a weak main effect of distance was found, which by computing the Bayes factor of BF01 = 3.40 actually favors the null hypothesis over the model including distance (given a Cauchy prior distribution of d~N(0,.707)). None of the additional interaction effects survived in the overall sample, specifically the theoretically interesting Distance by Hand gesture by Hand response interaction. One reason for the absence of an effect of perceived interpersonal space might have been the nature of our stimulus material. While using a similar paradigm as Liepelt et al. (2010) in which stimuli were presented using 2D pictures on a computer screen, our third study controlled for the subjective 'realness' of the stimulus set and in particular the distance manipulation. Even though the results in this study demonstrated an advantage for hand responses in close distance the follow-up comparisons only showed an effect of distance for open hand gestures irrespective of the type of response (i.e., participants made faster responses to open hand gestures that were presented in close distance).

The absence of the complementary hand effect in the third study may be related to the fact that when presenting the stimuli in 3D, subject's own hand responses were not visible during the task. The visibility of your own hands might be important given that response-effect associations are partly formed using forward models (Wolpert & Kawato, 1998) where a hand movement is accompanied by specific sensory effects. Based on these learned sensorimotor associations, we automatically generate expectations about visual feedback as soon as an action (e.g., hand movement) is initiated (i.e., akin to the ideomotor principle; Greenwald, 1970). The absence of visual feedback from the own hand might thus create a temporary mismatch in the response-effect association for the open hand response, which in turn might have affected response latencies. However, given that the error pattern (albeit small) reflected a tendency to perform complementary actions in response to open hand gestures this remains a speculative explanation.

Furthermore, in the third study the spatial position of gesture cues differed in terms of distance (close vs far) when comparing 3D pictures to the 2D referent pictures. That is, cues were presented more laterally in the close compared to the far distance trials for 3D stimuli due to the physical proximity of the stimuli (screen vs head-mounted display). Specifically, the head-mounted display provides a wide angle view so that the stimulus (actor) is perceived as much larger compared to the 2D pictures on a computer screen (larger in visual degrees), which affects the position of the hand gestures relative to the fixation mark (center of the screen). Previous studies have indicated that congruent motor priming can be modulated by distance in a 3D environment (Costantini et al., 2011). However, using 3D images to enhance the perception of distance does not seem to be a necessary requirement for the distance effect to be found (Costantini et al., 2010). This discrepancy between previous findings and our results, relates to a broader discussion about the different potential of affordances in real situations versus pictures depicting real situations (Wilson, 2014). Whereas here we have shown that pictures have the potential to facilitate hand responses as a proxy for physical responses, the role of distance is not evident. Perhaps this is due to a dissociation between processing the information that social actions provide and the possibility to physically respond to these cues, suggesting that pictures might be limited for conveying specific information. Therefore, future studies should experimentally distinguish settings in which participants can realistically interact with the environment from settings where participants passively observe it (Kourtis, Sebanz, & Knoblich, 2010; Schilbach, 2014; Wilson & Golonka, 2013).

Another explanation for the absence of the modulation of space on the complementary effect is that the effect actually reflects the automatic simulation of the partner's behavior during observation (Brass et al., 2000), which should not depend on distance. Some studies have shown how in interactive settings, people simulate actions from interacting partners prior to movement (Kourtis, Sebanz, & Knoblich, 2013) and show similar response inhibition patterns in no-go trials even when the no-go rule applies only to the co-acting partner in the task (Sebanz, Knoblich, Prinz, & Wascher, 2006). In the same vein, in the current experiment, participants could be simulating perceived communicative gestures as if they would perform the observed behavior themselves rather than responding to it. This seems plausible given that in our case the observed and performed gesture were similar (matching). However, this idea would go against Liepelt et al. (2010), van Schie et al. (2008) as well as Sartori, Bucchioni and Castiello (2012) in that here responses to complementary requests (in a picture or video presentation task) are interpreted as learned interactive responses rather than responses driven by a first-person simulation of the perceived gesture. In addition, a number of studies have suggested that when imitating observed behavior people tend to imitate movements in a mirrored (specular) fashion rather than in an anatomical similar way (Liepelt et al. (2010); Koski, Iacoboni, Dubeau, Woods, & Mazziotta, 2003). In sum, we think it is unlikely that our findings of a complementary effect for observed open hand gestures simply reflects a process of motor simulation; rather we suggest that the effect reflects the strong and automatic activation of learned associations by responding in a socially prescribed and overlearned fashion.

Besides the success in replicating the complementary effect we failed to replicate the congruency effect for intransitive gestures as demonstrated in Liepelt et al. (2010; i.e., a mirror congruency effect for fist gestures was expected). Interestingly, though, the meta-analysis across the four studies reported here did show the congruency effect. However, when comparing the congruency effect for fist gestures from the meta-analysis in our study to that found in Liepelt et al. (2010) it appears that the resulting effect size is roughly four times smaller. This also seemed to affect the Bayesian analysis, which suggested weak or anecdotal evidence for the existence of a congruency effect compared to the frequentist analysis. One of the ways in which our study deviated from this study is that in our study gestures were presented at spatially opposite sides of the screen so that the mirror congruency effect might have been confounded with spatial congruency. Based on Bertenthal et al. (2006) we know for example that when dissociated, spatially congruent hand movements in response to intransitive hand gestures are more strongly facilitated than anatomically congruent hand movements. Similarly, for communicative hand gestures, Flach et al. (2010) found that participants were faster in making hand movements when open hand gestures or arrows were subsequently presented at spatially congruent locations on the screen. However, based on these findings we should have expected a stronger rather than weaker congruency effect for intransitive gestures. A simpler explanation is that our stimuli differ from Liepelt et al. (2010) in that they depict a full body posture instead of only showing a hand gesture. While van Schie et al. (2008) as well as Ocampo and Kritikos (2010) showed an advantage of performing congruent hand responses using similar full-body postures in their studies this effect was driven by the task instruction to either perform congruent or incongruent responses in non-critical trials, something we did not manipulate here. Thus, by adding contextual information, we therefore perhaps strengthened the social information provided in the stimulus (full body vs single hand) which at the same time decreased the salience of the spatial attributes of the stimulus.

So even though some questions remain unsolved, here we have demonstrated and replicated a complementary motor priming effect in response to open hand gestures irrespective of the space in which these gestures are perceived. Note that this effect does not generalize to all complementary actions given that some complementary actions do not require physical interaction (e.g., throwing, catching). Specifically, they do not generalize to actions that require both hands and mutual response coordination (e.g., carrying a table), where actions need to be adjusted continually in order to perform a task successfully. In this case reachability is an essential precondition for interaction while in the current study set actions were essentially single shot responses where no coordination took place. Mutual actions are more complex and therefore are not easily translated to an experimental paradigm but recent innovative task settings might provide a suitable way to extend current findings (Kokal, Gazzola, & Keysers, 2009).

On the basis of the current findings we propose that motor priming is primarily a function of stimulus-response coding that does not necessarily correspond with motor priming in real situations where the physical opportunity to respond to objects affects motor responses. Given that effects of space on motor priming have been reported for objects (Costantini et al., 2010), we did not find effects of space on motor priming for hands which could highlight that the social relevance of a target object may automatically necessitate a response to it, which is something for future research to examine. Also, the fact that social and non-social requests differ in terms of the consequences (not responding to a handshake can be a sign of social miscommunication) could be investigated by directly comparing social and non-social target cues in a single design. The current set of studies extends our knowledge of motor priming and demonstrates the limitations of specific information provided in paradigms using pictures (e.g., distance) with respect to the information these pictures are designed to reflect (Yu, Abrams, & Zacks, 2014).

Supporting information

S1 Appendix. Reachability ratings.

S2 Appendix. Post-test ratings.

S1 Appendix. Post-test ratings

A post-test rating was administered to a separate group of raters. This test was done to further examine the way the pictures were perceived during the experiment and whether other attributes of peripersonal space were accurately perceived in the stimuli used in the task. One such attribute is the notion that peripersonal space requires people to see that objects or other people are within one's peripersonal space such that extending an arm would allow one to reach and manipulate a target object/ person.

Participants

In total, 46 participants (University of Amsterdam students) were included in the post-test. A 'catch question' was included to see if participants would accurately read the instructions. Data for three participants who filled in the catch question (when requested to not do so) was deleted, creating a total of 43 responses. The mean age was 21 years old (range = 18-66) and the sample consisted of 37 female participants and included 35 right-handed participants. All participants received course credit for participation and signed an informed consent before starting the experiment. The test was approved by the University of Amsterdam ethics committee (2015-SP-6354).

Stimuli and Results

For question 1 and 7 a picture was added from Study 1 depicting an actor sitting in either close or far distance making an open hand gesture with a right hand.

	Answers		
Questions	**Yes**	**No**	**Unclear**
1. Do you feel it is possible to touch the hand of the person on the opposite side of the table if this person was sitting opposite to the **short end of the table** (see picture below)?	30	11	2
For the following questions, imagine you are sitting opposite to the person in the picture (opposite to the **short end of the table**).			
2. Is the person within reach?	34	8	1
3. If you and the person would extend their arms, would you be able to touch each other?	41		0

4. Can you shake the person’s hand?	39		2
5. Could you understand the person if he/she would say something?	43		0
6. Could you catch a ball if this person were to throw one at you?	41		0
7. Do you feel it is possible to touch the hand of the person on the opposite side of the table if this person was sitting opposite to **the long end of the table** (see picture below)?	1	0	2
For the following questions, imagine you are sitting opposite to the person in the picture (opposite to the **long end of the table**).			
8. Is the person within reach?	1	1	1
9. If you and the person would extend their arms, would you be able to touch each other?	7	4	3
10. Can you shake the person’s hand?	3	7	3
11. Could you understand the person if he/she would say something?	42		
12. Could you catch a ball if this person were to throw one at you?	42		

For all options where unclear was chosen, participants were free to note reasons explaining the clarity of the pictures. With respect to question 2 some participants noted that “the full size of the table was not visible”, or “it is unclear what the size of the table is”. For question 4 one participant noted that “if the person would extend his arm just like in the picture, I would have to move across the table and it would therefore not be possible but if he would extend his full arm it would be possible”. Given the low rate of uncertain choices it can be concluded that the pictures and the distance manipulation presented therein were sufficiently clear. Also, looking at the answers on item 3 and 4, it seems that in terms of peripersonal space, the pictures were seen as depicting settings in which physical contact is possible while in the answers 7, 8, and 9 the majority of participants concluded this was not possible in terms of peripersonal space. One source of confusion could be the length of extension of the arm in the depicted scenes. With respect to the remaining questions, it is clear that in both the close and far distance settings, social interactions that do not require direct physical contact are seen as equally possible (see items 5, 6, 11, and 12).

Taken together, the current findings suggest that (the vast majority of) participants could accurately perceive the pictures as depicting settings in which physical contact would be (im)possible in terms of the interpersonal distance. The items used here might therefore be more specific in referring to the notion of peripersonal space than the original items used in the paper, which might explain the apparent discrepancy between the reachability ratings observed in the main studies and in the post-test.

S2 Appendix. Reachability ratings

In study 1-4 we looked at the effect of Reachability, as indicated by the imagined possibility to make physical contact with the perceived actor onscreen (taken from the exit questionnaire), on the main within-subjects design (Distance x Hand gesture x Hand response). We initially looked at Reachability separately for each study before and subsequently conducted a re-analysis of all studies combined

Reachability in each study was assessed by selecting participants who indicated that it would be possible to touch hands with the actor across from the short end of the table and impossible to touch hands with the actor across from the far end of the table

Study 1

Participants who thought it was possible to touch hands with the actor across from the short end of the table and impossible to touch hands with the actor across from the far end of the table were put in the first group (n = 16), the remaining participants were put in the second group (n = 24). By adding Reachability to the full 2 x 2 x 2 repeated measures analysis we found a significant interaction effect for Distance and Hand response, $F(1, 38) = 4.50$, $p = .041$, $\eta^2_p = .11$. For this interaction no simple main effects reached significance. More importantly, the interaction between Hand gesture and Hand response was still significant when including Reachability in the analysis, $F(1, 38) = 15.69$, $p < .001$, $\eta^2_p = .29$. The remaining effects including interactions between Reachability and within participants factors (Distance, Hand gesture and Hand response) were all non-significant (all Fs < 1). So adding Reachability did not affect the hypothesized three-way interaction between Distance, Hand gesture and Hand response.

Study 2

As in the first study we added Reachability as a between participants factor to see how ratings of the perceived ability to interact would affect response time data. Fifteen participants indicated that it was possible to make physical contact in the no frame trials and impossible in the glass frame trials. All remaining participants were put in the second group. If this variable was added to the analysis we again found an interaction effect for Hand gesture and Hand response, $F(1, 27) = 17.07$, $p < .001$, $\eta^2_p = .39$. Also the interaction between Frame and Hand gesture was still significant, $F(1, 27) = 6.60$, $p = .016$, $\eta^2_p = .20$. However, Reachability did not interact with any of the remaining variables.

Study 3

In terms of Reachability, one group was formed with twelve participants who indicated it was possible to make physical contact with the actor in close but not far distance trials, the remaining participants were put in the second group. Adding this grouping variable as between participants factor did not affect the Distance by Hand gesture interaction, $F(1, 19) = 5.63$, $p = .028$, $\eta^2_p = .23$. No other main and interaction effects were found.

Study 4

Based on the reachability ratings, two groups could be distinguished, including one group of 10 participants who thought it was possible to touch the hands of the actor across from the short end of the table and impossible from the far end of the table. The remaining participants were put in the second group. The interaction effect between Hand gesture and Hand response was still significant when including Reachability in the within-subjects analysis, $F(1, 18) = 15.35$, $p < .001$, $\eta^2_p = .46$. There was no main effect of Reachability nor did Reachability interact with the Hand gesture by Hand response interaction (both *F*s < 1).

Chapter 4

Perspective switching in social interactions: Does responding favor an egocentric perspective?

This chapter is based on: Faber, T. W, & Jonas, K. J. (2016). *Perspective switching in social interactions: Does responding favor an egocentric perspective?* Manuscript in preparation.

Abstract

Perspective taking has been argued to be a key factor in successfully participating in social interactions. Indeed, it has been shown how observers voluntarily take the perspective of an actor (allocentric perspective taking) who is simply present or performing object-directed actions. Even though these tasks elucidate the social function of perspective taking, so far experiments have been used that constrain the interactive nature of social settings (e.g., in which the observer takes a passive rather than active role). We predicted that perspective taking is facilitated when observing others perform manual actions but is less functional when a complementary request is observed which requires responding from a first-person (egocentric) perspective. Three studies are presented here measuring spatial perspective taking in a voluntary choice task (Study 1) or an explicit reaction time task (Study 2 and 3). Although our results suggest an increased difficulty in making spatial judgments in interactive settings this was true for both egocentric and allocentric perspective instructions. We discuss these findings in light of stimulus properties and task difficulty.

Introduction

Social interactions benefit from perspective taking if people successfully infer beliefs and intentions of others. To complete interactions however, people must switch to their own perspective in order to complement or prepare fitting responses in accordance with the goals of the interacting partner (Sartori & Betti, 2015; Hamilton, 2013). For example, when somebody throws you a ball it is important to infer whether the other expects you to throw it back or pass it to a third person, whereas actually catching the ball requires switching to a first-person perspective to accurately track the position of the ball in order to catch it. Although this example demonstrates how an interactive setting requires both taking the other's (allocentric) perspective as well as one's own (egocentric), this is hardly reflected in experimental operationalizations. In fact, most perspective tasks are restricted to passive rather than interactive displays to estimate perspective inference. Therefore, we seek to test whether changes in the interactive nature of a task facilitates or inhibits allocentric perspective taking with three studies.

Determinants and biases of perspective taking

Tracking of intentions or Theory of Mind (ToM) abilities have been studied in terms of their role in human interaction in providing an understanding of other minds by successfully dissociating one's thoughts or perspective from that of others (Frith & Frith, 1999; Baron-Cohen, Tager-Flusberg, & Cohen, 2000). Besides false-belief tasks, in which the ability is tested to dissociate one's own belief from that of other's regarding the same object, a number of studies have used spatial perspective tasks as a means to test (visual) perspective dissociation for both children and adults (Piaget & Inhelder, 1956; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010; Tversky & Hard, 2009; Mazzarella, Hamilton, Trojano, Mastromauro, & Conson, 2012; Furlanetto, Cavallo, Manera, Tversky, & Becchio, 2013). Typically in these tasks, people are asked to judge the position of objects or spatial cues from either their own perspective or that of an actor also present in the scene. Spatial perspective taking differs from visual perspective taking which is concerned with judging whether and how something is seen from another's perspective but evidence suggests that visual and spatial perspective taking use similar computational processes (Surtees, Apperley, & Samon, 2013). In both types of perspective taking a distinction can be made between level-1 and level-2 perspective taking (Flavell, 1974; Flavell, Everett, Croft, & Flavell, 1981). Whereas level-1 perspective taking is concerned with the understanding that something that is visible to the self does not have to visible to somebody else, level-2 perspective taking is concerned with understanding that something which is both visible to the self and the other leads to different visual experiences. It has been demonstrated that specifically level-2 perspective taking requires egocentric mental rotation which is defined as mentally rotating your perspective in order to compute what others, sitting at a different angle from you, are seeing (Surtees et al., 2013).

From a number of spatial perspective tasks it has become evident that primarily children and to some extent adults have a strong tendency to see and judge the world from their own perspective (Piaget, 1930; Birch & Bloom, 2004; Epley, Morewedge, & Keysar, 2004). This egocentric bias is a product of a cumulative learning process involved in interactions with others, objects and other stimuli, which generates sensorimotor connections based on egocentric sensory consequences of self-produced actions (Jackson, Meltzoff, & Decety, 2006; Vogeley & Fink, 2003). More specifically, when observing self-initiated movements this produces feedback from actions seen from a first-person perspective which at the same time can be dissociated from actions performed by others (Vogeley & Fink, 2003). Given that level-2 perspective taking tasks require mental rotation strategies, rotating one's egocentric perspective to the perspective of the other is quite effortful. An egocentric bias is then revealed by the difficulty in processing another's perspective if the perspective of the other produces an experience that is incongruent relative to congruent with one's own. Evidence for this bias is however restricted to tasks in which (spatial) perspective taking is measured using single judgments in a non-interactive setting or at a single time-point in an interaction. In real social interactions people might use perspective taking variably over time and in service of interactive tasks such as preparing upcoming responses. To (partly) resolve this, a number of studies have been done that have identified factors part of social interactions that facilitate (allocentric) perspective taking or interfere with responding from an egocentric perspective.

One such factor is gaze cueing which facilitates perspective taking by directing (joint) attention towards the direction of other's gaze, regarded as an early developmental marker of intention inference (Tomasello, 1995; Baron-Cohen, 1994; Shepherd, 2010, Driver, Davis, Ricciardelli, Kidd, Maxwell, & Baron-Cohen, 1999; Frischen, Bayliss, & Tipper, 2007; Samson et al., 2010). Although gaze cueing seems to automatically interfere with egocentric judgments in level-1 perspective taking (Samson et al., 2010) it is unclear to what degree this reflects a strictly social phenomenon given that a similar interference effect has been found when replacing social gaze cues (human avatars) with non-social directional cues (arrows; Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2014). Besides gaze cueing it has been shown that observing others perform simple object-directed movements can facilitate perspective taking. For example, in two studies using a level-2 spatial perspective taking task participants either observed pictures of an actor sitting at a table moving their hand towards an object (Tversky & Hard, 2009) or an actor holding an object while gazing straight forward (Mazzarella et al. 2012). The task instruction was to judge the position of the target object relative to a second object placed to the side which could be done from either the participant's own perspective or that from the actor in the scene. Results from these studies demonstrated that the frequency of taking the actor's perspective increased relative to pictures including only objects or an actor seated at a table not performing an action. Importantly, when combining gaze cueing and hand movements, perspective taking increased when the gaze direction was inconsistent (e.g., gaze cueing left – hand moving right) relative to consistent with the direction of the hand (Furlanetto et al., 2013). These results suggest that even though perspective taking is affected by an egocentric bias, information about action plans or uncertain action outcomes produced by gaze-action incongruency tends to facilitate (allocentric) perspective taking. This increase in perspective taking is interpreted in terms of the social relevance of the perceived action for the observer (Mazzarella et al., 2012) given that perceived actions may require interactive responses. Even though the data fit the idea that perspective taking is especially functional in social interactions the tasks used to test perspective taking still display non-interactive settings where either an actor is gazing straight (towards the subject) but not clearly planning to interact.

Perspective Switching

The fact that most spatial perspective tasks lack interactive properties is not problematic according to the idea that perspective taking is a function of (automatically) anticipating actions even when these actions do not require immediate behavioral responses (Tversky & Hard, 2009). However, this means that clear evidence for the role of perspective taking in interactive settings, which should benefit most from perspective taking, is still lacking. One option is that perspective taking increases when changing from a passive to an interactive (interdependent) situation given its increased functionality in these settings (Furlanetto et al., 2013). Alternatively, social interactions might benefit from effective perspective switching between egocentric and allocentric perspectives (Apperley et al., 2010). This means that perspective taking might be inhibited rather than facilitated depending on the (task) roles of those that take part in the interaction. As suggested in the introduction, computing the perspective of an interacting partner benefits the observer but might interfere with subsequently responding to a complementary request from an egocentric perspective (e.g., giving and receiving an object). A number of related studies on action simulation have demonstrated how action simulation (internally simulating observed movements) is modulated by the social or complementary role between the actor and observer in an interactive setting (Sartori, Bucchioni, & Castiello, 2013; Kourtis, Sebanz, & Knoblich, 2010). For example, individuals seem to simulate actions performed by an actor until the moment the observed action changes into a complementary request, at which point people (covertly) prepare fitting behavioral responses (Sartori et al., 2013). Although this fits with the notion of perspective switching, the link between simulating actions performed by others and perspective taking is not evident. Specifically, it is still unclear whether simulating actions performed by others implies simulating the action from one's own perspective or reflects a mental image of the observed behavior displayed by the other (Gallese, Keysers, & Rizzolatti, 2004; Ramnani & Miall, 2004; Sebanz & Frith, 2004). Moreover, the task in Sartori et al. (2013) does not require participants to estimate the actor's perspective and perhaps therefore does not require mental rotation. Nonetheless, given that perspective taking is a simulative process (Goldman, 2006) changing perspectives could interfere with the preparation of complementary responses which is something we aim to test here.

Current Research

We examined whether minimally changing the (social) interactive nature of the task setting facilitates or inhibits allocentric perspective taking relative to non-interactive settings. We expected that in line with the idea of perspective switching in social interactions, people would have a stronger tendency to voluntarily judge interactive situations (a depiction of an actor giving an object to the participant in the study) from an egocentric perspective than situations in which an actor is present only holding an object (H1). Additionally, since responding to a complementary request requires one to pay attention to the location of the object from one's own or egocentric perspective this should interfere with spatial judgments when explicitly instructed to take an allocentric perspective (H2). To test this we adapted the spatial judgment task used in Tversky and Hard (2009) and Mazzarella et al. (2012). As in the original version of Tversky and Hard (2009) we presented participants with pictures depicting an actor seated at a table on which two objects (bottle and book) were positioned. Depending on the picture, only the two objects were visible (object) or the actor was visible while holding one of the objects (action). Here, we added a condition in which the actor held an object while tilting it forward suggesting a complementary request (interaction). In the first study we used an implicit instruction of perspective taking, in which participants were asked to simply specify the spatial location of the bottle relative to the book (e.g., the bottle is left from the book) without defining from which perspective. Although this study allowed us to infer whether participants would voluntarily pick a perspective depending on the type of picture, it could not answer our second hypothesis. In order to test this, a second and third (pre-registered) study was done in which participants were explicitly instructed to make the same spatial judgment from either an egocentric or allocentric perspective. Here we measured response times as a way of assessing the ease with which participants would perform the spatial judgment task from an egocentric (or allocentric) perspective while suppressing an allocentric (or egocentric) perspective.

Study 1

Method

The first study used a 2 x 2 design with action type (action vs interaction) and description (mention object vs mention action and actor) as between subject factors with object as a control condition. Pictures were used displaying an actor seated at a table making an object directed movement towards one of two objects placed in front on the table. We manipulated the description of the scene to include either references to the actor present in the scene and the action performed or only a reference to the objects on the table (see Tversky & Hard, 2009, Study 2). The description was manipulated primarily to control for different interpretations of the task instruction depending on the details provided.

Participants

For the first study, 364 university students were recruited to take part in the study in the lab including 225 female participants with an overall mean age of 22 (range = 18-65). Participants signed informed consent before starting the experiment. The study was approved by the University of Amsterdam ethics committee (2015-SP-4306).

Materials

Novel pictures were created of a man seated at a table used in a number of different ways, which formed the conditions in the task (see Fig 1). In the object condition, only a table was visible on which a water bottle was placed on the left and a book on the right. In the action condition a man was seated in the same setting while holding the water bottle with the bottle still positioned on the table. The picture in the interaction condition was the same as in the action condition but here the actor had lifted the bottle up from the table and was seen to tilt it slightly forward. For each of these three conditions participants were asked to indicate: “Relative to the book where is the bottle”. In two additional conditions, the action and interaction pictures were used but here participants were asked to indicate “Relative to the book where is the bottle he is holding” in the action act condition and “Relative to the book

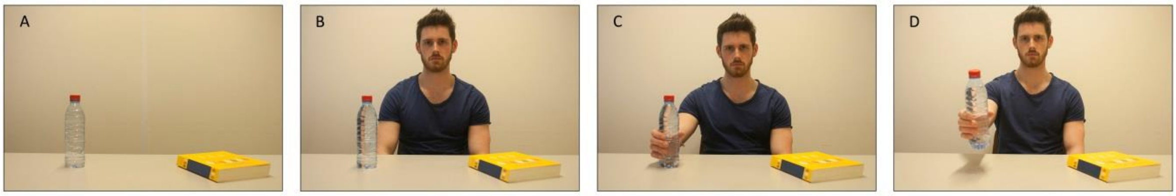


*Fig 1.*The stimuli used in all studies. (A) Object (B) Actor (C) Action (D) Interaction. (B) was used only in the second study

where is the bottle he is giving” in the interaction act condition. Note that contrary to Tversky and Hard (2009) we changed the wording of the question in the action act condition to include ‘holding’ instead of ‘placing’ given that placing might signal the end of an action whereas holding is relatively more ambiguous. Also, whereas the actor in Tversky and Hard (2009) was in fact still reaching for the object in the current setting he was holding it (see Mazzarella et al., 2012). We did not add a condition including the actor not manipulating any of the objects given that the description could not include a reference to the action being performed if no action was displayed. In all pictures the actor gazed straight forward (toward the screen) to control for differences in gaze cues.

Procedure

Along with a variety of psychological tasks, participants were presented with the current spatial judgment task. Participants were first randomly assigned to one of the five conditions: The object (n = 75), action (n = 73), interaction (n = 71), action act (n = 74) and interaction act (n = 71) condition. Pictures were presented on a computer screen and participants were allowed to write down their answer in an open format. In addition, subject ID and demographic variables were collected.

Results

Answers on the task were first categorized in terms of the perspective they implied. The bottle described as being left from the book was categorized as an egocentric response whereas the bottle seen as right from the book was categorized as allocentric. Responses that were undefined (e.g., "depends on which viewpoint you take") or included both egocentric and allocentric responses (e.g., "to the left side (from my point of view), otherwise on the right side") were categorized as neutral. For the object picture, 84 % of responses were coded as egocentric, 9.3 % as allocentric and the remaining as neutral. For the action picture, 57.5 % of responses were coded as egocentric whereas 27.4 % of responses were coded as allocentric. For the interaction picture, a total of 53.5 % responses were coded as egocentric, 21.1% as allocentric and the remaining responses as neutral. For the conditions including descriptions of the actor and action, responses were coded as (25.7 % egocentric; 40.5 % allocentric; 33.8 % neutral) for the action act and (32.4 % egocentric; 39.4 % allocentric; 28.2 % neutral) for the interaction act condition (See Fig 2).

Similar to the analysis in Furnaletto et al. (2013) we created three binary response scales using the egocentric, allocentric and neutral responses for each action type. The egocentric scale used scores of 1 if an egocentric response was made and 0 for all other responses (allocentric and neutral). Similarly, for the allocentric scale allocentric responses were scored as 1 and remaining responses as 0 while for the neutral scale neutral responses were coded as 1 and other responses as 0. Separate binary logistic regressions were performed for each response scale. For the first analysis we first ran a 2 x 2 binary logistic regression with action type (action vs interaction) and description (mention object vs mention action and actor) as between subject predictors using the egocentric response scale as DV. When including the two main effects and the interaction effect between action type and description in the model we found a significant overall model ($\chi^2(3) = 22.26$, $p < .001$). This seemed to be driven by a main effect for description (Wald $\chi^2 = 20.31$, $df = 1$, $p < .001$, odds ratio = 0.33) but not for

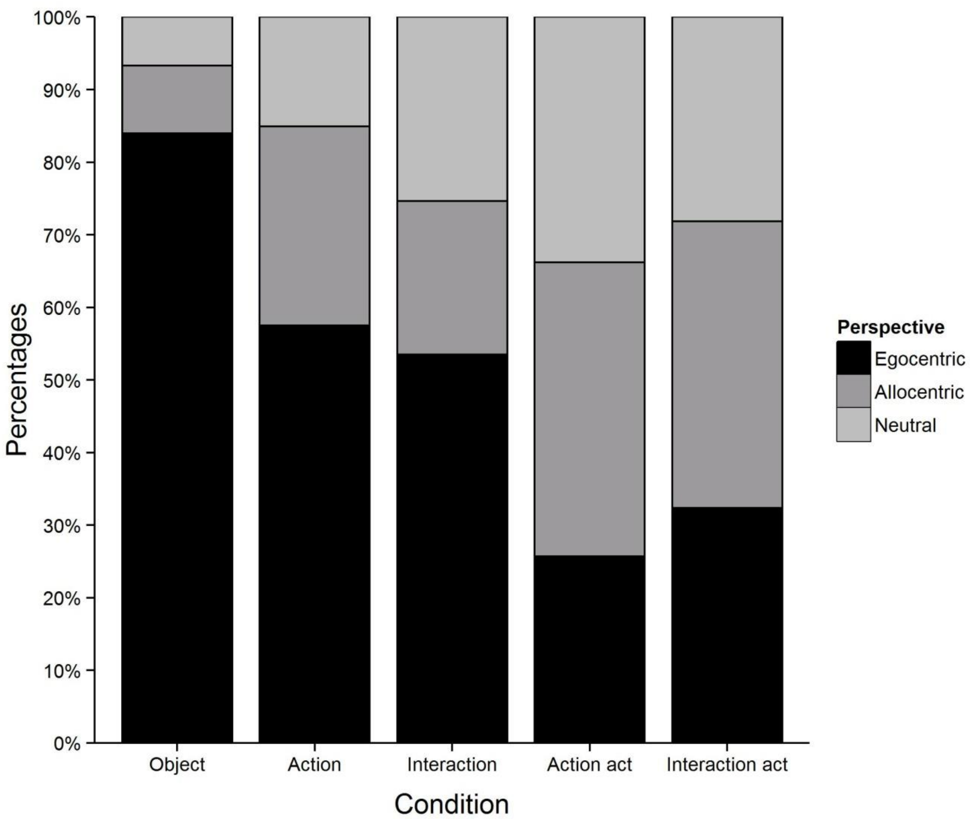


Fig 2. Percentages of each response type (egocentric, allocentric, neutral) for all conditions in Study 1. In the action and interaction conditions only the position of the objects was mentioned while in the action act and interaction act both actor and action were mentioned in the description of the picture. All graphs were created using ggplot2 (Wickham, 2009)

action type (Wald χ^2 = 0.11, *df* = 1, *p* = .741, odds ratio = 1.09) nor by an interaction between action type and description (Wald χ^2 = 0.97, *df* = 1, *p* = .325, odds ratio = 1.63). This means that across type of action (action and interaction), the percentage of egocentric responses was lower when both actor and action were mentioned (% 29.0) compared to only mentioning the objects in the scene (% 55.5). The same analysis was performed using the allocentric response scale. Following a significant model fit ($\chi^2(3)$ = 9.01, *p* = .029) again a main effect for description was found (Wald χ^2 = 8.16, *df* = 1, *p* = .004, odds ratio = 2.10) but no main effect for action type (Wald χ^2 = 0.56, *df* = 1, *p* = .453, odds ratio = 0.82) nor an interaction between action type and description (Wald χ^2 = 0.33, *p* = .567, odds ratio = 1.35), indicating a higher percentage of allocentric responses when target stimuli included descriptions of both the actor and action (% 40.0) compared to only mentioning the objects (% 24.3) irrespective of the type of action displayed. Finally, for the neutral response scale, the addition of both main effects and the interaction did not add significantly to the model fit ($\chi^2(3) = 7.45$, $df = 1$, $p = .059$) which suggests that there were no discernible differences with regards to neutral responses across conditions.

Additionally, we compared the number of egocentric responses in the object picture (control) to both action and interaction pictures separately for both description levels. This was done separately because of the unequal samples sizes between collapsed description levels and the object picture. We performed a binary logistic regression using the egocentric response scale as DV and all action type pictures (object vs action vs interact vs action act vs interaction act) as between subject predictors. Overall there was an effect of action type ($\chi^2(4) = 67.21$, $p < .001$) which was evident by a higher percentage of egocentric responses in the object condition relative to both the action condition ($\chi^2(1) = 11.82$, $p = .001$, odds ratio = 0.26), the interaction condition ($\chi^2(1) = 14.77$, $p < .001$, odds ratio = 0.22), the action act condition (χ^2 (1) = 43.55, $p < .001$, odds ratio = 0.07) and the interaction act condition ($\chi^2(1) = 35.05$, $p < .001$, odds ratio = 0.09). For the allocentric response scale a similar pattern was observed with an effect of condition ($\chi^2(4) = 27.36$, $p < .001$) showing higher percentages of allocentric responses in the action condition ($\chi^2(1) = 7.45$, $p = .006$, odds ratio = 3.67), the action act condition ($\chi^2(1) = 16.73$, $p < .001$, odds ratio = 6.62), the interaction act condition ($\chi^2(1) = 15.71$, $p < .001$, odds ratio = 6.33) but not in the interaction condition ($\chi^2(1) = 3.78$, $p = .052$, odds ratio = 2.60) relative to the object condition. Finally, a lower percentage of neutral responses was made in the object condition compared to the interaction condition ($\chi^2(1) = 8.42$, $p < .001$, odds ratio = 4.76), the action act condition ($\chi^2(1) = 14.07$, $p < .001$, odds ratio = 7.14), the interaction act condition ($\chi^2(4) = 10.22$, $p = .001$, odds ratio = 5.49), but not the action condition ($\chi^2(1) = 2.58$, $p = .108$, odds ratio = 2.48) as evident by a main effect of condition ($\chi^2(4) = 22.73$, $p < .001$).

Discussion

The results of the first study did not confirm our hypothesis (H1) in that no difference was found in the degree to which people took the perspective of the actor when comparing the action to the interaction condition. Only by adding a reference to both the actor and action in the task description did perspective taking increase relative to conditions in which the instruction referred only to the objects in the picture, which is in line with Tversky and Hard (2009). Also, compared to the object condition, the presence of an actor performing an action (any) increased the percentage of allocentric responses suggesting that settings involving manual actions increase perspective taking irrespective of the action's specific goal. Nonetheless, there were some issues with this study. First, given the relatively high rate of neutral responses across task descriptions, the task was potentially too ambiguous to provide a reliable estimate of spontaneous perspective taking. Moreover, it is argued that perspective taking differences are specifically sensitive to decisions made under time pressure or when explicitly estimating another's perspective which was not the case here (Surtees, Butterfill, & Apperley, 2012; Samson et al., 2010). The first study assessed perspective taking using an implicit instruction by only referring to the objects or actor and action in the picture without specifying the perspective participants had to take. Although this allowed us to test whether voluntary perspective taking was modulated by the type of action displayed, it did not demonstrate to what degree the situation facilitated or inhibited perspective taking when explicitly taking roles (Surtees et al., 2012; Mazzarella et al., 2012). In order to account for this, we performed a second study using an explicit instruction of perspective taking which allowed us to infer the relative ease in adopting an allocentric or egocentric perspective depending on the setting. Additionally, the second study allowed us to validate the findings in study 1 using a different measure of spatial perspective taking (RT).

Study 2

Method

In a second study we used the same setup as in the first study, but used an explicit instruction to estimate variability in perspective taking. This study had a 2 x 4 mixed design with condition (egocentric vs allocentric) as between subjects factor and action type (object vs actor vs action vs interaction) as within subjects factor. Participants in this study were instructed to either take their own perspective (egocentric) or the perspective of the actor (allocentric) providing an estimate for both direct and indirect measures of perspective taking (Surtees et al., 2012). Whereas egocentric trials reflect an indirect measure of perspective taking, where participants are instructed to take an egocentric perspective which can be affected by automatically computing another's perspective, allocentric trials reflect a direct measure where participants are instructed to take into account another's perspective while suppressing their own.

The same stimulus set was used in which participants were instructed to determine the relative position of the two objects (bottle and book) on the table. In addition we included a picture in which an actor was positioned at the table holding his arms alongside his body (see figure 1). This was done to match the conditions used in Mazzarella et al. (2012) and to provide a way to test if the mere presence of an actor would affect perspective taking. In contrast to the first study, participants were presented with all different pictures and responded by pressing buttons on a keyboard indicating the position (left or right from the book) of the bottle. To measure relative facilitation or inhibition of perspective taking in different settings we measured response times for correct decisions (Surtees et al., 2012). If an interaction setting facilitates egocentric perspective taking this would lead to increased response times in the interaction setting relative to the action setting when instructed to take an allocentric perspective. Alternatively, if the interaction setting facilitates allocentric perspective taking this would lead to lower response times relative to the action setting in allocentric trials. The action setting was chosen as the primary reference category given that this setting was hypothesized as the category that would most likely show an increase in allocentric perspective taking (Mazzarella et al., 2012).

Participants

In total, 62 participants, of which 30 female, took part in the second study with a mean age of 23 (range = 18-43). Participants first signed a written informed consent before starting the task and received course credit or monetary compensation in return for participation. The study was approved by the University of Amsterdam ethics committee (2015-SP-4306).

Procedure

After being seated in a lab cubicle participants were assigned to one of two conditions. In both conditions four different pictures were randomly presented (25.4 x 16.94 cm; in color) including the object, action and interaction pictures used in the first study as well as the newly added actor picture in which the actor was displayed seated at the table without acting on either of the objects. After each picture was presented, participants were asked to indicate if the bottle was presented to the left or right from the book by pressing either the B key (left) or H key (right) on the keyboard. Participants in the egocentric condition were instructed to determine the position of the bottle relative to the book always from their own perspective while participants in the allocentric condition were asked to determine the relative position of the bottle always from the perspective of the actor in the picture. The pictures were all horizontally flipped so that objects would appear both left and right in the pictures to make sure participants would pay close attention to the object's position. Pictures were randomly presented 20 times for each different picture (10 with the bottle left and 10 with the bottle right) creating a total of 80 trials. Before starting the main task participants received a short instruction and completed a practice run including 10 trials.

Results

Data for two participants was removed due to missing response times for the object picture as well as data for one subject for which response times exceeded more than 5 x *SD* above the mean of each picture (including this subject in the analysis did not alter the results). Two groups including 29 participants in the allocentric and 30 participants in the egocentric condition were used in the final analysis.

First off, the response times for correct trials were selected and used for the main analyses. Responses that exceeded 2.5 x *SD* above and below the mean response time per individual were filtered out. Since the RT data was not normally distributed across all conditions they were log transformed before analysis (means and SDs represent untransformed estimates; statistical findings were equal using transformed and untransformed scores). For the main analysis, we ran a 2 x 4 mixed design with perspective (egocentric vs allocentric) as between subjects factor and action type (object vs actor vs action vs interaction) as within subjects factor (see Fig 3). We found a main effect found for perspective showing faster response times in the egocentric (*M* = 492.67, *SD* = 220.34) relative to the allocentric condition (*M* = 856.58, *SD* = 363.09), $F(1, 57) = 36.09$, $p<.001$, $\eta^2_p = .39$. There was also a main effect for action type indicating different response times for different types of actions, $F(2.26, 128.71) = 6.20$, $p = .002$, $\eta^2_p = .10$. Due to a violation of the sphericity assumption ($\chi^2(5) = 35.95$, $p < .001$) we report Greenhouse-Geisser corrected values here ($\varepsilon < .75$). The interaction effect between perspective and action type was not significant, $F(2.26, 128.71) = 1.16$, $p = .323$ (for untransformed RT: $F(1.93, 109.88) = 2.65$, $p = .077$), so therefore we first analyzed post-hoc comparisons averaged over perspective conditions. Results from the first comparison showed faster response times for the action picture (*M* = 645.28, *SD* = 320.98) relative to the interaction picture (*M* = 679.21, *SD* = 347.26), $t(58) = -3.19$, $p = .002$, $dz = 0.42$. No difference was found between response times for the actor picture (*M* = 635.10, *SD* = 280.90) and the action picture ($t<1$) while response times for the actor picture were faster compared to the object picture, (*M* = 726.57, *SD* = 449.66), $t(58) = 2.81$, $p = .007$, $dz = 0.37$.

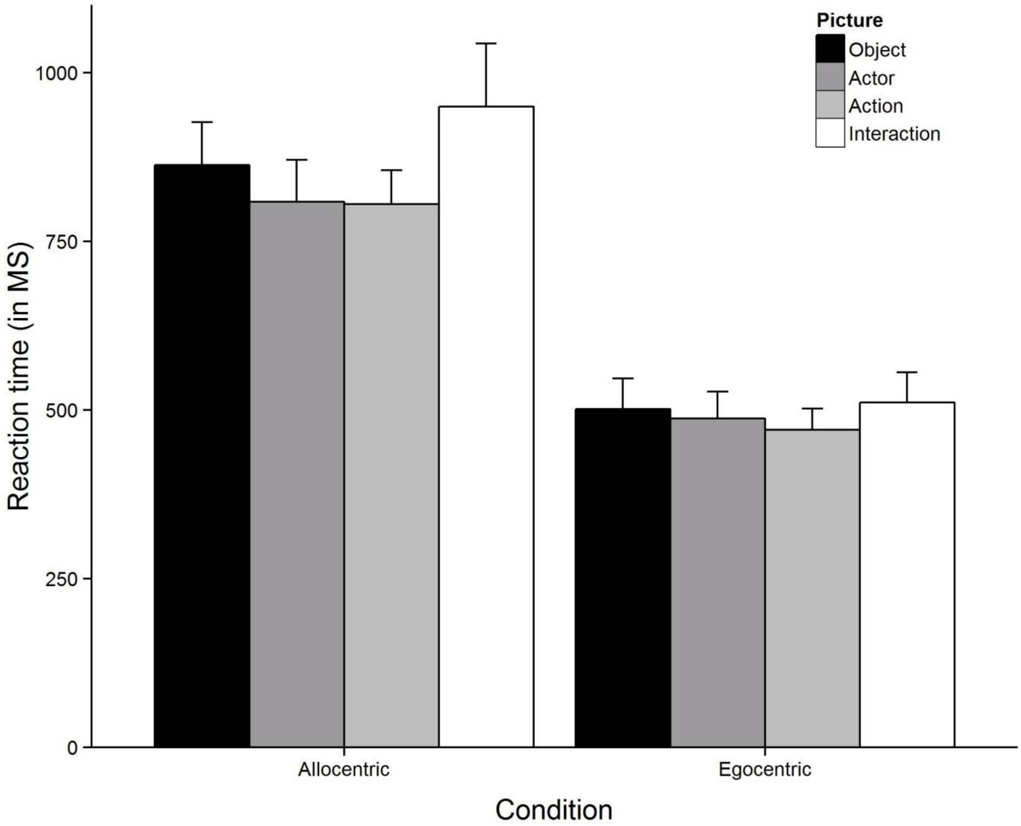


Fig 3. Mean reaction times from Study 2 for each action type in the Allocentric (Left) and Egocentric (Right) conditions separately. Bars represent from black to white: Object, Actor, Action and Interaction pictures. Error bars reflect 1 x SE

Additionally, we analyzed response times between the interaction and action picture separately by perspective condition in order to directly test our primary hypothesis (H2). In the allocentric condition, response times for action pictures (M = 808.49, SD = 332.98) were faster than interaction pictures (M = 863.20, SD = 342.46), $t(28) = 2.99$, $p = .006$, $d_z = 0.57$ (for untransformed RT: $t(28) = 2.63$, $p = .014$, $d_z= 0.50$), while no difference between action and interaction pictures was found in egocentric trials, $t(29) = 1.36$, $p = .184$. This result supports the idea of egocentric interference or the tendency to respond from an egocentric perspective when instructed to take an allocentric perspective.

Discussion

Results of the second study are in accordance with our second hypothesis (H2). The data demonstrated a stronger egocentric interference for spatial judgments in interaction settings relative to settings where an action was displayed for participants instructed to take an actor's perspective. This suggests that in interactive settings, participants had more difficulty to suppress their egocentric perspective when instructed to take an allocentric perspective. Interestingly, no response time difference was found when comparing action to actor conditions which is in contrast with earlier findings that have shown that the presence of an actor performing an object-directed movement triggers perspective taking relative to an actor merely being present (Mazzarella et al., 2012). However, there are a number of important issues that we propose to resolve with a third pre-registered study.

Pre-registration Study 3

The absence of an interaction effect is problematic and makes it difficult to interpret the egocentric interference effect. Primarily, interpreting simple main effects following a non-significant interaction is inappropriate according to the idea that the difference between a significant difference and a non-significant difference is not itself significant (Nieuwenhuis, Forstmann, & Wagenmakers, 2011). Unclear is whether this is due to the absence of differences in the interaction between action and interaction pictures or due to the similarity of the remaining pictures (object and actor) creating a similar response time pattern across between-subject conditions. The variability in response times in both egocentric and allocentric trials could be a function of the specific task instructions participants received. Whereas in Study 2 participants were asked to choose the correct location of the bottle by choosing either 'Right' or 'Left' as answer, Samson et al. (2010) as well as Surtees et al. (2012) instructed participants not only the perspective they had to take but also provided the (spatial) location of the target object. Rather than choosing the correct location as was done in Study 2, participants in their studies had to decide if the provided location (e.g., Right) was correct or not from the instructed perspective (e.g., Egocentric).

We propose two changes to the design in Study 2. Instead of using a between-subjects design we want to run the study using a full within-subjects design, in line with similar tests of perspective taking (Samson et al., 2010; Surtees et al., 2012). This could prevent participants from using specific spatial judgment 'rules' depending on the type of condition they are in what could increase variation in response latencies for the different pictures. Secondly, we want to change the task instruction so that instead of judging the position of the bottle as being either left or right from the book, participants will receive information about which perspective they need to take (allocentric or egocentric) as well as the location of the bottle (left or right) relative to the book. In this way, participants judge the correct location by responding with 'Yes' or 'No' if the bottle is in fact in the right location as seen from the instructed perspective.

Method

The proposed study (S3; pre-registered at https://osf.io/4tyxg/) will use the same stimuli as in the previous studies in a within-subjects design with perspective (egocentric vs allocentric) and picture (actor, vs action, vs interaction) as within-subject factors. Each participant will perform a response time task in which each trial will start with a perspective instruction presented in the center of the screen either depicting the word "HIJ" (he) or "JIJ" (you) for 750 ms, representing the allocentric and egocentric perspective respectively. Following the perspective instruction either the word "RECHTS" (right) or "LINKS" (left) will be displayed 500 ms later in the same position on the screen. After displaying the location for 750 ms, the target picture will be displayed after a delay of 500 ms. For each picture, participants will be asked to judge if the bottle is in the correct position as specified by the location word from the instructed perspective (e.g., is the bottle on the right from the book from your perspective). Randomly, the picture will depict an actor seated at the table on which the two objects are placed (actor), the actor grabbing the bottle still positioned on the table (action) or the actor handing the bottle towards the screen/ participant (interaction). The object picture will not be used here given that there is no possibility to take the perceptive of an actor if he is not present in the picture, something which was not clearly specified in Study 2. Only trials will be analyzed in which a correct (spatial) location is provided. In total each picture type for which a correct location is given (e.g., bottle appearing left when YOU and LEFT are specified) will be presented 20 times including both left and right locations creating a total of 240 trials counting incorrect location trials as well. Response times will be recorded from the onset of the target picture until a button is pressed.

Hypotheses and Analyses

For the analysis we will look at response times using a repeated measures ANOVA with perspective (egocentric vs allocentric) and action type (actor vs action vs interaction) as within-subject factors (action type is used instead of 'picture' which is used in the OSF document). As in the second study (Study 2) we will look at the interaction effect between perspective and action type. Although we specifically hypothesize an egocentric interference effect we still need to show that the pattern in response times is different between allocentric and egocentric trials. Provided there is an interaction effect between perspective and picture we expect a difference in response times between the action and interaction pictures in the allocentric perspective trials such that response times for correct decisions are expected to be larger for interaction pictures compared to action pictures in line with an egocentric interference effect. This would imply that interaction pictures trigger more egocentric interference compared to action trials given that interaction pictures require observers to take a first-person perspective. In egocentric perspective trials we expect no difference in response times between interaction and action pictures. This would confirm our findings in Study 2 as well as support findings showing an absence of an allocentric interference effect in egocentric perspective trials (Surtees et al., 2012). Follow-up tests will be performed using paired sample t-tests. No difference is expected between action and actor conditions in terms of mean response times in line with the findings in Study 2. Before starting the main analysis response time data will be cleaned by removing trials in which subjects respond 2.5 x SD above or below each subject mean response time (across picture type).

Power analysis

In order to test the proposed hypotheses we calculated the necessary sample size to obtain 80 % power using a within-subjects design. Power can be calculated based on the results from Study 1 and 2 but this can be partly problematic. First off, since we plan to rerun the paradigm tested in Study 2 rather than Study 1 we have to perform a power analysis based on a single effect size which is not sufficient (Hoenig & Heisey, 2001). Secondly, we would estimate power on the basis of a simple main effect following a non-significant interaction (in Study 2). We therefore tried to select related studies to estimate a reliable effect size. A number of papers have looked at the effect of changing angular disparity (the angular position at which an actor appears relative to the observer) on response times in spatial judgment tasks (Michelon & Zacks, 2006; Kessler & Thomson, 2010; Surtees et al., 2013). Typically these effects are quite strong and explain around +- 60-80 percent of the variance in judgment tasks (conditional on other effects studied in the specific design). This is relevant here given that spatial perspective taking as we found in our task (Study 2) similarly requires embodied transformation in the case of allocentric trials (i.e., putting yourself in the position of the actor). Although this is evident as a main effect in Study 2 as well it does not apply to the comparison between action and interaction pictures in allocentric perspective trials. The proposed changes in the design only match (reasonably) closely to the visual perspective taking task used in Surtees et al. (2012) in which the effect size for the interaction effect was

η^2 = .181 whereas the simple main effect for the critical (ambiguous) stimuli was d_z = 1.02. The simple main effect represents the hypothesized egocentric interference effect whereas the interaction effect is required to judge the interference as unique to the allocentric perspective trials. One option is to choose a mean estimate based on our own results and the study by Surtees et al. (2012) with a resulting mean effect size of d_z = 0.66 for the simple main effect and η^2 = .095 for the interaction. Using the most conservative estimate we choose to run 47 participants for a full within design (computed using GPower with a test of dependent means of d_z = 0.42).

Participants

As planned, 47 participants, of which 37 female, took part in the third study with a mean age of 25 (range = 18-59). Two of the 47 participants were left-handed. Participants first signed a written informed consent and received course credit or monetary compensation in return for participation. The study was approved by the University of Amsterdam ethics committee (2016-SP-6751).

Confirmatory analysis

For the analysis we ran the proposed within-subjects analysis with perspective (egocentric vs allocentric) and action type (actor vs action vs interaction) as within subject factors. We had to remove two participants who made over 106 errors in correct trials making it impossible to compute averages over all action type trials so data for two additional participants was collected. Response errors for the remaining participants were removed prior to the analysis phase. Also, we removed 1.4 % of total trials due to response times that were above the 2.5 x SD threshold relative to individual response times. As proposed, for the reaction time analysis we looked at trials in which the correct object location was specified (e.g., object appearing left when the object was described as being left from an egocentric perspective).

A main effect was found for perspective showing higher response times for allocentric (*M* = 1051.23, *SD* = 466.40) relative to egocentric trials (*M* = 859.65, *SD* = 389.96, $F(1, 46) = 41.41$, $p < .001$, η^2 = .47. Also, a main effect for action type was found, $F(1.56, 71.78) = 3.76$, $p = .037$, η^2 = .08, but not the critical interaction between perspective and action type, $F(1.66, 76.57) = 0.14$, $p = .836$ (See Fig 4). Across egocentric and allocentric trials, longer response times were found for interaction pictures (*M* = 985.23, *SD* = 419.68) compared to action pictures (*M* = 923.51, *SD* = 397.30), *t*(46) = 2.41, *p* = .020, d_z = 0.35, Hedges g_{av} = 0.15, as well as for actor (*M* = 957.58, *SD* = 408.22) compared to action pictures, *t*(46) = 2.20, *p* = .033, d_z = 0.32, Hedges g_{av} = .08.

Exploratory analysis

When exploring paired differences within each perspective condition in line with the second study we again found a significant positive difference for interaction (*M* = 1083.02, *SD* = 457.98) relative to action pictures (*M* = 1011.95, *SD* = 457.93) in allocentric trials, *t*(46) = 2.75, *p* = .008, d_z = 0.40, Hedges g_{av} = 0.15, but not in egocentric trials, *t*(46) = 1.19, *p* = .241 (controlling for multiple comparisons α / 2 = .025). In addition, we looked at response errors in correct trials (e.g., object said to appear left and appearing left from an egocentric perspective but judged as incorrect), using a repeated measures analysis with perspective (egocentric vs allocentric) and action type (action vs interaction) as within subject factors. We found a main effect of perspective across levels of action type, $F(1, 46) = 5.57$, $p = .023$, η^2 = .11, while no main effect for action type or an interaction between perspective and action type was found (both *p*s < .79). More errors were made in allocentric trials (*M* = 1.94, *SD* =2.76) compared to egocentric trials (*M* = 1.43, *SD* = 2.25). Since this pattern of results could be a function of a speed-accuracy tradeoff we computed inverse efficiency scores (IES; Bruyer & Brysbaert, 2011) and subsequently used the IES scores for the main repeated measures analysis but this did not affect our results.

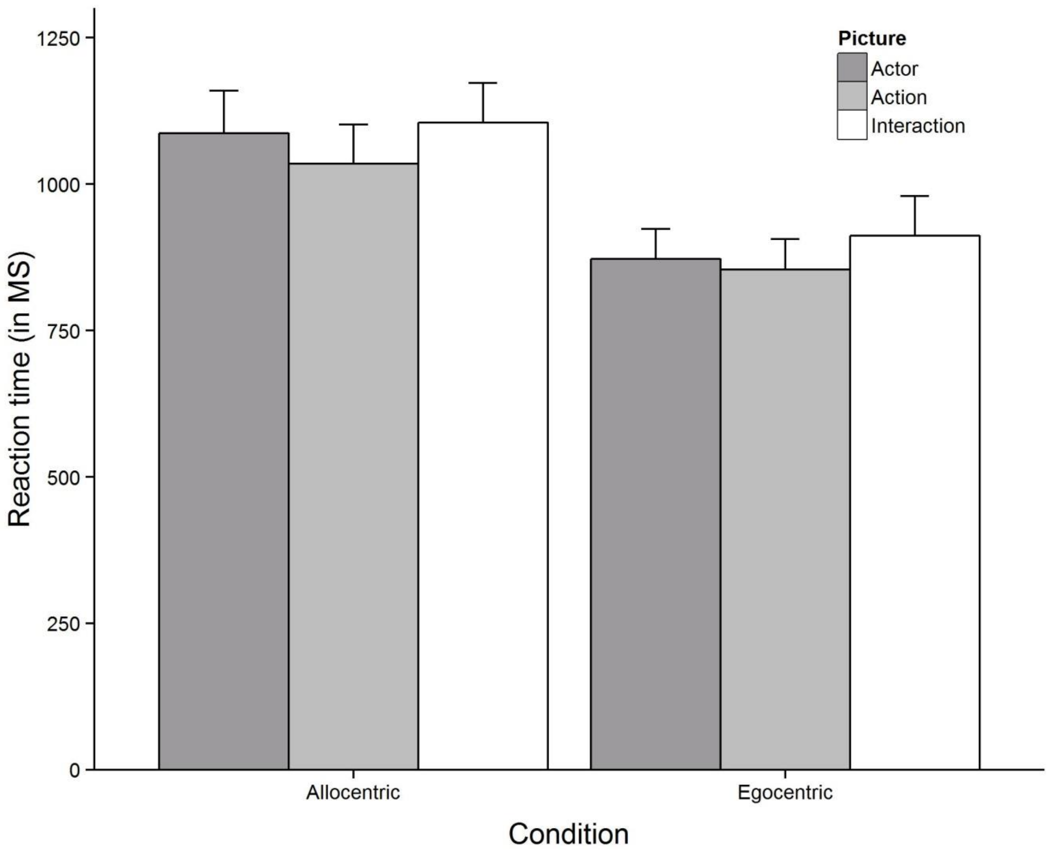


Fig 4. Mean reaction times from Study 3 for each action type in the Allocentric (Left) and Egocentric (Right) conditions separately. Bars represent from dark grey to white: Actor, Action and Interaction pictures. Error bars reflect 1 x SE

Discussion

The results of the third study are somewhat in line with the second study in that we found a main effect for both perspective and action type but no interaction between the two factors. Collapsed across levels of perspective we found longer response times for interaction pictures compared to action pictures suggesting that interactive settings increase the cost of making spatial judgments. Even though this increase in response times was only found in allocentric trials and not in egocentric trials which is consistent with Study 2 and with our hypothesis (H2), it is not possible to state that the difference between these paired comparisons is itself significant. Contrary to Study 2 we found a difference between actor and action pictures across perspectives suggesting that pictures of an actor grasping and holding an object facilitates spatial judgments in both egocentric and allocentric trials.

General discussion

The presented findings provide mixed support at best for our chosen operationalization of perspective switching in interactive settings. In the first study we found no difference in the degree to which participants took the actor's perspective when comparing action to interaction settings. However, in line with the findings in Tversky and Hard (2009), we did find an increase in allocentric responses when the description of the stimulus mentioned both the action and actor present in the scene relative to descriptions only about the objects on the table. Subsequently, in the second and third study we found a stronger egocentric interference effect for participants instructed to take the actor's perspective for interaction relative to action settings but the absence of a difference in this comparison across perspective instructions make this a problematic interpretation. As a result we cannot support the hypothesized idea of effective perspective switching in which (social) complementary requests trigger egocentric responding rather than increase (allocentric) perspective taking. These results partly mirror the findings in Surtees et al. (2012) who found evidence of level-2 perspective taking only using direct measures (allocentric instructions) and not using indirect measures (egocentric instructions). Similar to their study, we found no response time differences in the egocentric condition across Study 2 and 3 which could be due to an absence of (allocentric) interference when judging the spatial position of stimuli from an egocentric perspective. Otherwise, the fact that egocentric responses form the default response for an observer may have created a floor effect in response times across action types specifically in Study 2, eliminating any differences due to the type of action displayed by the actor. Resolving this by using a slightly different design in Study 3 in which perspective instructions and action types were intermixed did not change the overall response pattern.

An important point is the apparent difference in measurements used here. Whereas letting participants voluntary choose a perspective did not affect differences in perspective taking when comparing action to interaction settings, a different pattern was found when participants were explicitly instructed to take either one's own or another's perspective. Importantly, measures of instruction (explicit vs implicit) were 'confounded' with the level of measurement (open question = explicit; response time = implicit). This was a necessary attribute of the design given the high response rate of egocentric answers by default and the difficulty of evaluating implicit instructions using response time measures. This differs somewhat from Mazzarella et al. (2012) who found a similar pattern in allocentric vs egocentric responses using both implicit and explicit instructions while we found a different pattern depending on the specific instruction used. However, in their studies, both the proportion of egocentric and allocentric responses and the timing of the response (RT) were measured in each study. While a consistent pattern was found for proportion of responses, this was not the case for the RT estimates. We provide mixed support for the response proportion findings in Mazzarella et al. (2012), given that we found no differences in response times when comparing actor to action settings in Study 2 but to some extent in Study 3. We believe response time measurements are more useful here as pointed out earlier in reference to Surtees et al. (2012) that egocentric errors are most likely made when performing tasks quickly. Moreover, given that participants in Study 1 could voluntary choose either perspective as valid response, this might have increased the ambiguity of the task. As noted earlier, a reasonable number of participants in the interaction condition in the first study (approx. 25-28 % depending on the description condition) were uncertain about the goal of the task which was clear from several responses (e.g., “for the actor left, for me right”; “from my perspective, left of the book, from his perspective, right from the book”).

Besides the discrepancy in outcomes for the first compared to the last two studies, there is a problem with the response pattern for different action types in Study 2 and 3. The absence of an interaction between perspective and action type suggests that the type of action affected response times irrespective of the instructed perspective. It is possible that stimulus features such as showing a hand grasping an object provided an additional (low-level) spatial cue relative to seeing the actor sitting at the table not reaching for the bottle. Considering that this spatial cue would primarily facilitate responding from an egocentric perspective but could at the same time inhibit responding from an allocentric perspective this likely is not an issue given the overall response pattern. Moreover, since in both action and interaction settings the actor was grasping and either holding or handing over the object this additional hand cue would be equal for these two settings. Given that across Study 2 and 3 differences were found in response times there must be another reason that produced differences in response times not represented in stimulus features. Overall, it seems that changing the design in Study 2 from a mixed to a within-subjects design in Study 3 as well changing the task instructions did affect task difficulty. Compared to Study 2, mean response times in allocentric trials (across pictures) were around 195 ms higher in Study 3 while mean response times in egocentric trials (across pictures) were around 367 ms higher. However, consistent with Samson et al. (2010), using a separate or intermixed presentation did not seem to affect the overall response time pattern besides the difference between action and actor pictures in Study 3 but the absence of this difference in Study 2. One related issue concerns the specification of the correct or incorrect location prior to showing the target picture (in Study 3), which allowed participants to anticipate the correct location prior to presenting the target stimulus. In fact, a number of participants in Study 3 indicated they had used this technique in order to improve response performance. Nonetheless, a similar task setup was successfully used to test perspective interference effects in both Samson et al. (2010) and Surtees et al. (2012) and it is unlikely that this technique had a significant influence given the overall higher response times in allocentric trials. Finally, one possibility is that interaction settings were more difficult to process regardless of the instructed perspective. Perhaps due to the increased (social) relevance of the situation, this distracted the participant when judging the spatial location of the object rather than triggered the participant to pay better attention to the object location from an egocentric perspective. At least from the error rates this seems unlikely since more errors were made in allocentric relative to egocentric trials irrespective of the action type which suggests a general increase in difficulty due to embodied transformation.

Nonetheless, in most studies that have used complementary and interactive action displays (van Schie, Waterschoot, & Bekkering, 2008; Ocampo & Kritikos, 2010; Faber, van Elk, & Jonas, 2016) the task is to manually respond to the observed action rather than to judge the spatial location of the action or the object related to it. Arguably, an intrinsic part of the decision to respond to others is to identify the spatial location of relevant objects, and given the absence of an explicit (manual) task this might make judging the spatial location less relevant. Alternatively, the task to judge the spatial location from a certain perspective, which can be done without relying on the actor's behavior in any way, might have interfered with the tendency to respond to the complementary request which slowed responses across conditions. Some recent evidence suggests that it's the interdependency between actor and observer that drives errors in perspective taking. Using a physical interactive perspective taking task, Elekes, Virga, and Király (2016) found that level-2 perspective taking was affected specifically when the task of the interacting partner overlapped with the task the participant had to perform compared when both tasks were different. Whereas this 'mutual awareness' was known to the participant, no such interdependency was implied in our spatial judgment task which might be a useful factor in future studies.

As proposed in Tversky and Hard (2009), allocentric perspective taking can increase when observing others in order to better anticipate their actions even when no behavioral response is necessary or immediate. Also, as argued in Zwickel and Müller (2010), perspective taking might be driven by the observed relevance of the actor's mental state possibly in order to prepare subsequent responses. Rather than an increase in perspective taking, here we tried to demonstrate that perspective taking decreases or is inhibited when responding is facilitated by taking an egocentric rather than an allocentric perspective. Given that we only used pictures suggesting an interaction, which only slightly differed from pictures displaying object-directed actions, the mean deviation in response times is remarkable. Nonetheless, we cannot conclude that the increase in response times in interaction pictures is a function of egocentric interference or interferes with embodied transformation per sé. Whether perspective switching is therefore relevant to dissociate perspective taking in different points in time during social interactions remains unclear at this point. Still, we have provided a well-powered demonstration of spatial perspective taking in social interactions and have shown that by means of minimally changing a passive into an interactive display we can affect response interference in spatial perspective taking. We believe that extending the research line by trying to identify additional boundaries for our effects in an explicit task setting should ultimately provide valuable information.

Compared to the previous two empirical chapters the current chapter extends the first two by showing how context dependent learning affects responses that either reflect imitative vs complementary actions. In all of the three chapters either reaction time or visual attention measures were used. This allowed us to infer both the role of overt behavioral responses as well as the attention processes leading up to these responses. Nonetheless, additional, neuropsychological processes precede and regulate attention as well as overt behavior that can help to clarify the full path leading from perception to action. Accordingly, in the next chapter we will use EEG measurements to see what happens prior to changes in attention and overt action performance. On the downside, this further restricts the type of stimuli and task setup given the measurement sensitivity of EEG but we can use it to better understand where action sequences originate and whether they reflect a passive (only responding to the environment) or active (predicting upcoming stimuli) mechanism. Our goal was therefore to use EEG to see how contextual cues can inform an observer about incoming sensory information which can be applied to sequential task perception (Chapter 2), response preparation (Chapter 3) as well as perspective switching (Chapter 4).

Chapter 5

Dissociating predictive from online motor activation: an EEG study

This chapter is based on: Faber, T. W., Bekkering, H., Jonas, K. J., & van Elk, M. (2016). *Dissociating predictive from online motor activation: an EEG study*. Manuscript in preparation.

Abstract

Currently the precise functional role of motor activation during action observation is unclear. Motor activation may primarily subserve the prediction of upcoming action goals, or alternatively: motor activation may be used to mirror or understand ongoing actions. In this study we aim to disentangle these two possible mechanisms, by focusing on mu- and beta-oscillations - as a measure of motor activity - during anticipation and observation of high and low predictable actions. Our goal was to see whether an increase in motor activity for predictable (unpredictable) actions prior to observation leads to reduced (increased) motor activity during observation. We found a stronger decrease in beta-power, but not in mu, prior to action observation for high relative to low predictable actions. No differences in beta- or mu-power between high and low predictable were found during observation. The findings suggest that motor activity - as measured with beta-oscillations – is primarily involved in action prediction, potentially reflecting a detailed simulation of the upcoming action.

Introduction

During the last two decades, accumulating evidence has shown that the brain areas involved in our own goal-directed actions, are also activated when observing actions performed by others. Motor involvement during action observation has been suggested to support action understanding, by inferring action goals on the basis of the perceived kinematics (Rizzolatti, Fogassi, & Gallese, 2001; Rizzolatti & Craighero, 2004; but see Hickock, 2009). An alternative proposal suggests that motor activation primarily involves a simulation of future actions, which follows rather than produces goal inference (Csibra, 2008; Southgate, Johnson, Osborne, & Csibra, 2009; Urgesi et al., 2010; Wilson & Knoblich, 2005). Evidence to support the latter proposal has shown that, when upcoming actions are predictable, motor activity can be measured prior to action onset (Kilner, Vargas, Duval, Blakemore, and Sirigu, 2004; Southgate et al., 2009; Southgate, Johnson, El Karoui, & Csibra, 2010; Umiltà et al., 2001). One potential source for the divergence in views is that most studies have either focused selectively on the anticipation or the observation phase of actions. This raises the question: If an action can be anticipated in advance does motor activation prior to action onset determine motor activation during observation?

The theoretical framework that we use to answer this question, proposes that the motor system relies on both forward and inverse models that allow an observer to anticipate the sensory consequences of one's actions and to subsequently monitor these actions in real-time (Wolpert, Doya, & Kawato, 2003; Wilson & Knoblich, 2005; Csibra, 2008; Schütz-Bosbach & Prinz, 2007; Kilner, 2011). This allows the motor system to form a loop, whereby incoming sensory information incongruent to the predicted input is used to update one's forward model (Wolpert et al., 2003). During action observation, the forward model specifies the sensory consequences based on the predicted intention of the actor, which allows the observer to deduce the goal of the action that is most likely given the predicted input (Kilner, Friston, & Frith, 2007). This implies that updating the predicted input during action observation is determined by the degree to which an action can be anticipated. For example, when an upcoming action can be anticipated there is less error between the predicted and observed action and therefore less ‘need’ to predict its development during observation. In accordance with this idea, some have argued that processing of predicted or anticipated information is computationally easier (Wilson & Knoblich, 2005) or is suppressed (Kilner et al., 2007; Friston, 2010) relative to unpredicted information. However, such arguments are mostly based on research looking at the visual processing of predicted events (e.g., Rao & Ballard, 1999) while there is limited empirical evidence showing how anticipated actions subsequently affect motor activation during observation. To get a complete picture of the role of forward models in action observation it is important to compare settings in which action outcomes can be anticipated (Kilner et al., 2004) and those where outcome information is lacking as is common in action observation studies that use random trial presentation (e.g., Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005).

In this EEG study, we used a pre-cueing task in which participants passively observed videos depicting object-directed hand movements (either a whole hand grasp or a precision grip) in which a color cue determined the predictability of subsequently observed actions (see Figure 1 and Supplementary Material Online for example stimuli). The cue was either highly predictable in that it was followed by one action type in 70% of trials (e.g., whole hand grasp) or low in predictability in that two actions were equally likely to follow the cue (i.e., either a whole hand grasp or a precision grip). In high predictable trials, the remaining 30% of trials showed an unexpected action (e.g., precision grip), which differed from the predicted action (e.g., whole hand grasp). Finally, an additional cue fully predicted a video in which no hand movement was made (the hand remained idle) which served as a control condition. Analyses focused on changes in mu- (sensorimotor alpha) and beta-oscillations as well as ERPs (e.g., the lateralized readiness potential and the contingent negative variation), which have been found to reflect motor activity both prior (Kilner et al., 2004; Southgate et al., 2009; Kilner, Bott, & Posada, 2005; Tzagarakis, Ince, Leuthold & Pellizer, 2010) and during action observation (Hari, 2006; Stapel, Hunnius, van Elk, & Bekkering, 2010; Kilner, Baker, Salenius, Hari, & Lemon, 2000; Koelewijn, van Schie, Bekkering, Oostenveld, & Jensen, 2008).

Three hypotheses were tested by looking at mu and beta-power, the squared amplitude over oscillations in each frequency-band, prior to action onset (the anticipation phase) as well as during action observation. First, conform the idea that motor activation primarily reflects action prediction, in which actions are anticipated if predictable and where observing actions reflects monitoring of (un-)predicted outcomes, anticipating high compared to low predictable actions would result in a stronger decrease in mu and beta-power prior to action onset (Southgate et al., 2009; Tzagarakis et al., 2010; Zaepffel, Trachel, Kilavik, & Brochier, 2013). Subsequently, a stronger decrease in power would be expected when observing low relative to high predictable actions (Csibra, 2008; Stapel et al., 2010; Hypothesis 1). Secondly, we expected a decrease in power for high but not low predictable action sequences relative to sequences in which no action (no movement) was displayed in both temporal phases (Hypothesis 2). The inclusion of a no movement condition is critical in order to dissociate anticipation of an outcome involving movement from an outcome involving no movement (see Kilner et al., 2004). Thirdly, we expected a reduction in mu and beta-power when observing unexpected actions (actions with different kinematic properties and goals) compared to actions observed in high as well as low predictable trials (Stapel et al., 2010; van Elk, van Schie, den Heuvel, & Bekkering, 2010; Hypothesis 3). Specifically, we predicted a power decrease when observing unexpected compared to low predictable outcomes given that unexpected actions are expected to occur infrequently whereas in low predictable trials actions are expected to occur equally often. This hypothesis is important given that most studies on action prediction have used unexpected outcome trials (e.g., Stapel et al., 2010) but rarely unpredictable trials even though the latter are more common in randomized task paradigms. Besides changes in time-frequency components we intended to replicate the observed lateralized readiness potential (LRP) prior to action onset over sensorimotor regions contralateral to the observed hand for high predictable actions relative to predictable no movement sequences (Kilner et al., 2004).

Method

Experimental design and procedure

Central to the design were a series of video clips depicting object-directed hand movements comparable to those used in Kilner et al. (2004). Each trial started with a fixation marker presented in the center of the screen for 2000 ms which appeared throughout the trial in the same position (overlaying all subsequent stimuli). Participants were asked to fixate on the marker during the entire trial in order to prevent eye movements. Following the fixation marker, a still image was presented depicting a right hand seen from an egocentric perspective positioned on a table with in front of it a circular object which afforded both a precision grip and a whole hand grasp (see Fig 1). Although the motor preparation effect in Kilner et al. was found irrespective of perspective (egocentric or allocentric) other studies have shown preferential processing of actions shown from an egocentric compared to allocentric perspective (Jackson, Meltzoff, & Decety, 2006; Oosterhof, Tipper, & Downing, 2012) and therefore we decided to use an egocentric perspective as well. After the image was presented for 1000 ms the hand changed color by applying a color layer (medium opacity) to the hand in the image for another 1000 ms (color cue). Rather than changing the color of the object (Kilner et al., 2004) we changed the color of the hand to increase attention to the kinematics of the upcoming hand movement. After the color cue the still image reappeared (i.e., the color disappeared) for 1000 ms before a video (25 fps) was presented for 1520 ms showing either the hand reaching and grabbing the smaller tube extending from the upper part of the object (precision grip) or reaching and grasping the full object at the lower part (whole hand grasp). To prevent that small differences in the initial hand position could be used to predict the upcoming action we created four different videos for each action outcome.

Still images were created by choosing the first frame of each video. For the color cues, four color layers were used each representing one of the stimulus conditions (see Table 2). A green colored hand (cue 1) cued a precision grip in 70 % of trials while in the remaining 30 % of trials the object was picked up with a whole hand grasp (high predictable trial). When a blue colored hand (cue 2) was shown the reverse probabilities applied (70 % whole hand grasp outcome; 30 % precision grip outcome). In low predictable trials, a red colored hand (cue 3) was presented, cueing either a precision grip or a whole hand grasp with equal probability (both 50 % of trials). In the final control condition a yellow colored hand (cue 4) appeared which remained stationary in 100 % of trials (no movement; see Kilner et al., 2004). Colors were pseudo-randomly matched to their corresponding outcome probabilities between participants to prevent systematic color differences between our experimental conditions of interest.

In a practice block prior to the main task, participants were informed about the color cues and their corresponding action outcomes (e.g., "a blue colored hand will be followed by a whole hand grasp in most cases") and were shown each trial in a short instruction task. To make sure participants would pay adequate attention in the main task, a number of catch trials was added in which trials were cut before the action onset following which participants were asked to predict the upcoming action. Participants could either answer this question by choosing one of three options: "Precision grip", "Whole hand grasp" or "I don't know/ No movement" by pressing one of three buttons on the keyboard. Crucially, the unique button corresponding to each answer option was changed randomly for each catch trial to prevent response preparation prior to the display of the question. Every 10 trials a 10 second break was presented allowing participants to make eye blinks while every 40 trials a longer break was presented for 1 minute. In total 8 blocks were presented containing 40 trials each summing up to a total of 320 trials with an additional 64 catch trials evenly spread across blocks and randomly within each block (8 trials per block). All target trials were presented pseudo-randomly to prevent unique trial repetition and to distribute a consistent amount of trial types per block. The task was programmed and presented using Presentation (Neurobehavioral Systems, Inc., Berkeley, California). Following the main task an exit questionnaire was presented in which participants were asked about their strategy in answering low predictable trials (e.g., guessing the outcome) and their thoughts about the goal of the experiment.

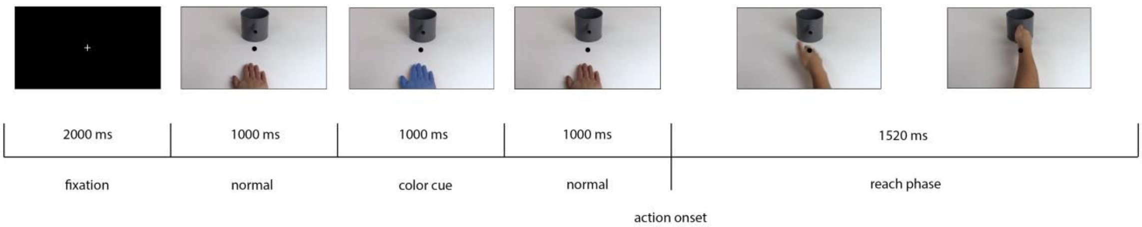


Fig 1. Trial design. Shown here is a trial sequence in which a blue color cue is followed by a precision grip. For each type of action outcome (i.e., precision vs. whole hand grip) four different clips were used.

Table 1. Trial conditions

Condition	Outcome (nr of trials)	
Cue 1 (HP)	70 % precision (56)	30 % whole hand (24)
Cue 2 (HP)	30 % precision (24)	70 % whole hand (56)
Cue 3 (LP)	50 % precision (40)	50 % whole hand (40)
Cue 4 (PN)	100 % no-movement (80)	

Table 2. Trial conditions including number of trials for each cue type in percentages and absolute number (between brackets). HP = high predictable trials; LP = low predictable trials; NP = no movement trials. The 30% outcomes represent the unexpected trials (UP).

Electrophysiological recordings

We used a BioSemi 64 EEG-headcap with a 64 electrode setup placed according to the 10-10 system with reference electrodes placed on the left and right mastoids. Data was recorded using the BioSemi active-electrode system (BioSemi inc., Amsterdam, the Netherlands), which was amplified and processed using the BioSemi ActiveTwo system (DC coupled). Data was recorded at a sampling rate of 2048 Hz. All offline analyses were performed using the Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) in Matlab (MathWorks Inc.). In offline preprocessing, electrodes were re-referenced to the common average rather than the mastoid channels due to excessive noise on mastoid channels for a number of participants. We subsequently applied a bandpass filter between 0.1 and 100 Hz as well as a DFT filter for 50 Hz to remove line noise from the re-referenced data.

For data analysis, segments were created around the video sequence starting from 2800 ms prior to action onset, to 1800 ms after action onset. Taking a long segment starting from cue onset to action onset allowed us to inspect the full sequence in which participants had (partial or full) information about upcoming actions. Also, the time windows of action anticipation and observation were comparable to previous studies looking at either mu-power or ERP-indices of motor activation (Kilner et al., 2004; Southgate et al., 2009). However, this also came at the risk of having to remove more trials due to artifacts and the subsequent loss of participants because of an insufficient number of trials per condition. After creating segments, a baseline correction was applied by subtracting the raw signal from the average across each segment (demean). Artifacts were manually checked and removed before analysis: if more than 50 % of trials was removed for any of the trial types for a particular subject this subject was not included in the final analysis. Before averaging across trials, whole hand grasp and precision grip trials were collapsed since we had no specific predictions for the type of action outcome and we wanted to retain as much trials as possible before averaging. This resulted in four stimulus conditions: high predictable actions (HP), low predictable actions (LP), predictable no movements (PN; 100 % certain of no movement) and unexpected trials in which high predictable action cues were followed by unexpected actions (UP) after action onset.

Analysis

Time-frequency data was computed by applying a Hanning taper with a fixed window length of 500 ms in steps of 50 ms (2 Hz spectral resolution) over the full segment (-2800 to 1800 ms with respect to the onset of the action). A baseline correction was applied from -2500 to 2000 ms prior to action onset in order to determine both the change in power in the anticipation phase and the subsequent change in power during action observation. This baseline was chosen given that no information about the upcoming cue or the type of action outcome was available at this point. Furthermore, this baseline allowed us to disentangle general stimulus anticipation (prior to cue onset; i.e., the cue was presented at -2000 ms relative to action onset) from movement expectancy prior to action onset (i.e., the action was initiated at 0 ms). Electrodes and frequency bands of interest were specified a priori by choosing electrodes ipsilateral (C4, CP4) and contralateral to the visible hand to in the video sequence (C3, CP3), in order to test the laterality of the power difference, within the alpha (8-14 Hz) and beta frequency range (20-30 Hz) where the decrease in power was predicted to be strongest over contralateral electrodes (Southgate et al., 2009; Pfurtscheller, Graimann, Huggins, Levine, & Schuh, 2003; Streltsova, Berchio, Gallese, & Umiltà, 2010). Sequential (F-) testing was used to compare conditions for each 50 ms time-window across the trial segment of interest (-2000 to 1500 ms relative to action onset). To control for multiple comparisons we chose a method used in van Elk, Bousardt, Bekkering and van Schie (2012). Since the full trial segment contained 71 samples (each lasting 50 ms) this means the false alarm rate (FAR) would be (1-0.95^71) = .97 when uncorrected. By selecting only segments containing at least 3 consecutive significant samples (71*(0.05^3)) the FAR was reduced to .009.

For the ERP analysis we first downsampled the raw data to 256 Hz for further analyses. We were primarily interested in the lateralized readiness potential (LRP) prior to action onset over sensorimotor areas contralateral to the observed hand (as in Kilner et al., 2004), as well as the ERP in response to the observation of unexpected actions in high predictable trials. For the analysis we used the same sequential testing procedure as for the time-frequency data. For each stimulus condition the ERP amplitude was averaged over 20 ms segments for the full trial period (from -2000 to 1500 ms relative to action onset) and we only looked at segments containing 4 or more subsequent significant samples in order to correct for multiple comparisons (176*(0.05^4) = .0011).

Below we will report the outcomes of three different analyses, to test our hypotheses. In the first analysis HP trials were directly compared to LP trials in order to see whether the predictability of upcoming actions determines the level of motor activity prior and post action onset. Secondly, we compared HP (high predictable) and LP (low predictable) to PN (no movement) trials in order to investigate whether motor processing for low and high predictable actions could be dissociated from predictable sequences that do not involve movement. Finally, HP, LP and PN trials were compared to UP trials in order to see how outcome predictability is related to processing of unexpected actions. In all analyses we report effects on mu and beta-power and ERP effects.

Results

Participants

In total 31 participants were tested for the experiment, but 13 were excluded: one participant turned out to be left-handed (Edinburgh Handedness Inventory range from 50 to 100 across participants; *SD* = 12.6) while data for the remaining participants was removed due to considerable loss of more than half of trials for at least one trial type as a consequence of excessive eye or movement artifacts as well as the noise levels on specific electrodes. The selected participants included 16 Females with a mean age of 23 years (range = 18-50).

Exit-questionnaire

In the exit-questionnaire we asked participants if they had used a strategy to anticipate action outcomes in low predictable action trials. Out of the original sample, 22 participants (13 in the final sample) said they consistently chose ‘I don’t know’ as answer in the catch trials, which suggests the majority did not bet on the action outcome after the color cue was observed. The remaining participants thought they could either detect patterns in trial sequences or (in some trials) notice differences in the hand position at the start of each video that they could use to predict upcoming actions. This seemed to be confirmed by the actual responses from the catch trials as participants chose ‘I don’t know’ in 64 % of low predictable trials while correctly choosing the action outcome in 90 % of high predictable trials. However, it is possible that these estimates were partly affected by changing the response keys on each trial.

To assess at which point in the trial the type of action outcome could be fully disambiguated, we asked two additional participants who did not take part in the experiment to choose a frame in the action video at which the action outcome (precision grip or whole hand grasp) could be predicted with certainty. The two independent raters were shown each video, without the colored action cues, and were asked to skip frame by frame and select the first frame that they thought showed clear kinematic differences predicting the action outcome (e.g., width of the grip aperture). We subsequently looked at time estimates conditional on the fact that the action outcomes were correctly predicted. Averaged across precision grip and whole hand grasp sequences we found that the outcome of the action could be anticipated 445 ms after the hand started moving for participant 1 and after 475 ms for participant 2. Inter-rater reliability was $\alpha = .91$. Mean time estimates separate for each action outcome ranged between 440 and 480 ms.

High vs low predictable actions

We first performed F-tests for each time sample with condition (HP x LP) and hemisphere (Left x Right; corresponding to electrodes C3 and CP3 = Left; C4 and CP4 = Right) as within subject factors on estimates of mu and beta-power. First off, the effects of condition and hemisphere on mu-power were not significant (max $F = 4.73$). We did however find a stronger decrease in beta-power for HP compared to LP trials prior to action onset (from -1800 to -1650 ms; from -1350 to -1150 ms; from -950 to -800 ms; $F(1, 17) \geq 4.91$, $p < .05$) across the right and left hemisphere (see Fig 2C; See additional analyses in the Supplementary Material).

Movement vs No movement

Next we ran the same analysis including HP and PN trials. In terms of mu-power, a main effect was found for condition, reflected by a decrease in power for HP compared to PN trials after action onset (400 to 1500 ms; $F \geq 5.84$, $p < .05$)[4]. In the beta range we found a decrease in power for HP compared to PN trials prior to action onset (from -850 to -550 ms and -350 to -150 ms; $F(1, 17) \geq 4.56$, $p < .05$) as well as post action onset (from 250 to 1300 ms; $F(1, 17) \geq 4.64$, $p < .05$). In addition, we found an interaction effect between condition and hemisphere from -1800 to -1200 ms prior to action onset ($F(1, 17) \geq 5.12$, $p < .05$) as evidenced by a

[4] We chose a fixed alpha frequency range in order to keep all analyses consistent. However, it is optimal to use an individual alpha frequency since alpha peak frequency tends to vary from person to person (Haegens, Cousijn, Wallis, Harrison, & Nobre, 2014). In a separate analysis, individual alpha frequency (IAF) was determined by choosing each subject's 3-Hz frequency band which showed the strongest attenuation in mu-power during the observation phase (from 0 to 1500 ms relative to action onset; $M(\text{IAF}) = 10.61$ Hz, $SD = 1.82$ Hz). When comparing HP and LP to PN trials we found a decrease in mu-power not only during action observation but also prior to onset for HP trials relative to PN trials (from 1550 to 1100 ms prior to action onset) as well as for LP compared to PN trials (from 1550 to 1400 ms prior to action onset) but not for HP vs LP trials. This analysis thus yields a slightly different outcome than the analyses using the pre-specified mu-frequency band (8-14 Hz) – a topic we will return to in the General Discussion.

decrease in beta-power for HP compared to PN trials over the left (from -1800 to -1200 ms; $F(1, 17) \geq 5.08$, $p < .05$) but not over the right hemisphere (max $F = 1.73$; see Fig 2C and Fig 2D). When running the same analysis for LP x PN trials we found a similar decrease in mu-power for LP compared to PN trials following action onset (from 400 to 1450 ms; $F(1, 17) \geq 4.70$, $p < .05$; see Fig 2A) as well as a decrease in beta-power for LP relative to PN trials post but not prior to action onset (i.e., from 300 to 1200 ms; $F(1, 17) \geq 4.60$, $p < .05$)[5].

For the ERP analysis we found a main effect for condition, prior to (from -360 to -40 ms) and post action onset (from 20 to 340 ms), reflected in a negative potential for HP relative to PN trials over bilateral electrodes in both time segments (see Fig 3A). In addition a main effect of hemisphere was found, reflected in a stronger positive slow wave over the left compared to the right hemisphere prior to action onset (from -1380 to -1120 ms; $F(1, 17) \geq 4.46$, $p < .05$) and a more negative slow wave following action onset (from 760 to 1280 ms; $F(1, 17) \geq 5.00$, $p < .05$; see Fig 3B). A stronger negative slow wave for LP relative to PN trials was found only post action onset (from 120 to 320 ms; $F(1, 17) \geq 4.68$, $p < .05$) across electrodes.

(Un-)predicted vs Unexpected actions

When comparing HP to UP trials over the action observation phase we found no differences in mu or beta-power. Similarly, no difference was detected in either mu or beta-power when comparing LP to UP trials over the same time window (Fs < 4.45). For the ERP analysis we did find a stronger positive slow wave for UP compared to HP trials following action onset (from 1060 to 1140 ms; $F(1, 17) \geq 5.13$, $p < .05$) and for UP compared to LP trials (from 1060 to 1180 ms; $F(1, 17) \geq 4.68$, $p < .05$; see Fig 3A) as well as for UP compared to PN around the same time after action onset (800 to 1160 ms; $F(1, 17) \geq 4.67$, $p < .05$). Finally, a stronger negative slow wave was found over left compared to right electrode sites across HP and UP trials (from 760 to 1500 ms post onset; $F(1, 17) \geq 4.81$, $p < .05$) as well as LP and UP trials (from 800 to 1500 ms post onset; $F(1, 17) \geq 4.64$, $p < .05$).

[5] Although an additional interaction effect was found for condition and hemisphere prior to onset (-900 to -750 ms; $F(1, 17) \geq 4.89$, $p < .05$), no difference in power was found between LP and HP conditions in either the left or right hemisphere.

Cluster-based permutation tests

The ERP results suggest that motor activation prior to action onset and the processing of unexpected actions is not necessary reflected in lateralized motor activity (i.e., contralateral to the observed hand) as hypothesized, but might be more bilaterally or centrally distributed (i.e., the interaction between Condition and Hemisphere was not significant). To investigate this possibility, in an exploratory analysis we performed additional cluster-based permutation tests using the event-related data, which do not require a priori specification of the electrodes of interest (for a summary see Table 2). Shortly explained, cluster-based permutation testing is a form of nonparametric testing in which conditions are compared for each channel and time point. In the process, clusters are formed on the basis of adjacency in terms of neighboring channels and time points. The significant probability of each cluster level statistic (e.g., sum of t-values) is subsequently compared to a large number of random assignments of data to the experimental conditions, computed for each channel and time point. The cluster level statistic is considered significant if it exceeds a specified threshold using two-tailed testing (alpha = .025) in order to correct for multiple comparisons (for more information see Maris & Oostenveld, 2007).

When comparing HP to PN trials we found a stronger negative cluster for HP relative to PN trials starting -405 ms prior to action onset primarily over central electrodes (around Cz), which gradually became more posteriorly localized until 474 ms post onset (cluster p = .002; see Fig 3C). We also found a negative going potential when comparing HP to PN trials starting from 575 ms post action onset, persisting until the end of the action sequence, primarily visible over occipital regions and strongly lateralized to the left hemisphere. Finally, a stronger positive cluster was found (for HP vs PN trials) starting from 528 ms until 1134 ms post action onset over fronto-central electrodes (cluster p = .002). Approximately the same pattern was found when comparing LP to PN trials albeit the timing and specific electrodes involved were somewhat different for the first segment which started around action onset (first negative segment = 3-442 ms; second negative segment = 575-1500 ms; third positive segment = 521-1500 ms; all cluster ps < .05).

When comparing UP to HP trials during the action observation phase we found a stronger positivity for UP relative to HP trials starting from 800 until 1192 ms post action onset primarily over centro-parietal electrodes (cluster p = .016). The same pattern was found when comparing UP to LP trials from 903 until 1196 ms (cluster p = .024; see Fig 3C) reflecting a slow wave effect similar to that found in van Elk et al. (2012) in which a positive amplitude difference was found for unexpected relative to predicted action outcomes. Compared to the conventional analysis over lateral electrodes, we found a pattern of positivity over fronto-central electrodes from 591 until 1267 ms post onset (cluster p = .008) for UP compared to PN trials while a negative going potential reflected the difference in observing movement relative to no movement over left occipital electrodes (580 to 1263 ms; cluster p = .004; see Fig 3C).

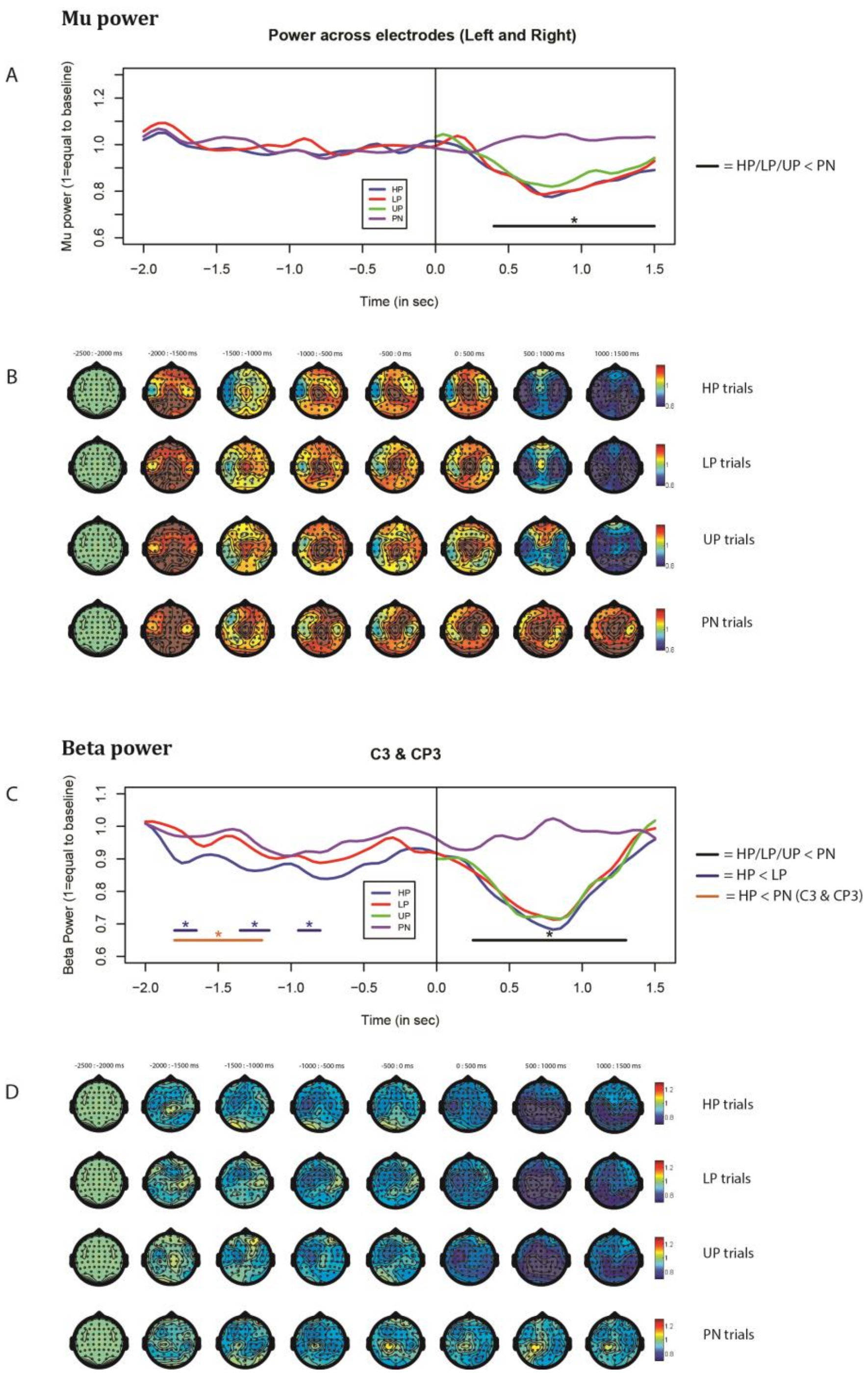


Fig 2. (A) The change in mu-power over time relative to the baseline period (-2500 to -2000 ms relative to action onset) averaged over left (C3, CP3) and right (C4, CP4) electrodes for all stimulus conditions in time bins of 50 ms; HP = high predictable trials; LP = low predictable trials; UP = unexpected trials; PN = no movement trials. (B) Topographic distribution of mu-power (8:14 Hz) from -2500 ms prior to action onset to the end of the action sequence for all stimulus conditions averaged in time-bins of 500 ms; (C) The change in beta-power over time relative to the baseline period (-2500 to -2000 ms) over left (C3, CP3) electrodes for all stimulus conditions in time bins of 50 ms; Significant differences between conditions are marked with an asterisk (*) for the different contrasts. (D) Topographic distribution of beta-power (20:30 Hz) from -2500 ms prior to action onset until the end of the action sequence for all stimulus conditions averaged in time-bins of 500 ms. Significant differences in power over time are marked if at least three subsequent samples were significant (using sequential F-testing)

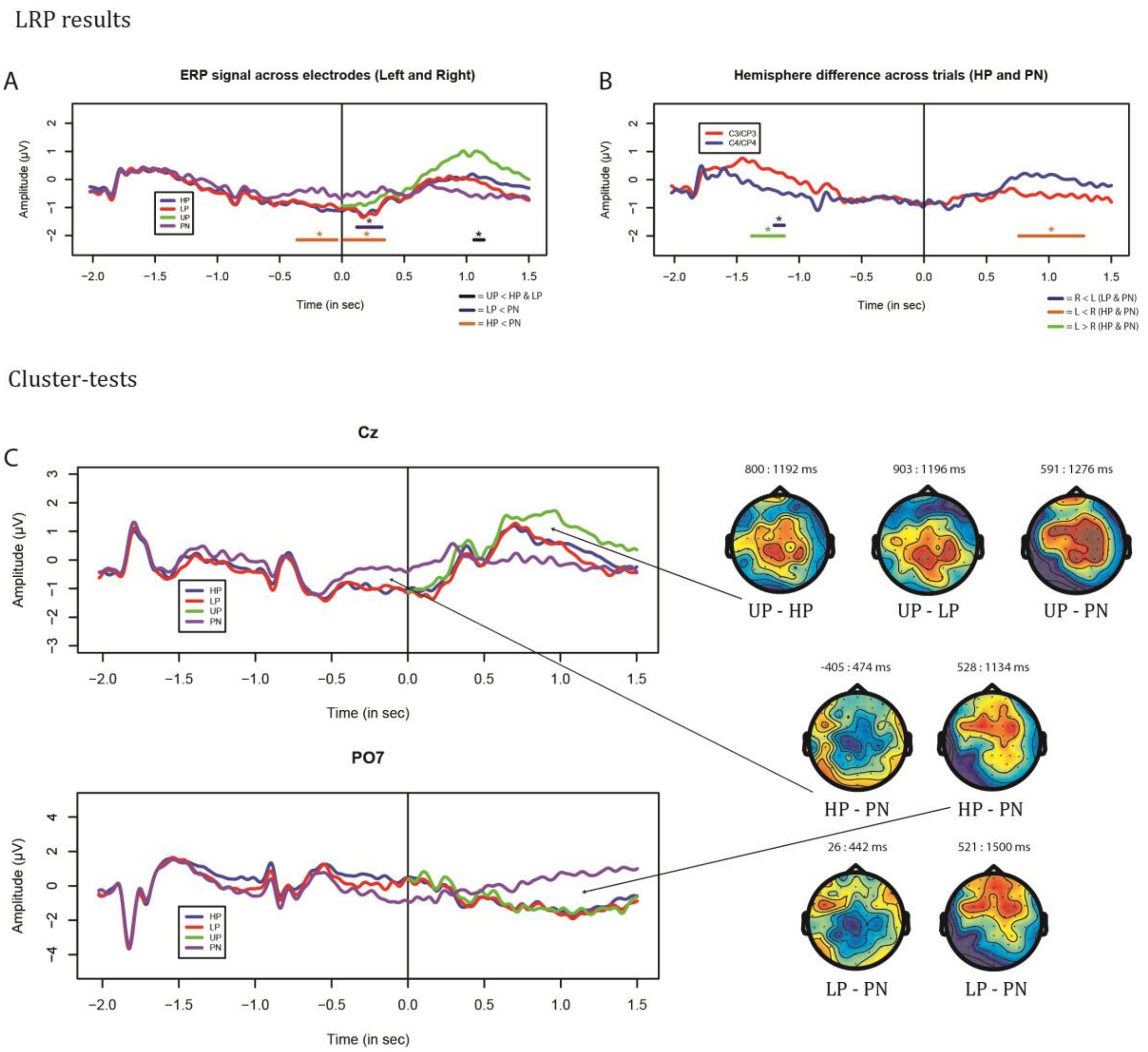


Fig 3. (A) The averaged ERP-signal over electrodes (C3, CP3, C4 and CP4) for all stimulus conditions (HP = high predictable; LP = low predictable; UP = unexpected; PN = no movement trials). All plots display the preprocessed and subsequently averaged signal, which was low pass filtered with a cut-off frequency of 30 Hz using a zero-phase FIR digital filter to prevent phase distortion. Analyses were performed using the unfiltered data; (B) Display of the lateralized ERP-signal reflecting the difference between left (C3, CP3) and right (C4, CP4) hemisphere electrodes averaged over HP and PN trials. Significant differences in amplitude are marked with an asterisk (*). Also marked is the significant difference in amplitude between left and right electrodes averaged over LP and PN trials. (C) *On the left*: Graphs representing the averaged ERP signal over electrode Cz and PO7 which were selected as representative electrodes of the clusters. *On the right*: Topoplots for the ERP signal averaged over the time-window for which the cluster permutation test yielded significant differences between conditions for UP vs. HP, UP vs. LP and UP vs. PN trials. At the bottom the average negative difference is displayed between HP and PN trials and LP and PN trials prior and post action onset including the negative difference over occipital electrodes post onset

Table 2. Cluster permutation tests

Comparison	Time (in ms)	Location	p-value (cluster)
HP < PN	-405 to 474	centro-parietal	$p = .002$
HP < PN	575 to 1500	left-occipital	$p = .002$
HP > PN	528 to 1134	fronto-central	$p = .004$
LP < PN	3 to 442	centro-parietal	$p = .012$
LP < PN	575 to 1500	left-occipital	$p = .002$
LP > PN	521 to 1500	fronto-central	$p = .002$
UP > HP	800 to 1192	centro-parietal	$p = .016$
UP > LP	903 to 1196	centro-parietal	$p = .024$
UP > PN	591 to 1276	fronto-central	$p = .008$
UP < PN	580 to 1263	left-occipital	$p = .004$

Table 3. Results for the cluster-permutation tests. Comparisons between stimulus conditions are displayed along with the time-windows of the significant clusters (relative to action onset) as well as the spatial location and the corresponding cluster p-values. For a visual representation of the respective ERP effects, see Figure 3C

Discussion

In line with the first hypothesis, the analyses revealed a stronger decrease in beta-power (but not in mu-power) prior to the observation of high relative to low predictable actions. Additionally, a stronger decrease in beta-power was found over the left hemisphere (contralateral to the observed hand) only for high and not for low predictable actions compared to trials in which no movement was anticipated. This suggests that actions are anticipated only when sufficiently predictable and that beta-power but not mu-power reflects these anticipation effects. While a power decrease in mu and beta was found during action observation for all actions (high, low and unexpected) relative to no movement trials, this decrease was not modulated by differences in predictability. Accordingly, the findings support a predictive view in which motor activity reflects anticipation of enfolding actions which does not necessarily translate into increased (or decreased) action monitoring during action observation.

High vs Low predictable actions

Critically, we found a stronger decrease in beta-power for high relative to low predictable trials between the onset of the cue and the action. This supports the role of outcome predictability on beta-power although the effect here was not restricted to contralateral electrodes but was found across bilateral electrode sites. A decrease in beta power has been found to affect response preparation prior to initiating a manual response (Tzagarakis et al., 2010) as well reflect response uncertainty during action performance and motor imagery (van Elk et al., 2010; Brinkman, Stolk, Dijkerman, de Lange, & Toni, 2014; Tan, Wade, & Brown, 2016). Nevertheless, it is disputed to what degree beta-power modulation reflects knowledge about motor-specific parameters given its insensitivity to, among others, the force and grasp type of an upcoming response (for an overview see Kilavik, Zaepfell, Brovelli, MacKay, & Riehle, 2013). Importantly, we found no subsequent difference between high and low predictable trials in the observation phase, which rejects the idea that observing low predictable actions subsequently increases motor activity (i.e., due to a process of action monitoring). This finding resonates with a study by Tzagarakis et al. (2010), in which a modulation of beta-power by outcome predictability, determined by the directional (un)certainty of the upcoming response, was found prior to target onset but remained unchanged after the target was presented and gradually disappeared until a manual response was initiated. One possibility is that if unpredicted actions were not anticipated prior to action onset there might be no ‘need’ to monitor potential outcomes. In this case, monitoring is limited to cases where people actually have predictions about upcoming events. Alternatively, it might be that enhanced monitoring is primarily reflected in increased sensory processing of unpredicted actions which is revealed in different frequency bands of cortical oscillations involved in feed-forward connections (e.g., gamma; Arnal & Giraud, 2012; Bastos et al., 2012; van Pelt et al., 2016).

A potential concern could be that since catch trials were always presented prior to action onset, participants did not monitor the actions during the subsequent observation phase, which might explain the absence of an effect for high compared to low predictable trials during action observation. The fact that we observed a clear ERP response to unexpected compared to expected action outcomes argues against this possibility and suggests that participants were paying sufficient attention to each trial sequence.

Movement vs no movement

Additionally we found a decrease in beta-power for high compared to no movement trials. Given that this difference was absent for low predictable compared to no movement trials it seems that, in line with Kilner et al. (2004) and Southgate et al. (2009), motor anticipatory processes might be limited to anticipating high (or fully) predictable actions. This finding is corroborated by the replication of the pre-movement readiness potential (RP) found in Kilner et al. (2004) prior to observing high but not low predictable trials compared to no movement trials. In contrast to Kilner et al. (2004), the effect we found was stronger over centroparietal electrodes (rather than at lateral electrodes) perhaps reflecting a concurrent contingent negative variability in association with the anticipation of action onset (CNV; Jentzsch, Leuthold, & Ridderinkhof, 2004). Interestingly, the laterality of the power difference between high predictable and no movement trials, is comparable with activity patterns found during the preparation and execution of manual responses (Zaepffel et al., 2013) which suggests our findings do not reflect the mere anticipation of sensory changes.

The absence of concurrent differences in the mu frequency band might be due to different functions ascribed to alpha (mu) and beta oscillations in relation to action prediction and observation (Tzagarkis, West, Pellizer, 2015). Although both mu and beta-power reflect similar modulations of motor system activity and are often strongly correlated (Neuper, Wörtz, & Pfurtscherller, 2006; de Lange, Jensen, Bauer, & Toni, 2008), beta oscillations have been linked to action selection or planning of effector specific movements whereas alpha oscillations reflect inhibitory processing in regions irrelevant to the task (Brinkman, Stolk, Dijkerman, de Lange, & Toni, 2014). Nonetheless, when computing mu-power using an individually defined alpha frequency band, a stronger decrease was found between both high and low predictable trials relative to no movement trials during the anticipation phase overlapping in time with the beta-power decrease. Therefore, it seems that individual determined alpha frequency bands that are most responsive to a specific manipulation, might be more sensitive in capturing motor processes involved in action prediction (Klimesch, 1999).

Unexpected actions

Finally, we detected a late positive slow wave effect as shown by a stronger positive slow wave for unexpected actions compared to both high and low predictable actions, comparable to earlier ERP effects on action violations (van Elk et al., 2012; Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008; de Bruijn, Schubotz, & Ullsperger, 2007). This suggests that unexpected effects that were in conflict with one's predictions, can be dissociated from effects that were not predicted (i.e., in the case of low predictable trials). Interestingly, the processing of unexpected actions was not reflected in oscillatory dynamics but only produced changes in terms of event-related potentials which is inconsistent with earlier accounts of mu or beta involvement in monitoring unexpected outcomes (Stapel et al., 2010; van Elk et al., 2010; Koelewijn et al., 2008). Since unexpected actions in the current task were only unexpected due to a violation of expectancy and were not by definition unexpected or extraordinary actions as in van Elk et al. (2010) and Stapel et al. (2010) perhaps this did not require an (online) update of the predicted action outcome but only induced a more general level of action monitoring (de Bruijn et al., 2007).

Prediction strategies

Whereas increased predictability about action outcomes seems to facilitate action anticipation, it is unclear how participants dealt with the uncertainty in low predictable sequences. Given that participants differed in response strategies (e.g., guessing, not preparing at all) there might be multiple ways of preparing for an unpredictable outcome producing varying effects in action observation. Some pre-cueing studies looking at the role of outcome information (full or partial) on response preparation suggest that, while full information provides a response benefit relative to partial information, outcome uncertainty in the latter is partly ‘resolved’ by preparing multiple responses in parallel (for an overview see Cisek &

Kalaska, 2010). For example, in a pre-cueing task with monkeys, neural firing rates specific to either multiple directional cues or different response types (whole hand grasp or precision grip) were comparable, not only for trials in which a single target was cued (full information), but similarly when multiple targets were cued (partial information; Bastian, Riehle, Erlhagen, & Schöner, 1998; Baumann, Fluet, & Scherberger, 2009). In a similar task with human participants, Jentzsch et al. (2004) found that ambiguous or directional pre-cueing (e.g., cueing upward finger movements with either a left or right hand) led to a smaller activity pattern of contingent negative variability (CNV) prior to the target cue over contralateral but a larger pattern over ipsilateral motor areas compared to trials in which only the response hand was cued (e.g., left finger movement). Therefore, it seems that the degree of motor preparation for each outcome is proportional to the number of response outcomes (i.e., capacity-sharing model; Pellizzer & Hedges, 2003) specifically when upcoming actions are mutually suppressive (Praamstra, Kourtis, & Nazarpour, 2009). Whereas the difference in high vs low predictable trials seems in line with this account (i.e., a stronger beta-power decrease was observed for high compared to low predictable actions), the absence of a difference between low predictable and no movement trials is not. Partial knowledge about an upcoming action compared to knowing that no movement will occur would be expected to result in a relative increase in motor activity regardless of outcome uncertainty. Nonetheless, the dissociation between processing unexpected actions and actions observed in low predictable action sequences speaks against the idea that participants would actively guess and therefore anticipate only a single outcome in low predictable sequences which is in line with ratings from the exit questionnaire.

Conclusion

In closing, we provide support for the role of beta (and not mu) oscillations in action processing and suggest that outcome predictability primarily affects motor activation during action anticipation. It seems that the anticipation of actions is restricted to instances where an observer has sufficient (high) certainty about the action outcome compared to events in which no movement is anticipated. Furthermore, we found that both high and low predictable outcomes can be differentiated from unexpected action outcomes, which allow us to functionally dissociate unpredictable from unexpected action outcomes. How actions are processed in low predictable action sequences and to what degree this is affected by the simultaneous preparation of multiple outcomes provides an interesting topic for future investigation. In addition, it is important to assess in which way motor processing, the way it is measured here, is restricted to motor simulative processes or whether it can be recruited to anticipate and track a host of motor and perceptual events outside of an observer's motor repertoire (Press & Cook, 2015). Nevertheless, we have demonstrated here that motor cortical oscillations in the beta-band primarily support action prediction, possibly through specifying a (detailed) motor plan of the upcoming action which is absent when the specific action outcome is not known in advance.

Chapter 6

Discussion

Discussion

This dissertation comprises of four empirical chapters focusing on the interplay between imitative and complementary actions. Whereas imitative actions refer to the display of actions that are congruent to observed actions (e.g., making an open hand movement when observing an open hand movement), complementary actions are social actions, often incongruent to observed actions, that have the aim of attaining joint goals (e.g., carrying a table, pair dancing). Overall, I have identified several variables including the social context, peri-personal space, spatial perspective and outcome predictability that either contribute or do not contribute to the facilitation of both action types. These studies have been inspired and contribute to the idea that even though imitation is a useful tool in human development, automatic imitation is not a pervasive element of ordinary situations, specifically social situations. While I acknowledge that automatic imitative behavior (copying identical actions) is a product of repeated learning, this is restricted to mostly passive instances of learning (e.g., observational learning) and is not the only possible outcome of sensorimotor coupling. I argue here that in a lot of instances complementary actions, which are learned in active social and non-social interactions (i.e., with objects), are more useful than imitation, which is often counter effective. Critical to this idea is the notion that imitation is not an attribute of a build-in mechanism, but rather a product of learning.

Evidence for this notion is accumulating. For example, a recent longitudinal study failed to find evidence for neonatal imitation over a wide variety of gestures and behaviors (Oostenbroek et al., 2016). Gestures included mouth opening, index finger protrusion and emotional expressions, none of which were reliably imitated in the first time point (1 week old) when comparing congruent (similar) to incongruent control gestures. The evidence corroborates an extensive re-analysis by Ray and Heyes (2011) critically examining the existing evidence on neonatal imitation. Together, these findings are in contrast with the idea of a hard-wired matching system, which translates observed in (identical) performed actions (Meltzoff, 1988), thereby solving the correspondence problem. Rather, they are in line with the view that domain-general processes produce imitative behavior over time through associative learning (Heyes, 2015). Direct evidence for the role of learning has shown that learning in a conditioning task in the first months of an infant's life was predictive of performance in 5, 9 and 12 months, following the initial learning phase (Reeb-Sutherland, Levitt, & Fox, 2012). Whereas humans might thus not be born with the capacity to imitate, they might be born with a tendency to attend to close others, predisposed to perceive faces or be more attentive to a variety of social cues (Heyes, 2015).

Although the current dissertation is not about the ontogeny of imitation, the evidence put forward here supports the key point that imitation is ‘only’ one of the outcomes produced by associative learning though repeated coupling between observed and performed actions. This idea has been demonstrated with simple hand gestures which, after a short period of retraining, can trigger the performance of actions (covert as well as overt) that are either similar or dissimilar to observed actions (Catmur, Walsh, & Heyes, 2007). The time-course at which these response effects take place seem to be equal, which suggests that performing dissimilar actions does not require suppressing the automatic tendency to perform similar (congruent) movements (Cavallo, Heyes, Becchio, Bird, & Catmur, 2014).

Before I start discussing the findings that contribute to this proposal I want to highlight and re-iterate several ways in which imitative and complementary actions are shaped in different ways through repeated learning. Imitative actions (covert and overt) can be seen to result from at least three types of learning: 1) through mirror learning and self-observation 2) by copying goal-directed behavior 3) by being imitated (Brass & Heyes, 2005). Mirror learning relates to the visual feedback people get from physical mirrors (i.e., reflective mirrors) producing visual feedback in mirror image (i.e., looking at yourself in the mirror) while self-observation produces feedback from a first-person perspective (e.g., seeing your own hand move). Besides mirror learning, it has been argued that imitation is not only based on simple sensorimotor coupling but can be driven by more complex and higher-order processes (e.g., goal imitation; Bekkering, Wohlschläger, & Gatis, 2000; Ondobaka, de Lange, Newman-Norlund, Wiemers, & Bekkering, 2011). Goal imitation is not restricted to the performance of similar behavior but can also constitute emulative behavior (copying the same goals using different means) and is, similar to being imitated by others (e.g., in parent-infant interactions), driven by input or feedback from a third-person’s (allocentric) perspective. Especially in this case, when observing movements performed by others, imitation requires a transformation from the observed action seen from a third-person perspective into a first-person perspective (Jeannerod & Frak, 1999; ter Horst, van Lier, & Steenbergen, 2010; Vogeley & Fink, 2003). There is support for the role of visual perspectives in imitation as shown by stronger congruency effects (performing faster responses for similar vs dissimilar movements) for hand gestures that appear in mirror perspective (Brass et al., 2000; Koski et al., 2000) or for gestures perceived from a first relative to a third person perspective (Caggiano et al., 2011; Jackson, Meltzoff, & Decety, 2006).

In contrast to imitative actions, complementary actions are learned in interactive settings in which observed actions trigger non-identical responses viewed from a first-person perspective and therefore do not require a transformation from a third to a first-person perspective. Furthermore, complementary actions are not driven by mirror (or passive) learning but by observing others in a strictly interactive setting. In social situations, in which actions performed by others require complementary responses, taking into account the actions performed by an interacting partner is primarily useful to determine one's own behavioral response. Possibly in these situations, covertly imitating actions performed by others is used to encode or track the observed action in order to subsequently prepare a behavioral response (Sartori, Cavallo, Bucchioni, & Castiello, 2011; see paragraph on 'Mechanisms of imitative and complementary actions').

Here, I have presented a collection of four chapters (including 11 studies) that build on the role of (associative) learning in the production of imitative and complementary actions. I will first provide a short summary of each chapter and discuss both merits and limitations. Finally, I will review some issues that are still unclear or remain unanswered and which new issues have come up along the way. In particular, I will discuss the role of control in automatic imitation and the mechanisms involved in producing imitative and complementary actions.

Chapter summaries

The chapters discussed here all relate to factors that play a part in motor learning and specifically, the learning of complementary compared to imitative actions. The second chapter discusses the role of contextual cues in learning complementary actions. This was assessed by looking at attention for means (e.g., objects) that are used to respond to social category members (e.g., police officers), but only in a certain context. Since complementary actions are typically learned in specific contextual situations (e.g., you usually play football on a football field, not in a retirement home), using these cues when observing social category members can trigger actions that are appropriate only in a certain context. Instead of commonly used reaction time paradigms, I used an eye-tracking paradigm in which I measured attention (i.e., looking time) for objects presented on different parts of the screen, following pictures of a context (e.g., stadium) and a social category word (e.g., athlete). The results show that attention for objects that were related to a social category (e.g., criminal – pistol) was largest in contexts in which the object was most useful. For example, people attended to objects related to defensive behavior predominantly when seeing the word criminal coupled with an image of a dark alley compared to that of a courtroom. This implies that gaze behavior is not (strictly) guided by matching semantic connections between social categories and associated behavior (e.g., criminal – violence – pistol). Rather, people use the context in which a social category plays a specific role (criminal in a dark alley vs in a courtroom) to direct attention to response-related objects. These studies were done in order to compare complementary actions to matching or imitative actions. We have shown that observing actions performed by others or behavior typically associated with social categories (e.g., athlete – running) is not the only source driving attention (criminal – gun) but that attention can also be a function of complementary, interactive actions depending on the context (dark alley – criminal – defending).

Even though the two studies discussed in Chapter 2 provide a useful demonstration of the role of contextual cues in complementary actions, it has some limitations. For example, there was no specific control condition in which contextual cues were absent. It is important to assess whether the absence of contextual cues drives attention in terms of matching associations with social categories and not complementary actions per se. However, numerous studies have been done showing that when no context is used, it is possible to trigger covert and even overt behavior upon priming participants with social categories that is consistent with behavior (stereo-)typically performed by members of this category (Bargh & Pietrommonaco, 1982; Chen & Bargh, 1997). This could suggest that social categories typically prime associated behaviors, but can prime complementary actions as well if contextual cues are present. Priming of behavior, however, is quite a controversial topic. Non-replications of behavioral priming experiments (Doyen, Klein, Pichon, & Cleeremans, 2011) as well as larger scale tests of priming paradigms (Manylabs 1; https://osf.io/wx7ck/) have made it difficult to trust priming studies. Even though the studies performed in Chapter 2 refer to this type of priming, the dependent measure was restricted to attention as pre-cursor for actions. Specifically, I based my design on the idea that perception is shaped by action plans in which preparing to perform an action directs attention to environmental cues that are facilitative for action (Bruner, 1957). Attention is therefore typically not described as a behavioral measure but as a precursor to it.

Another, more theoretical problem with the studies is that typically context and social category information are not observed as separate or sequential pieces of information but are commonly seen together (as integrated pieces of information). For example, observing a tiger in a cage when in a zoo will probably be less frightening because you expect the environment to be safe. In contrast, when you spot a tiger on safari, the situation will be less safe and as a result will affect your response options (Cesario, Plaks, Hagiwara, Navarrete, & Higgins, 2010) and attentional choices. In two additional studies (not part of this dissertation) I found that when presenting the context and social category (either humans or non-human animals) simultaneously (e.g., a tiger in a cage or in the wild), this directs attention to objects used as a defensive means to the (social) category, irrespective of the context. In the case of one of these studies, ratings of the social category directly determined the level of mean attention to response-related means rather than an interaction between context and category. Although this is in contrast with the studies described in Chapter 1, presenting integrated pictures misses the point that being in a context first allows you to anticipate upcoming events. Therefore, ideally a study incorporates a sequential presentation procedure in which a social category is added to, rather than following the contextual cue, something which has been used often in more low-level stimulus expectancy tasks (Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008; Demiral, Malcolm, & Henderson, 2012). Future studies should use such a sequential procedure as well as investigate in more depth the relationship between attention and subsequent behavioral measures.

Since the theoretical background of Chapter 1 was based on the display of overt behavior I used a more direct measure of behavioral responses in Chapter 2. Also, because I wanted to see how observed responses trigger the (overt) performance of either similar or dissimilar actions I directly compared the performance of imitative and complementary actions in response to observed gestures.

In Chapter 3 I used a very simple manual response task in which subjects were asked to copy hand gestures they saw on a computer screen either with their hand in mirror perspective (observe right hand – perform left hand movement) or with their opposite hand (observe right hand – perform right hand movement). Either an open hand or closed hand was displayed by an actor on the screen so that responses reflected a mirror congruent response (mirroring a right closed hand with a left closed hand) or a complementary response (performing a right open hand upon observing a right open hand; handshaking), depending on the color of the hand in each trial. In addition to the type of gesture, I also manipulated the perceived distance between the actor onscreen and the participant. I reasoned that as a result of (associative) learning, complementary actions are commonly (although not exclusively) performed when others are within one's peripersonal space. Whereas there is no space restriction to imitative behavior, an individual can copy another's actions irrespective of interpersonal distance; complementary actions such as handshaking, can only be performed when others are within reach. This design extends the findings in Chapter 1 by directly comparing imitative to complementary actions and by using a different contextual cue (i.e., distance) to see whether this modulates task performance. Five studies were done in order to carefully assess the hypothesis that distance would modulate performance of complementary actions but not of imitative actions. Across the studies, no evidence was found for distance as modulating factor. Rather, open hand gestures primed complementary responses (i.e., handshaking) irrespective of distance whereas (weak) support was found for imitative responses to closed hand gestures across distance. This implies that complementary hand responses can be triggered upon perceiving hand gestures but that this effect is not modulated by the (perceived) ability to actually perform the action. Since the modulating role of space was demonstrated previously in the context of object affordances (Costantini, Ambrosini, Tieri, Sinigaglia, & Committeri, 2010), the results suggest that there might be something to social gestures that affects behavioral responses in a different way than non-social objects. One simpler reason for the discrepancy might be that the task used in Costantini et al. (2010) required participants not to respond directly to the object but rather to copy a hand gesture that was either congruent or incongruent to the object affordance (i.e., direction of the ear of a cup). This suggests that the distance effect found in their study might be driven by the expectancy of the visual consequences of an action typically perceived in that setting rather than by the direct response to the object (Tucker & Ellis, 1998). Although these processes are not mutually exclusive, this makes it difficult to directly compare the current findings to theirs.

Setting aside earlier findings, the third chapter provides evidence for a stimulus-response effect that is not bounded to contextual factors or to affordance criteria. Accordingly, automatic responses (imitative or complementary) might be driven by both the physical and imagined possibility to respond to (social) stimuli. To further investigate this an additional (unpublished) study was performed in which the same task was transformed into a physical setting with dyads performing imitative or complementary actions in a mutual response task in which interpersonal distance was manipulated by placing a screen between the two participants in a subset of trials. Unfortunately, no conclusive evidence was found for response facilitation for both gestures nor was there an effect of distance. Only faster responses to open hand compared to control gestures were found, irrespective of response hand (left or right) or distance (close or far). In a future setup increased experimental control is needed to create a task that is better comparable to the studies used in the second chapter.

Besides practical issues there are some further limitations that are partly addressed in Chapter 3. For example, the distance manipulation might not have been clear enough, given the variability in subjective ratings of distance in study 1 through 4. From the data analyzed in the supplementary material it seems that subjective ratings do not affect response measures at all. This is in contrast with earlier studies in which subjectively judging distance is linked to neurophysiological correlates of distance in object perception (Valdés-Conroy, Sebastián, Hinojosa, Román, & Santaniello, 2014). In Valdés-Conroy et al. distance was physically manipulated rather than onscreen. Perhaps therefore the ratings I used were perceived as more ambiguous compared to those used in a real life setting.

Furthermore, while it is common in motor tasks to use a response interference paradigm in which participants are asked to ignore motor information (e.g., hand gesture), I asked participants to directly respond to the motor stimulus in the task. While ignoring motor properties tests facilitation or inhibition of automatic response priming, directly responding to the (motor) properties of the task does not (Brass, Bekkering, Wohlschläger, & Prinz, 2000). As evident in Brass et al. (2000) a congruency effect was found, as evident by faster responses to congruent compared to incongruent hand gestures, only when participants were instructed to respond to irrelevant task properties (i.e., performing a hand gesture based on a numerical cue) and not when copying the perceived gesture directly. This congruency effect seems to be driven partly by the mirror similarity of the hand gesture (seeing a hand gesture in mirror image) and by the gesture congruency (seeing a right hand in opposite view and using your right hand; Bertenthal, Longo, & Kosobud, 2006). My manipulation somewhat falls in the middle. While participants responded to the color cue of the hand, they were instructed to copy the observed hand gesture (similar to van Schie, Waterschoot, & Bekkering, 2008) with either the same hand (in mirror perspective) or the opposite hand. In this case, participants always directly copied the observed gesture, but did so based on a non-motor stimulus property (the hand color). Given that I found a consistent effect across studies, in line with previous research (Flach, Press, Badets, & Heyes, 2010; van Schie et al., 2008), this should not complicate the results.

Building on the type of low-level interactions studied in Chapter 3, Chapter 4 looked at the role of perspective taking in interactive settings. In Chapter 2 and 3 the premise was that due to associative learning, people would learn to respond to others by performing either imitative or complementary actions. Depending on factors as social context (Chapter 2) or interpersonal distance (Chapter 3), associative links are formed specifically within the context in which they are learned (Heyes, 2016). Chapter 4 builds on the same idea but uses a perspective taking task to examine how interactive settings are learned mainly through a first-person perspective, as explained in the first part of the discussion. Whereas imitation is often linked to (allocentric) perspective taking, or taking the perspective of others to understand what they are seeing or experiencing (Jackson et al., 2006; Kessler & Thompson, 2010), interactive behavior is driven by egocentric perspective taking (taking a first-person perspective). Therefore, perspective taking is a useful measure to directly compare imitative to complementary actions in social interactive situations.

In Chapter 4, three studies were conducted, using a spatial perspective taking task. Participants were shown pictures on which an actor was shown sitting at a table on which two objects were positioned. Participants were asked to determine the relative position of one of the objects while the actor either grasped the object (active), handed it over (interactive) or did not manipulate the object at all, but was simply present in the scene (actor). I found that the frequency of voluntary perspective taking did not decrease when the setting displayed an interactive compared to active setting. Only when using an interference paradigm in which participants were instructed to take their own perspective (egocentric) or the perspective of the actor (allocentric) while suppressing the urge to take the perspective of the actor (their own perspective), I found that spatial judgments were more difficult (as evidenced by delayed response times), when participants took the perspective of the actor in interactive settings in which a complementary request was made, relative to active settings. This suggests that when asked to take the perspective of the actor in an interactive setting, participants automatically took their own perspective leading to delayed response times. However, across two studies, no difference of this increased difficulty between actor and interactive trials was found when participants took their own or the actor's perspective. When directly comparing interactive to active settings within egocentric and allocentric trials, evidence of delayed response times in allocentric trials was found, but no interaction effect when considering response times for all trial types at once. This suggests that irrespective of the instructed perspective, interactive settings increased the difficulty of making spatial judgments. Interpreting this finding is quite complex. Since the visual differences between trial types were quite subtle, it is unlikely that specific stimulus characteristics could have produced differences in response times.

Furthermore, even with minimal stimulus differences I still found large differences between active and interactive settings in response times across two studies. One option is that although the two control settings (actor and active) did not directly reflect interactive settings, they could be interpreted as developing into interactive settings over time. In this case, the anticipation to receive a complementary request could have minimized response differences following both task instructions. This reasoning fits with Tversky and Hard (2009), who used a similar paradigm and argued that merely perceiving an actor acting on an object would be sufficient to increase perspective taking due to the increased relevance for the observer perhaps by triggering the preparation of a complementary response.

Another potential explanation for the results relates to embodied transformation, or the transformation of one's own perspective to that of another person by mental rotation. It is assumed and empirically demonstrated in Tversky and Hard (2009) and Mazzarella, Hamilton, Trojano, Matromauro, and Conson (2012), that displaying an actor either gazing towards or manipulating an object increases (allocentric) perspective taking. Given that the frequency of participants taking another's perspective is quite low across these studies, it is unclear whether perspective taking and therefore embodied transformation is truly functional in every day, interactive situations or whether it is a product of task properties in these studies. For example, evidence for perspective taking is primarily found in task interference studies (Surtees, Butterfill, & Apperley, 2012; Surtees, Apperley, & Samson, 2013) where participants are explicitly instructed to take another's perspective which increases response times when judging the spatial position of an object incongruent compared to congruent with one's own perspective. It is argued that the process of embodied transformation is interrupted by one's intention to judge everything from an egocentric perspective, an effect coined egocentric interference. Since I did not find differences in egocentric and allocentric trials it might be possible that different mechanisms besides embodied transformation inhibited response times in interactive settings such as the tendency to perform a complementary response which interfered with the spatial judgment task.

It might proof beneficial to integrate the findings of Chapter 3 with those in Chapter 4. Some have argued for a direct connection between observing other's actions and simulating these actions in one's own motor system to get a first person's perspective of the observed action (Singer, 2006). Based on my findings, I would suggest that the type of action (e.g., interactive) affects the connection between covert action processing and (spatial) perspective taking by either facilitating or inhibiting perspective taking depending on the task roles in an interaction. Nonetheless, it might be that for some tasks perspective taking or embodied transformation is not a necessary element unless it is explicitly part of the task instruction. This is discussed in more detail at the end of Chapter 4.

The fifth chapter uses a different setup and dependent measure than the first three chapters. While the first three chapters concern motor responses (imitative or complementary) in response to observed others (or actions), motor responses are often predictive in nature (Wilson & Knoblich, 2005). In line with associative sequence learning (ASL), repeated learning creates connections between observed actions (or others) and the performance of either similar (imitative) or different (complementary) actions. Over time this stimulus-response relationship produces predictions about upcoming actions prior to action observation. As illustrated in the summary of the first paper, being in a context creates expectancies about people who will appear in this context and what actions one might need to prepare. This type of predictive processing can affect gaze tracking (Chapter 2), by means of anticipatory gaze to parts of a (context) scene that provide valuable information (Ambrosini, Reddy, de Looper, Costantini, & Sinigaglia, 2013; Faber & Jonas, unpublished data) or it could affect complementary responses if the outcome of an interaction can be predicted in advance. Even perspective taking can be anticipatory if upcoming situations are expected to be self-relevant or not although this has not yet been explicitly tested. Prediction therefore serves as a tool building on ecological cues (such as social context or distance) in order to anticipate stimulus events. It is important to note that the fifth chapter does not directly relate to the main question comparing imitative to complementary actions but rather focuses on predictive processing of simple object-directed actions.

The study reported in Chapter 5 provides a more global picture and explicitly compares the role of motor processing during observation, which is commonly used to test covert imitative responses to observed actions, to motor processing prior to observation. This study was done using an EEG setup in which I looked at areas in the brain involved in both the performance as well as observation of motor tasks (e.g., hand movements). My aim was to show that motor involvement during action observation can be partially driven by predictive processes that occur prior to the onset of an action. To do this, I manipulated the degree of outcome predictability to see how this would affect motor processing in both time windows.

In the study I showed participants a series of object-directed hand movements (e.g., hand grasping a cup with a whole hand grasp or a precision grip) that were either highly predictable (occurring in 70% of trials) or non-predictable (50-50%) prior to movement onset. Subsequently I measured motor activity (covert motor processing) prior and post movement onset to estimate whether having more (or less) information concerning the outcome of an upcoming action, would affect motor processing in both time periods. The results indicated that highly predictable actions led to stronger motor processing compared to low predictable as well as predictable no-movement actions (trials in which no movement was predicted to occur) prior to action onset. Following action onset, no differences were found in terms of motor activity between high and low predictable action trials. I did however find that highly predictable trials followed by an unexpected action could be detected in terms of a positive slow wave difference between unexpected and highly predictable actions. Interestingly, a similar distinction was found between unexpected actions and actions observed in low predictable trials, suggesting that outcomes in low predictable actions were not guessed prior to action onset, which was consistent with the majority of the ratings from the exit-questionnaire. Taken together, this study has shown that the level of predictability about potential action outcomes affects motor processing primarily prior to action onset, while motor processing during onset only reflects monitoring of unexpected outcomes.

Given the setup of the final study it is more complex to couple the findings to those in the first three chapters. Although most studies relate to complementary actions, different measures were used (e.g., attention, manual responses) that all play different parts in an action sequence. Nonetheless, I think that the last study nicely supports the role of prediction in motor processing which can be applied to interactive settings as well. In contrast to a viewpoint in which observed movements are interpreted and translated into similar movements, associative learning produces predictive links between observed and performed actions over time. This seems to be an adaptive quality specifically in interactive situations as already shown by Ménoret et al. (2014) and Sartori et al. (2011). However, the objects I used in Chapter 5 are not agents and therefore it is not possible to anticipate intentional movements, which is crucial in the case of interaction partners. Consequently, the study in Chapter 5 only captures motor (predictive) processing for self-performed movements (or movements seen from a first-person perspective) and not processing of other-performed movements (from a third-person perspective). From numerous findings we know, however, that motor processing of movements observed from a first- or third-person perspective overlaps quite strongly, even when observing non-human agents or when anticipating actions

(Oberman, Pineda, & Ramachandran, 2007; Kilner et al., 2004; Oberman, McCleery, Ramachandran, & Pineda, 2007).

Another important finding from the final chapter pertains to the laterality of motor processing in the anticipation phase which seemed to be different when comparing high to low predictable trials and high predictable to predictable no-movement trials. Whereas the high vs low comparison yielded stronger motor processing for high predictable trials over bilateral sensorimotor regions, the high vs no-movement comparison yielded stronger motor processing specifically over the left part of the brain, contralateral to the observed hand. This could suggest that enhanced processing in high relative to low predictable trials was smaller (relative to high vs no-movement), because low predictable trials yielded a relative increase in motor processing compared to no-movement trials. Eyeballing the graph in Chapter 5 (Figure 2) seems to support the view that motor processing for low predictable trials fell somewhat in the middle of no-movement and high predictable trials even though this was not supported statistically. Another option is that no-movement trials in fact led to increased predictive control given the fact that only these trials were fully predictable across the experiment. Some related evidence suggests that event-related synchronization precedes stimulus onset primarily in the alpha frequency band (Klimesch, Sauseng, & Hanslmayr, 2006). From the same (EEG) power graph in Chapter 4 you can similarly observe patterns of enhanced synchronization in motor regions contralateral to the observed hand for predictable no-movement trials. However, these patterns were evident primarily in the beta-frequency range and were stronger during rather than prior to the observation phase. Nonetheless, inhibitory processes might have affected the difference in motor processing between high predictable and no-movement trials.

Before providing an overview of the findings and a take-home message that can be distilled from this dissertation it is important to review some of the larger questions pertaining to the presented findings. In particular, I will review the mechanisms that drive imitative and complementary actions and the role of action control in transforming covert to overt imitative as well as complementary actions.

Mechanisms of imitative and complementary actions

Debate has been ongoing about the mechanisms that produce either imitative or complementary actions in social interactions. Research by Sartori and colleagues suggests that action observation in social settings follows a two-step process (Sartori, et al., 2011; Sartori, Cavallo, Bucchioni, & Castiello, 2012), in which observing object-directed hand movements triggers covert simulation of similar (imitative) actions in order to prepare for a subsequent complementary request. This implies that, depending on the type of context, action simulation is not restricted to similar actions (imitative), but can trigger complementary actions specifically when predicted to occur next in an action sequence. Accordingly, when present in a social setting in which no complementary request is observed, covert imitation might be less relevant. This might have partly contributed to the small congruency effect for closed hand gestures reported in Chapter 3. While performing congruent actions upon observing simple meaningless gestures might be primarily driven by mirror learning (Brass et al., 2000; although see Berthental et al., 2006), seeing a full posture might have increased self-other dissociation (seeing oneself as separate from another) and subsequently decreased the need to copy another's gestures.

The two-step process also partly overlaps with the reasoning in Keysers and Gazzola (2014). To reiterate some points from the introduction, their paper revolves around the role of Hebbian learning in perception-action coupling. They argue that, in line with an associative learning account, sensorimotor coupling is a product of contiguity and contingency learning. That is, learning is both driven by the probability that the execution of an action is followed by the observation of the same action (contiguity), as well as the probability that observing the action is not preceded by the execution of the same action, which weakens stimulus-response relationships (contingency). Initially, sensorimotor coupling is shaped by direct connections between observed and executed actions. For example, the motor command of grasping a cup provides direct visual feedback of the same action. In fact, due to the delay in neural communication between visual and motor areas, visual feedback follows the motor command to perform an action rather than occurring at the same time (Wolpert, Doya, & Kawato, 2003). As a consequence, visual feedback provides information about what action to perform next, by generating motor commands that are useful in an upcoming event. Based on the reasoning by Keysers and Gazzola (2014), both imitative and complementary actions could be triggered based on the type of setting and the time point in an interaction. If passively observing actions performed by others can trigger covert imitation, anticipating a complementary request can trigger covert complementary responses prior to observing this request. My study in Chapter 5 partly supports this idea by demonstrating that the level of predictability modulates motor processing prior to action observation.

A related discussion on the mechanisms of imitation and complementary actions deals exclusively with the role of mirror neurons in action observation. While mirror neurons have typically been suggested to code for similar actions (direct-matching) and therefore to provide a foundation for imitative behavior (Iacoboni et al., 2009), some have suggested that mirror neurons respond more strongly to complementary actions (Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007). Similarly, mirror neurons have been argued to facilitate social responding (Hamilton, 2013), which partly overlaps with the proposal in Keysers and Gazzola (2014). Furthermore, mirror neurons have been coupled to (social) action selection in which mirror activity reflects the predicted consequences of upcoming motor events (Hickok, 2009; Kilner, 2011). It is still debated to what extent mirror neurons are involved in coding higher order concepts involved in predicting actions and preparing responses. For example, it has been shown that whereas mirror neurons respond to low-level action goals, inferring the intentions of others is restricted to differential activity in the mentalizing network (de Lange, Spronk, Willems, Toni & Bekkering, 2008). Besides the distinction between lower and higher-order action processing, the key point is that mirror neurons are not dedicated to imitation per se, but are likely to respond to different non-identical actions as well providing a potential unified mechanism underlying imitative and complementary actions (Cook, Bird, Catmur, Press & Heyes, 2014; Hamilton, 2015).

Action Control

Another important topic which speaks to research discussed in this dissertation as well as in the broader context of imitation is the concept of control. A large number of neuro-imaging and electrophysiological studies on imitation commonly focus on covert action responses (i.e., motor simulation) where participants passively observe (object-directed) actions (Rizzolatti & Craighero, 2004). Even though covert motor processing has been argued to be a potential source for imitation, for example through mirror neurons (Iacoboni et al., 2009), covert responses do not consistently translate to overt responses (as in echopraxia). There is evidence that when explicitly instructed to perform manual responses (as in Brass et al., 2000), performance is affected by the congruency of observed gestures. Nonetheless, there is a discrepancy between response inhibition (or facilitation) when instructed to copy behavior by an external cue and the voluntary decision to copy observed actions.

One line of research has paid attention to the role of control in dissociating internally produced actions from actions performed in response to others (Brass, Ruby, & Spengler, 2009). Brass et al. suggest that mentalizing areas are involved in explicitly distinguishing self and other representations which allows for control or inhibition of automatic imitative behavior. This suggests that automatic imitation might be driven by self-other overlap, while responding to others typically requires self-other dissociation. Evidence in support of this idea has shown that learning to control imitation of hand gestures improves self-other dissociation and in turn performance on theory of mind tasks (Santiesteban et al., 2011). While action inhibition tasks speak to the role of specific brain areas in exerting control on automatic imitation, it is unclear whether the same mechanisms translate covert into overt behavior. Recent evidence suggests that the pre-supplementary motor area (SMA; Spieser, van den Wildenberg, Hasbroucq, Ridderinkhof, & Burle, 2015) as well as the inferior frontal gyrus (IFG; Herz et al., 2014) play a relevant part in modulating overt responses. Interestingly, parts of the IFG has also been argued to be part of the mirror neuron system (Kilner, Nael, Weiskopf, Friston, & Frith, 2009). Similarly, a subset of mirror neurons has been found to show suppressive responses during action observation which inhibit overt action responses when observing actions performed by others (Mukamel, Ekstrom, Kaplan, Iaocoboni, & Fried, 2010). Another recent study suggests that simply instructing people to perform a hand gesture is sufficient to modulate covert action responses (Bardi, Bundt, Notebaert, & Brass, 2015). When participants in Bardi et al. were instructed to perform hand movements that were either similar or dissimilar to an observed hand gesture in a subsequent part of the experiment, enhanced MEPs were recorded in line with the instructed action (e.g., observing open hand – performing closed hand in later block) even in the absence of repeated training and preceding the actual manual response task. This suggests that covert responses can be controlled and selected, although it is unclear whether these findings can be extended to complementary gestures as well. One way to test this is to use the instructions in Bardi et al. (2015) in order to see how this affects covert action responses by comparing imitative to complementary actions. Given that studies on imitation and complementary actions selectively use covert and overt responses as dependent measures, it is important to incorporate action control to understand how these types of responses relate to each other.

It is also important to reiterate the importance of goals or motives for imitation (see also the Introduction) that exceed previous examples of low-level control. While the current dissertation exclusively covers automatic accounts of imitation and specifically the neural underpinnings and development of imitative vs complementary actions, imitation in day to day situations is strongly determined by the intentions of the actor. Besides that these intentions regulate the translation between observing and performing observed actions, they structure which actors to attend to and what benefit imitating others, but also complementing them, can bring an individual (Over & Carpenter, 2012). Ideally, an account that is based on sensorimotor (or associative) learning should incorporate higher order mechanisms (e.g., intentions) to get a better understanding of the development of imitative and complementary actions.

Conclusion

This dissertation has argued that through associative learning, connections are produced between observed and performed behavior that can be both similar (imitative) as well as different and meaningful (complementary). I have built on collective evidence that a) imitation is not an inborn ability but a product of sensorimotor learning; b) (covert) imitative responses can easily be retrained; c) imitative responses are not useful in most social interactive situations. I have tried to expand these findings by showing that a) context directs attention to objects relevant for complementary actions; b) complementary actions are triggered automatically irrespective of interpersonal distance; c) interactive situations inhibit perspective taking, and d) motor processing precedes action observation if outcomes are sufficiently predictable. Taken together, these studies demonstrate partial support for the role of associative learning in producing imitative and complementary actions. The presented evidence contributes to the idea that sensorimotor learning is flexible by showing that automatic complementary actions share similar features to imitative actions and highlights several factors that contribute to this learning process. This may provide a source for future research.

In future studies, the focus should be on the role of control in dissociating imitative from complementary actions, for example, by looking at the role of error monitoring in imitative and complementary actions. Furthermore, the role of prediction should be used to understand the time course of imitative and complementary actions in social interactions. It is likely that observers switch between imitative and complementary actions variably over time, which cannot be captured by passive observation tasks that are common in most experimental setups. Although preliminary evidence has been done looking at response performance in physical interactions (Ménoret et al., 2014), it is unclear how motor processing is affected by degrees of predictability of upcoming events prior to action observation.

References

Aarts, H., Dijksterhuis, A., & De Vries, P. (2001). The psychology of drinking: Being thirsty and perceptually ready. *British Journal of Psychology, 92*, 631-642. doi/10.1348/000712601162383.

Ambrosini, E., & Costantini, M. (2013). Handles lost in non-reachable space. *Experimental Brain Research*, *229*, 197-202. doi:10.1007/s00221-013-3607-0

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PLoS One*, *8*, e67916. doi:10.1371/journal.pone.0067916

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *The Quarterly Journal of Experimental Psychology*, *63*, 1201-1217. doi:10.1080/17470210903281582

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*, 390-398. doi:10.1016/j.tics.2012.05.003

Bandura, Albert (1977). *Social Learning Theory*. Oxford, England: Prentice-Hall.

Bardi, L., Bundt, C., Notebaert, W., & Brass, M. (2015). Eliminating mirror responses by instructions. *Cortex*, *70*, 128-136. doi:10.1016/j.cortex.2015.04.018

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230-244. doi: 10.1037/0022-3514.71.2.230

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437-449. doi:10.1037/0022-3514.43.3.437

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Current Psychology of Cognition, 13*, 207-239.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (Eds.). (2000). *Understanding other minds: Perspectives from Developmental Cognitive Neuroscience.* Oxford: Oxford University Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-609. doi:10.1.1.4.5511

Bastian, A., Riehle, A., Erlhagen, W., & Schöner, G. (1998). Prior information preshapes the population representation of movement direction in motor cortex. *Neuroreport*, *9*, 315-319. doi:10.1097/00001756-199801260-00025

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*, 695-711. doi:10.1016/j.neuron.2012.10.038

Baumann, M. A., Fluet, M. C., & Scherberger, H. (2009). Context-specific grasp movement representation in the macaque anterior intraparietal area. *The Journal of Neuroscience*, *29*, 6436-6448. doi:10.1523/JNEUROSCI.5479-08.2009

Baumeister, R., Bratslavsky, E., Finkenauer, C., & Vohs, K. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323-370. doi: 10.1037//1089-2680.5.4.323

Bekkering, H., Wohlschläger, A., & Gattis, M (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology*, *53*, 153-164. doi:10.1080/713755872

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 210-225. doi:http://dx.doi.org/10.1037/0096-1523.32.2.210

Birch, S. A., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8*, 255-260. doi:10.1016/j.tics.2004.04.011

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242-261. doi: 10.1207/S15327957PSPR0603_8

Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, *43*, 260-267. doi:10.1016/j.neuropsychologia.2004.11.012

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*, 3-22. doi:10.1016/S0001-6918(00)00024-X

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*, 124-143. doi:10.1006/brcg.2000.1225

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences, 9,* 489-495. doi:10.1016/j.tics.2005.08.007

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behavior and social cognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*, 2359-2367. doi:10.1098/rstb.2009.0066

Brinkman, L., Stolk, A., Dijkerman, H. C., de Lange, F. P., & Toni, I. (2014). Distinct roles for alpha-and beta-band oscillations during mental simulation of goal-directed actions. *The Journal of Neuroscience*, *34*, 14783-14792. doi:10.1523/JNEUROSCI.2039-14.2014

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123-52.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, *51*, 5-13. doi:http://doi.org/10.5334/pb-51-1-5

Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A., & Casile, A. (2011). View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Current Biology, 21*, 144-148. doi:10.1016/j.cub.2010.12.022

Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., & Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science, 324*, 403-406. *doi:*10.1126/science.1166818

Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral Cortex, 15,* 1243-1249. doi:10.1093/cercor/bhi007

Carroll, W. R., & Bandura, A. (1990). Representational guidance of action production in observational learning: A causal analysis. *Journal of Motor Behavior*, *22*, 85-97. doi:10.1080/00222895.1990.10735503

Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology*, *17*, 1527-1531. doi:10.1016/j.cub.2007.08.006

Cavallo, A., Heyes, C., Becchio, C., Bird, G., & Catmur, C. (2014). Timecourse of mirror and counter-mirror effects measured with transcranial magnetic stimulation. *Social Cognitive and Affective Neuroscience, 9*, 1082-1088. doi:10.1093/scan/nst085

Cesario, J., & Jonas, K. J. (2014). Replicability and models of priming: What a resource computation framework can tell us about expectations of replicability. Understanding priming effects in social psychology. *Social Cognition, 32*, 124-136. doi:10.1521/soco.2014.32.supp.124

Cesario, J., Plaks, J. E., Hagiwara, N., Navarrete, C. D., & Higgins, E. T. (2010). The ecology of automaticity how situational contingencies shape action semantics and social behavior. *Psychological Science, 9,* 1311-1317. doi:10.1177/0956797610378685

Cesario, J., Plaks. J. E., & Higgins, E. T. (2006). Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology, 90,* 893-910. doi:10.1037/0022-3514.90.6.893

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893-910. doi:http://dx.doi.org/10.1037/0022-3514.76.6.893

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*, 541-560. doi:10.1006/jesp.1997.1329

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, *33*, 269-298. doi:10.1146/annurev.neuro.051508.135409

Clark, A. (1997). *Being there.* Cambridge: MA: MIT Press.

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: from origin to function. *Behavioral and Brain Sciences*, *37*, 177-192. doi:10.1017/S0140525X13000903

Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 840-852. doi:http://dx.doi.org/10.1037/a0019256

Cooper, R. P., Cook, R., Dickinson, A., & Heyes, C. M. (2013). Associative (not Hebbian) learning and the mirror neuron system. *Neuroscience Letters*, *540,* 28-36. doi:10.1016/j.neulet.2012.10.002

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*, 1314-1329. doi: 10.1037//0022-3514.83.6.1314

Correll, J., Wittenbrink, B., Park, B., Judd, C. M., & Goyle, A. (2011). Dangerous enough: Moderating racial bias with secondary threat cues. *Journal of Experimental Social Psychology, 47,* 184-189. doi:10.1016/j.jesp.2010.08.017

Costantini, M., Ambrosini, E., Scorolli, C., & Borghi, A. M. (2011). When objects are close to me: affordances in the peripersonal space. *Psychonomic Bulletin & Review*, *18,* 302-308. doi:10.3758/s13423-011-0054-4

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., & Committeri, G. (2010). Where does an object trigger an action? An investigation about affordances in space. *Experimental Brain Research*, *207,* 95-103. doi:10.1007/s00221-010-2435-8

Crostella, F., Carducci, F., & Aglioti, S. M. (2009). Reflexive social attention is mapped according to effector-specific reference systems. *Experimental Brain Research*, *197*, 143-151. doi:10.1007/s00221-009-1900-8

Csibra, G. (2008). Action mirroring and action understanding: an alternative account. In P. Haggard, Y. Rosetti and M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: attention and performance XXII*. (pp.435 - 459). Oxford: Oxford University Press

Darwin, C., (1989), The Descent of Man, and Selection in Relation to Sex: Part One. In P. H. Barrett, & R. B. Freeman (eds), *The Works of Charles Darwin* (vol 21). London: Pickering.

Dawkins, R. (1976). *The selfish gene*. Oxford : Oxford University Press.

de Bruijn, E. R., Schubotz, R. I., & Ullsperger, M. (2007). An event-related potential study on the observation of erroneous everyday actions*. Cognitive, Affective, & Behavioral Neuroscience, 7, 278-285.* doi:10.3758/CABN.7.4.278

De Lange, F. P., Jensen, O., Bauer, M., & Toni, I. (2008). Interactions between posterior gamma and frontal alpha/beta oscillations during imagined actions. *Frontiers in Human Neuroscience, 2: 7.* doi:10.3389/neuro.09.007.2008

De Lange, F. P., Spronk, M., Willems, R. M., Toni, I., & Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current Biology*, *18*, 454-457. doi:10.1016/j.cub.2008.02.057.

Demiral, Ş. B., Malcolm, G. L., & Henderson, J. M. (2012). ERP correlates of spatially incongruent object identification during scene viewing: Contextual expectancy versus simultaneous processing. *Neuropsychologia*, *50*, 1271-1285. doi:10.1016/j.neuropsychologia.2012.02.011

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, *33*, 1-40. doi:10.1016/S0065-2601(01)80003-4

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res, 91*, 176–80. doi:10.1007/BF00230027

Doyen, S., Klein, O., Pichon, C. L., & Cleeremans, A. (2012). Behavioral priming: it's all in the mind, but whose mind? *PloS One, 7*: e29081. doi:http://dx.doi.org/10.1371/journal.pone.0029081

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*, 509-540. doi:10.1080/135062899394920

Duchowski, A. T. (2002). A breadth-first survey of eye tracking applications. *Behavior Research Methods, Instruments, & Computers, 34*, 455-470. doi: 10.3758/BF03195475

Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition, 41*, 93-103. doi:10.1016/j.concog.2016.02.010

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40,* 760-768. doi:10.1016/j.jesp.2004.02.002

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608-2611.

Faber, T. W., & Jonas, K. J. (2016). *Perspective switching in social interactions: When responding favors an egocentric perspective*. Manuscript in preparation, Department of Psychology, University of Amsterdam, Netherlands.

Faber, T. W., van Elk, M., & Jonas, K. J. (2015). Complementary hand responses occur in both peri- and extrapersonal space. *PLoS One, 11:e0154457.* doi:http://dx.doi.org/10.1371/journal.pone.0154457

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme information. *Journal of Personality and Social Psychology, 38*, 889-906. doi: 10.1037/0022-3514.38.6.889

Flach, R., Press, C., Badets, A., & Heyes, C. (2010). Shaking hands: Priming by social action effects. *British Journal of Psychology*, *101*, 739-749. doi:10.1348/000712609X484595

Flavell, J. (1974). The development of inferences about others. In T. Mischel (Ed.), *Understanding other persons* (pp. 66–116). Oxford: Blackwell.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology*, *17*, 99-103. doi:http://dx.doi.org/10.1037/0012-1649.17.1.99

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694-724. doi:http://doi.org/10.1037/0033-2909.133.4.694

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*, 127-138. doi:10.1038/nrn2787

Frith, C. D., & Frith, U. (1999). Interacting minds-a biological basis. *Science*, *286*, 1692-1695. *doi:*10.1126/science.286.5445.1692

Furlanetto, T., Cavallo, A., Manera, V., Tversky, B., & Becchio, C. (2013). Through your eyes: incongruence of gaze and action increases spontaneous perspective taking. *Frontiers in Human Neuroscience*, *7*:455. doi:10.3389/fnhum.2013.00455

Gaetz, W., & Cheyne, D. (2006). Localization of sensorimotor cortical rhythms induced by tactile stimulation using spatially filtered MEG. *Neuroimage*, *30*, 899-908. doi:10.1016/j.neuroimage.2005.10.009

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593-609. doi:http://dx.doi.org/10.1093/brain/119.2.593

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396-403. doi:10.1016/j.tics.2004.07.002

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature, 415*, 755-755. doi:10.1038/415755a

Goldman, A. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. New York: Oxford University Press. doi:10.1093/0195138929.001.0001

Graziano, M. S., Hu, X. T., & Gross, C. G. (1997). Coding the locations of objects in the dark. *Science, 277,* 239-241. doi:10.1126/science.277.5323.239

Greenwald, A. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review, 77*, 73-99. doi:http://dx.doi.org/10.1037/h0028689

Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., & Nobre, A. C. (2014). Inter-and intra-individual variability in alpha peak frequency. *Neuroimage*, *92*, 46-55. doi:10.1016/j.neuroimage.2014.01.049

Hamilton, A. F. (2013). The mirror neuron system contributes to social responding. *Cortex, 49,* 2957-2959. doi:10.1016/j.cortex.2013.08.012

Hamilton, A. F. (2015). The neurocognitive mechanisms of imitation. *Current Opinion in Behavioral Sciences*, *3*, 63-67. doi:10.1016/j.cobeha.2015.01.011

Hari, R. (2006). Action–perception connection and the cortical mu rhythm. *Progress in Brain Research, 159,* 253-260. doi:10.1016/S0079-6123(06)59017-X

Heed, T., Habets, B., Sebanz, N., & Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Current Biology*, *20*, 1345-1349. doi:10.1016/j.cub.2010.05.068

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7,* 498-504. doi: 10.1016/j.tics.2003.09.006

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science, 16,* 219-222. doi: 10.1111/j.1467-8721.2007.00507.x

Herz, D. M., Christensen, M. S., Bruggemann, N., Hulme, O. J., Ridderinkhof, K. R., Madsen, K. H., & Siebner, H. R. (2014). Motivational tuning of fronto-subthalamic connectivity facilitates control of action impulses. *The Journal of Neuroscience*, *34,* 3210-3217. *doi:10.1523/JNEUROSCI.4081-13.2014*

Hess, U., & Fischer, A. (2013). Emotional mimicry as social regulation. *Personality and Social Psychology Review*, *17*, 142-157. doi:10.1177/1088868312472607

Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, *5,* 253-261. doi:10.1016/S1364-6613(00)01661-2

Heyes, C. (2009). Evolution, development and intentional control of imitation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364,* 2293-2298. doi:10.1098/rstb.2009.0049

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*, 463-483. doi:10.1037/a0022288.

Heyes, C. (2015). When does social learning become cultural learning? *Developmental science,* 1-14. doi:10.1111/desc.12350

Heyes, C. (2016). Homo imitans? Seven reasons why imitation couldn't possibly be associative. *Philosophical Transactions of the Royal Society of London B: Biological Science, 371*, 20150069. doi:10.1098/rstb.2015.0069

Heyes, C. (2016). Imitation: Not in Our Genes. *Current Biology*, *26,* 412-414. doi:10.1016/j.cub.2016.03.060

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, *22*, 233-240. doi:10.1016/j.cogbrainres.2004.09.009

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience, 21*, 1229-1243. doi:10.1162/jocn.2009.21189

Hoenig, J. M., & Heisey, D. M. (2012). The abuse of power. *The American Statistician, 55*, 19-24. doi:10.1198/000313001300339897

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849-878. doi: 10.1007/s00426-009-0234-2

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, *60*, 653-670. doi:10.1146/annurev.psych.60.110707.163604

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526-2528. doi:10.1126/science.286.5449.2526

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med, 2*: e124. http://dx.doi.org/10.1371/journal.pmed.0020124

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *Neuroimage, 31*, 429-439. doi:10.1016/j.neuroimage.2005.11.026

Jeannerod, M., & Decety, J. (1995). Mental motor imagery: a window into the representational stages of action. *Current Opinion in Neurobiology*, *5*, 727-732. doi:10.1016/0959-4388(95)80099-9

Jeannerod, M., & Frak, V. (1999). Mental imaging of motor activity in humans. *Current Opinion in Neurobiology*, *9*, 735-739. doi:10.1016/S0959-4388(99)00038-0

Jeffreys, H. (1961). *Theory of probability.* Oxford, England: Oxford University Press.

Jentzsch, I., & Leuthold, H. (2004). Beneficial effects of ambiguous precues: Parallel motor preparation or reduced premotoric processing time?. *Psychophysiology*, *41*, 231-244. doi:10.1111/j.1469-8986.2004.00155.x

Jonas, K. J., & Sassenberg, K. (2006). Knowing how to react: Automatic response priming from social categories. *Journal of Personality and Social Psychology, 90*, 709-721. doi: 10.1037/0022-3514.90. 5.709

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773-795. doi:10.2307/2291091

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. *Cognition*, *114*, 72-88. doi:10.1016/j.cognition.2009.08.015

Keysers, C., & Gazzola, V. (2014). Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 369:* 20130175. doi:10.1098/rstb.2013.0175

Kilner, J. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences, 15*, 352-357. doi:10.1016/j.tics.2011.06.005

Kilner, J. M., Baker, S. N., Salenius, S., Hari, R., & Lemon, R. N. (2000). Human cortical muscle coherence is directly related to specific motor parameters. *The Journal of Neuroscience, 20*, 8838-8845.

Kilner, J., Bott, L., & Posada, A. (2005). Modulations in the degree of synchronization during ongoing oscillatory activity in the human brain. *European Journal of Neuroscience*, *21*, 2547-2554. doi:10.1111/j.1460-9568.2005.04069.x

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*, 159-166. doi:10.1007/s10339-007-0170-2

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience, 29*, 10153-10159. *doi:10.1523/JNEUROSCI.2668-09.2009*

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S. J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, *7*, 1299-1301. doi:10.1038/nn1355

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, *29*, 169-195. doi:10.1016/S0165-0173(98)00056-3

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain Research Reviews*, *53*, 63-88. doi:10.1016/j.brainresrev.2006.06.003

Koelewijn, T., van Schie, H. T., Bekkering, H., Oostenveld, R., & Jensen, O. (2008). Motor-cortical beta oscillations are modulated by correctness of observed action. *Neuroimage*, *40*, 767-775. doi:10.1016/j.neuroimage.2007.12.018

Kokal, I., Gazzola, V., & Keysers, C. (2009). Acting together in and beyond the mirror neuron system. *Neuroimage*, *47*, 2046-2056. doi:10.1016/j.neuroimage.2009.06.010

Koski, L., Iacoboni, M., Dubeau, M. C., Woods, R. P., & Mazziotta, J. C. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology*, *89*, 460-471. doi:10.1152/jn.00248.2002

Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: social interaction modulates action simulation. *Biology Letters*, *6*, 758-761. doi:10.1098/rsbl.2010.0478

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience, 8*, 31-42. doi:10.1080/17470919.2012.694823

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science, 14,* 334-339. doi:10.1111/1467-9280.14481

Lamberts, K., Tavernier, G., & d'Ydewalle, G. (1992). Effects of multiple reference points in spatial stimulus-response compatibility. *Acta Psychologica, 79*, 115-130. doi:10.1016/0001-6918(92)90028-C

Leder, H., Tinio, P. P., Fuchs, I. M., & Bohrn, I. (2010). When attractiveness demands longer looks: the effects of situation and gender. *The Quarterly Journal of Experimental Psychology, 63*, 1858-1871. doi: 10.1080/17470211003605142

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences*, *19*, 688-699. doi:http://dx.doi.org/10.1016/j.tics.2015.08.005

Leighton, J., & Heyes, C. (2010). Hand to mouth: automatic imitation across effector systems. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1174-1183. doi: 10.1037/a0019953

Lewin, K. (1939). Field theory and experiment in social psychology: Concepts and methods. *The American Journal of Sociology, 44,* 868-896. doi:10.2307/2769418

Liepelt, R., Prinz, W., & Brass, M. (2010). When do we simulate non-human agents? Dissociating communicative and non-communicative actions. *Cognition, 115*, 426-434. doi:10.1016/j.cognition.2010.03.003

Loftus, E. F. Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus". *Law and Human Behavior, 11*, 55-62. doi: 10.1007/BF01044839

Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: Compound stimulus strategies for the explicit task-cuing procedure. *Psychonomic Bulletin & Review, 11,* 832-840. doi: 10.1037/0096-1523.29.3.575

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., et al. (2015). JASP (Version 0.7) [Computer software].

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris, 102,* 59-70. doi: 10.1016/j.jphysparis.2008.03.004

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*, 177-190. doi:10.1016/j.jneumeth.2007.03.024

Mazzarella, E., Hamilton, A., Trojano, L., Mastromauro, B., & Conson, M. (2012). Observation of another's action but not eye gaze triggers allocentric visual perspective. *The Quarterly Journal of Experimental Psychology*, *65*, 2447-2460. doi:10.1080/17470218.2012.697905

McCabe, M. A., & Uzgiris, I. C. (1983). Effects of model and action on imitation in infancy. *Merrill-Palmer Quarterly, 29*, 69-82. doi:http://www.jstor.org/stable/23086192

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, *24*, 470-476. doi:http://dx.doi.org/10.1037/0012-1649.24.4.470

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838. doi:http://dx.doi.org/10.1037/0012-1649.31.5.838

Meltzoff, A, N, & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting, 6*, 179-192. doi:10.1002/(SICI)1099-0917(199709/12)6:3/4<179::AID-EDP157>3.0.CO;2-R

Ménoret, M., Varnet, L., Fargier, R., Cheylus, A., Curie, A., Des Portes, V., ... & Paulignan, Y. (2014). Neural correlates of non-verbal social interactions: a dual-EEG study. *Neuropsychologia*, *55*, 85-97. doi:10.1016/j.neuropsychologia.2013.10.001

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & Psychophysics*, *68*, 327-337. doi:10.1016/j.cognition.2009.08.015

Molenkamp, B. (2010). UvA faceLAB IEyeTracker2 interface for Presentation (Version 1.1) [Software]. Available from http://www.fmg.uva.nl/top/home.cfm

Muthukumaraswamy, S. D., Johnson, B. W., & McNair, N. A. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research*, *19*, 195-201. doi:10.1016/j.cogbrainres.2003.12.001

Neuper, C., Wörtz, M., & Pfurtscheller, G. (2006). ERD/ERS patterns reflecting sensorimotor activation and deactivation. *Progress in Brain Research*, *159*, 211-222. doi:10.1016/S0079-6123(06)59014-4

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*, *10*, 817-818. doi:10.1038/nn1911

Oberman, L. M., McCleery, J. P., Ramachandran, V. S., & Pineda, J. A. (2007). EEG evidence for mirror neuron activity during the observation of human and robot actions: Toward an analysis of the human qualities of interactive robots. *Neurocomputing*, *70*, 2194-2203. doi:10.1016/j.neucom.2006.02.02

Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Social Cognitive and Affective Neuroscience*, *2*, 62-66. doi:10.1093/scan/nsl022

Obhi, S. S., & Sebanz, N. (2011). Moving together: toward understanding the mechanisms of joint action. *Experimental Brain Research*, *211*, 329-336. doi:10.1007/s00221-011-2721-0

Ocampo, B., & Kritikos, A. (2010). Placing actions in context: motor facilitation following observation of identical and non-identical manual acts. *Experimental Brain Research*, *201*, 743-751. doi:10.1007/s00221-009-2089-6

Ondobaka, S., de Lange, F. P., Newman-Norlund, R. D., Wiemers, M., & Bekkering, H. (2011). Interplay between action and movement intentions during social interaction. *Psychological Science*, *23*, 30-35. doi:10.1177/0956797611424163

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., ... & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology*, *26*, 1334-1338. doi:10.1016/j.cub.2016.03.047

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011,* 1-9. *doi:*10.1155/2011/156869

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Viewpoint (in) dependence of action representations: an MVPA study. *Journal of Cognitive Neuroscience*, *24,* 975-989. doi:10.1162/jocn_a_00195

Over, H., & Carpenter, M. (2012). Putting the social into social learning: explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*, 182-192. doi:10.1037/a0024555

Parker, G. A. (1974). Assessment strategy and the evolution of animal conflicts. *Journal of Theoretical Biology*, *47*, 223-243. doi: 10.1016/j.bbr.2011.03.031

Parsons, L. M., Fox, P. T., Downs, J. H., Glass, T., Hirsch, T. B., Martin, C. C., ... & Lancaster, J. L. (1995). Use of implicit motor imagery for visual shape discrimination as revealed by PET. *Nature, 375*, 54-58. doi:10.1038/375054a0

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181-192. doi: 10.1037/0022-3514.81.2.181.

Pellizzer, G., & Hedges, J. H. (2003). Motor planning: effect of directional uncertainty with discrete spatial cues. *Experimental Brain Research*, *150,* 276-289. doi:10.1007/s00221-003-1453-1

Pfurtscheller, G., Graimann, B., Huggins, J. E., Levine, S. P., & Schuh, L. A. (2003). Spatiotemporal patterns of beta desynchronization and gamma synchronization in corticographic data during self-paced movement. *Clinical Neurophysiology*, *114*, 1226-1236. doi:10.1016/S1388-2457(03)00067-1

Piaget, J. (1930). *The child's conception of physical causality.* (Trans. by M. Gabain.) London: Kegan Paul, Trench, Trubner, & co., Ltd.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge.

Pineda, J. A. (2005). The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing”. *Brain Research Reviews, 50*, 57-68. doi:10.1016/j.brainresrev.2005.04.005

Praamstra, P., Kourtis, D., & Nazarpour, K. (2009). Simultaneous preparation of multiple potential movements: opposing effects of spatial proximity mediated by premotor and parietal cortex. *Journal of Neurophysiology*, *102*, 2084-2095. doi:10.1152/jn.00413.2009

Press, C., & Cook, R. (2015). Beyond action-specific simulation: domain-general motor contributions to perception. *Trends in Cognitive Sciences, 19*, 176-178. doi:http://dx.doi.org/10.1016/j.tics.2015.01.006

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann, & W. Prinz (Eds.), *Relationships between Perception and Action: Current Approaches* (pp. 167-201). Springer Berlin Heidelberg.

Prinz, W., Aschersleben, G., & Koch, I. (2009). Cognition and action. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford Handbook of Human Action* (pp. 35-71). Oxford: Oxford University Press.

Ramachandran, V. S. (1995). Decade of the Brain lecture at the Silver Jubilee meeting of the Society for Neuroscience. Retrieved from https://www.edge.org/conversation/mirror-neurons-and-imitation-learning-as-the-driving-force-behind-the-great-leap-forward-in-human-evolution

Ramnani, N., & Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nature Neuroscience*, *7*, 85-90. doi:10.1038/nn1168

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2,* 79-87. doi:10.1038/4580

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental Science, 14*, 92-105. doi:10.1111/j.1467-7687.2010.00961.x

Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye movements and visual encoding during scene perception. *Psychological Science, 20,* 6-10. doi: 10.1111/j.1467-9280.2008.02243.x

Reeb-Sutherland, B. C., Levitt, P., & Fox, N. A. (2012). The predictive nature of individual differences in early associative learning and emerging social behavior. *PLoS One, 7:* e30511. doi:http://dx.doi.org/10.1371/journal.pone.0030511

Richardson, D. C., Hoover, M. A., & Ghane, A. (2008). Joint perception: gaze and the presence of others. *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 309-314). Austin, TX: Cognitive Science Society.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Reviews of Neuroscience, 27*, 169-192. doi:10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science, 277*, 190-191. doi:10.1126/science.277.5323.190

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2,* 661-670.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16,* 225-237. doi:10.3758/PBR.16.2.225.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1255-1266. doi:10.1037/a0018729.

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing?. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 929. doi:10.1037/a0035175

Sartori, L., Begliomini, C., Panozzo, G., Garolla, A., & Castiello, U. (2014). The left side of motor resonance. *Frontiers in Human Neuroscience*, *8*: 702. doi:http://dx.doi.org/10.3389/fnhum.2014.00702

Sartori, L., & Betti, S. (2015). Complementary actions. *Frontiers in Psychology*, *6*: 557. doi:http://dx.doi.org/10.3389/fpsyg.2015.00557

Sartori, L., Betti, S., Chinellato, E., Castiello, U. (2015). The multiform motor cortical output: Kinematic, predictive and response coding. *Cortex*, *70,* 169-178. doi:10.1016/j.cortex.2015.01.019

Sartori, L., Bucchioni, G., & Castiello, U. (2013). When emulation becomes reciprocity. *Social Cognitive and Affective Neuroscience*, *8*, 662-669. doi:10.1093/scan/nss044

Sartori, L., Cavallo, A., Bucchioni, G., & Castiello, U. (2011). Corticospinal excitability is specifically modulated by the social dimension of observed actions. *Experimental Brain Research*, *211*, 557-568. doi:10.1007/s00221-011-2650-y

Sartori, L., Cavallo, A., Bucchioni, G., & Castiello, U. (2012). From simulation to reciprocity: the case of complementary actions. *Social Neuroscience*, *7*, 146-158. doi:10.1080/17470919.2011.586579

Schilbach, L. (2014). On the relationship of online and offline social cognition. *Frontiers in Human Neuroscience*, *8*:278. doi:http://dx.doi.org/10.3389/fnhum.2014.00278

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: action-induced modulation of perception. *Trends in Cognitive Sciences*, *11*, 349-355. doi:10.1016/j.tics.2007.06.005

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70-76. doi:10.1016/j.tics.2005.12.009

Sebanz, N., & Frith, C. (2004). Beyond simulation? Neural mechanisms for predicting the actions of others. *Nature Neuroscience*, *7*, 5-6. doi:10.1038/nn0104-5

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of Cognitive Neuroscience, 18*, 859-870. doi:10.1162/jocn.2006.18.5.859

Seeing Machines (2009). faceLAB (Version 5) [Hardware and software]. Available from www.seeingmachines.com

Shepherd, S. V. (2010). Following gaze: gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, *4*: 5. doi:10.3389/fnint.2010.00005

Shibata, T., Kim, J., Hoffman, D. M., & Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, *11*, 1-29. doi:10.1167/11.8.11

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *17*, 1-8. doi:10.1177/0956797611417632

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews, 30*, 855-863. doi:10.1016/j.neubiorev.2006.06.011

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, *20*, 2037-2057. doi:10.1162/jocn.2008.20143

Smith, E. R., & Semin, G. R. (2007). Situated social cognition. *Current Directions In Psychological Science*, *16*, 132-135. doi: 10.1111/j.1467-8721.2007.00490.x

Southgate, V., Johnson, M. H., El Karoui, I., & Csibra, G. (2010). Motor system activation reveals infants' on-line prediction of others' goals. *Psychological Science*, *21*, 355-359. doi:10.1177/0956797610362058

Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009). Predictive motor activation during action observation in human infants. *Biology Letters*, *5*, 769-772. doi:10.1098/rsbl.2009.0474

Spieser, L., van den Wildenberg, W., Hasbroucq, T., Ridderinkhof, K. R., & Burle, B. (2015). Controlling your impulses: Electrical stimulation of the human supplementary motor complex prevents impulsive errors. *The Journal of Neuroscience, 35,* 3010-3015. doi:10.1523/JNEUROSCI.1642-14.2015.

Stančák, A., Riml, A., & Pfurtscheller, G. (1997). The effects of external load on movement-related changes of the sensorimotor EEG rhythms. *Electroencephalography and Clinical Neurophysiology*, *102*, 495-504. doi:10.1016/S0013-4694(96)96623-0

Stapel, J. C., Hunnius, S., van Elk, M., & Bekkering, H. (2010). Motor activation during observation of unusual versus ordinary actions in infancy. *Social Neuroscience, 5*, 451-460. doi:10.1080/17470919.2010.490667

Streltsova, A., Berchio, C., Gallese, V., & Umiltà, M. A. (2010). Time course and specificity of sensory-motor alpha modulation during the observation of hand motor acts and gestures: a high density EEG study. *Experimental Brain Research*, *205*, 363-373. doi:10.1007/s00221-010-2371-7

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1746-1759. doi:http://dx.doi.org/10.1037/0096-1523.26.6.1746

Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, *129*, 426-438. doi:10.1016/j.cognition.2013.06.008

Surtees, A. D., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*, 75-86. doi: 10.1111/j.2044-835X.2011.02063.x

Tan, H., Wade, C., & Brown, P. (2016). Post-movement beta activity in sensorimotor cortex indexes confidence in the estimations from internal models. *The Journal of Neuroscience, 36,* 1516-1528. doi:10.1523/JNEUROSCI.3204-15.2016

Ter Horst, A. C., van Lier, R., & Steenbergen, B. (2010). Mental rotation task of hands: differential influence number of rotational axes. *Experimental Brain Research*, *203*, 347-354. doi:10.1007/s00221-010-2235-1

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103 –130). Hillsdale, NJ: Lawrence Erlbaum.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*, 495-511. doi:10.1017/S0140525X0003123X

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 830-846. doi:http://dx.doi.org/10.1037/0096-1523.24.3.830

Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, *4*, 173-187. doi**:**10.1207/s15326969eco0403_3

Tversky, B. (2005). Functional significance of visuospatial representations. In P. Shah & A. Miyake (Editors.), *Handbook of higher-level visuospatial thinking* (1-34). Cambridge: Cambridge University Press.

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*, 124-129. doi:10.1016/j.cognition.2008.10.008

Tzagarakis, C., Ince, N. F., Leuthold, A. C., & Pellizzer, G. (2010). Beta-band activity during motor planning reflects response uncertainty. *The Journal of Neuroscience, 30*, 11270-11277. doi:10.1523/JNEUROSCI.6026-09.2010

Tzagarakis, C., West, S., & Pellizzer, G. (2015). Brain oscillatory activity during motor preparation: effect of directional uncertainty on beta, but not alpha, frequency band. *Frontiers in Neuroscience*, *9*:246. doi:http://dx.doi.org/10.3389/fnins.2015.00246

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, *31*, 155-165. doi:10.1016/S0896-6273(01)00337-3

Urgesi, C., Maieron, M., Avenanti, A., Tidoni, E., Fabbro, F., & Aglioti, S. M. (2010). Simulating the future of actions in the human corticospinal system. *Cerebral Cortex, 20*, 2511-2521. doi:10.1093/cercor/bhp292

Uzgiris, I. C. (1981). Two functions of imitation during infancy. *International Journal of Behavioral Development, 4*, 1-12. doi: 10.1177/016502548100400101

Valdés-Conroy, B., Sebastián, M., Hinojosa, J. A., Román, F. J., & Santaniello, G. (2014). A close look into the near/far space division: A real-distance ERP study. *Neuropsychologia*, *59*, 27-34. doi:10.1016/j.neuropsychologia.2014.04.009

van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S. A. (2012). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping*, *33,* 2322-2333. doi:10.1002/hbm.21365

van Der Schalk, J., Fischer, A., Doosje, B., Wigboldus, D., Hawk, S., Rotteveel, M., & Hess, U. (2011). Convergent and divergent responses to emotional displays of ingroup and outgroup. *Emotion*, *11*, 286-298. doi:http://dx.doi.org/10.1037/a0022582

van Elk, M., Bousardt, R., Bekkering, H., & van Schie, H. T. (2012). Using goal-and grip-related information for understanding the correctness of other's actions: an ERP study. *PloS One*, *7*: e36450. http://dx.doi.org/10.1371/journal.pone.0036450

van Elk, M., van Schie, H. T., & Bekkering, H. (2009). Action semantic knowledge about objects is supported by functional motor activation. *Journal of Experimental Psychology: Human Perception and Performance*, *35,* 1118-1128. doi:10.1037/a0015024

van Elk, M., van Schie, H. T., van Den Heuvel, R., & Bekkering, H. (2010). Semantics in the motor system: motor-cortical beta oscillations reflect semantic knowledge of end-postures for object use. *Frontiers in Human Neuroscience, 4:* 8. doi:10.3389/neuro.09.008.2010

van Pelt, S., Heil, L., Kwisthout, J., Ondobaka, S., van Rooij, I., & Bekkering, H. (2016). Beta-and gamma-band activity reflect predictive coding in the processing of causal events. *Social Cognitive and Affective Neuroscience*, *11,* 973-980. doi:10.1093/scan/nsw017 van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2008). Understanding action beyond imitation: Reversed compatibility effects of action observation in imitation and joint action. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1493–1500. doi:10.1037/a0011750

Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, *7*, 38-42. doi:10.1016/S1364-6613(02)00003-7

Wang, Y., & Hamilton, A. F. D. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. Towards a neuroscience of social interaction. *Frontiers in Human Neuroscience*, *6*:153. 10.3389/fnhum.2012.00153

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of The Royal Society of London B: Biological sciences, 364*, 2417-2428. doi:10.1098/rstb.2009.0069

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer Publishing Company, Incorporated.

Willer, S. (2009). "Imitation of Similar Beings": Social Mimesis as an Argument in Evolutionary Theory around 1900. *History and philosophy of the Life Sciences*, *31*, 201-213.

Wilson, A. (2014, October 8). Limits on action priming by pictures of objects. Retrieved from http://psychsciencenotes.blogspot.nl/2014/10/limits-on-action-priming-by-pictures-of.html

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, *4*:58. doi:http://dx.doi.org/10.3389/fpsyg.2013.00058

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological bulletin*, *131*, 460-473. doi:http://dx.doi.org/10.1037/0033-2909.131.3.460

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 358*, 593-602. doi:10.1098/rstb.2002.1238

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317-1329. doi:10.1016/S0893-6080(98)00066-5

Yu, A. B., Abrams, R. A., & Zacks, J. M. (2014). Limits on action priming by pictures of objects. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1861-1873. doi:http://dx.doi.org/10.1037/a0037397

Zaepffel, M., Trachel, R., Kilavik, B. E., & Brochier, T. (2013). Modulations of EEG beta-power during planning and execution of grasping movements. *PloS One*, *8*: e60060, doi:http://dx.doi.org/10.1371/journal.pone.0060060

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes, 28*, 81-88. doi: 10.1080/01638539909545070

Zwickel, J., & Müller, H. J. (2010). Observing fearful faces leads to visuo-spatial perspective taking. *Cognition*, *117*, 101-105. doi:10.1016/j.cognition.2010.07.004

Articles

Bélanger, J. J., Faber, T., & Gelfand, M. J. (2013). Supersize my identity: when thoughts of contracting swine flu boost one's patriotic identity. *Journal of Applied Social Psychology, 43*, 153-155. doi: 10.1111/jasp.12032

Bramson, B., Faber, T. W., & Jonas, K. J. (in prep). *Giving replication effects a fair chance: An example and suggestions for a contingency check for direct replications*. Manuscript submitted for publication.

Faber, T. W., van Elk, M., & Jonas, K. J. (2016). Complementary Hand Responses Occur in Both Peri-and Extrapersonal Space. *PloS one, 11:* e0154457. doi:http://dx.doi.org/10.1371/journal.pone.0154457

Faber, T. W., & Jonas, K. J. (2013). Perception in a social context: Attention for response-functional means. *Social Cognition*, *31*, 301-314. doi:10.1521/soco.2013.31.2.301

Köpetz, C., Faber, T., Fishbach, A., & Kruglanski, A. W. (2011). The multifinality constraints effect: how goal multiplicity narrows the means set to a focal end. *Journal of personality and social psychology*, *100*, 810-826. doi:10.1037/a0022980

Author contributions :

Faber and Jonas (2013; Social Cognition): data collection (TF), data analysis (TF), writing paper (TF,KJ)

Faber et al. (2016; PloS one): data collection (TF), data analysis (TF), writing paper (TF,MvE,KJ)

Nederlandse samenvatting

Imitatiegedrag wordt door veel wetenschappers gezien als drijfveer van menselijke evolutie en als belangrijke factor in de ontwikkeling van cultuur (Legare & Nielsen, 2015). Je leert er niet alleen door fietsen, in de rij te staan in de supermarkt, maar imitatiegedrag kan ook invloed hebben op je kledingstijl of zelfs bepalen wat voor werk je doet. Het (letterlijk) kopiëren van andermans gedrag lijkt simpel, maar is vrij complex. Een voorwaarde voor imitatie is namelijk het leren dat wat andere mensen doen gelijk is aan wat jezelf doet of kan doen (bijv. zien dat iemand z'n rechterhand beweegt – zelf je rechterhand bewegen). Vooral in de eerste levensjaren levert dit een probleem op. Wanneer je gedrag van anderen ziet (bijv. je moeder die naar je glimlacht als je in de wieg ligt), zie je dit vanuit tegenovergesteld perspectief (vertoond door een ander) en zie je alleen de visuele gevolgen van het gedrag. Tegelijkertijd wordt het zelf produceren van gedrag (bijv. glimlachen; je hand bewegen) eerst bepaald door een motorcommando waardoor spieren worden aangespannen en zie je vervolgens de handeling vanuit je eigen perspectief (of zelfs helemaal niet). Deze discrepantie leidt tot de volgende vraag: Hoe weet je welke spieren te gebruiken als je alleen maar de visuele gevolgen ziet van andermans gedrag vanuit een omgekeerd perspectief? Dit is het zogenaamde *correspondence problem*. Wat dit nog ingewikkelder maakt, is dat tot een bepaalde leeftijd kinderen niet beschikken over zelfherkenning of het perspectief in kunnen nemen van een ander. Toch kunnen zelfs onder deze condities kinderen vrij snel leren imiteren.

Er is al tijden een debat gaande over het mechanisme dat dit correspondence problem oplost dat sterk verwant is aan het nature-nurture debat. Waarbij sommige wetenschappers stellen dat mensen geboren zijn met een ingebouwde module dat gedrag van anderen 'vertaalt' naar de productie van dezelfde handeling (Meltzoff, 1988), denkt een andere groep dat er meer algemene mechanismes aangeboren zijn die zowel imitatie alsook ander type gedrag kunnen faciliteren als gevolg van associatieve leerprocessen (Heyes, 2016). Deze leerprocessen zijn gebaseerd op een vrij simpele regel waarbij het regelmatig zien en uitvoeren van handelingen associatieve koppelingen vormen. Recent empirisch onderzoek en een her analyse van 20 jaar aan babyonderzoek ondersteunt de laatste visie en toont aan dat er geen bewijs is dat pasgeboren baby's kunnen imiteren en dus naar waarschijnlijkheid niet geboren zijn met een zogenaamde 'transformatie' module. Wel kunnen baby's leren om te imiteren doordat ze geïmiteerd worden door anderen (voornamelijk ouders), door in de spiegel te kijken of simpelweg hun eigen handen te volgen als ze die bewegen. Deze ervaringen creëren associatieve koppelingen tussen visuele en motorgebieden in de hersenen die over tijd ervoor zorgen dat het zien van gedrag van anderen imitatiegedrag stimuleert. Bijvoorbeeld, als elke keer als een baby lacht een ouder terug lacht, leert de baby over tijd de productie van zijn eigen lach te koppelen aan het zien van een glimlach, wat vervolgens imitatiegedrag kan stimuleren. Door het belang van ervaring zijn deze associatieve koppelingen sterker voor gedrag waar je meer ervaring mee hebt. Je kunt het gevoel hebben mee te bewegen als je bijvoorbeeld als balletdanseres een balletoptreden ziet, als je als basketballer naar een basketbalwedstrijd kijkt, maar misschien minder wanneer je als basketballer een balletoptreden ziet.

Hoewel imitatie belangrijk is in verschillende omstandigheden, is het vooral een passief proces. Daarnaast is imitatie in veel situaties geen aangeleerde vorm van gedrag. Bijvoorbeeld, als iemand je een bal gooit, vang je deze, als iemand je glas inschenkt, breng je je glas naar voren en als iemand huilt, is troosten een meer gebruikelijke reactie. Het mooie aan associatieve leerprocessen is dat dit type gedrag (complementair gedrag) wordt geautomatiseerd met behulp van dezelfde mechanismes die imitatie tot stand brengen. Omdat modules in de hersenen die hierbij een rol spelen niet speciaal ontwikkeld zijn voor imitatie kunnen koppelingen tussen het zien en produceren van ongelijke handelingen op dezelfde manier worden aangeleerd als imitatiegedrag. Hoewel het herhaaldelijk zien van een lachende moeder wanneer een kind glimlacht een gelijke koppeling creëert, zijn er dus tal van complementaire situaties waarbij een koppeling niet gelijk is (bijv. een bal gooien – vangen). Niettemin wordt onderzoek binnen (sociale) psychologie gedomineerd door imitatie en in het specifiek de automatische koppeling tussen het zien en uitvoeren van gelijke (identieke) handelingen.

In de huidige dissertatie hebben we gekeken naar de distinctie tussen imitatie en complementair gedrag en we beargumenteren dat imitatie in veel gevallen niet aangeleerd is en niet functioneel is in voornamelijk sociale (interactieve) situaties. De eerste stap was om gebruik te maken van klassieke studies in het veld en zo te kijken naar de grenzen van imitatiegedrag binnen een dergelijke studie. Een voorbeeld is een serie studies die heeft laten zien dat afbeeldingen van stereotype groepen associatief gedrag oproept dat past bij gedrag van groepsleden (bijv. crimineel – agressie). Het eerste hoofdstuk gaat hier verder op in en beschrijft een tweetal studies waar gebruik gemaakt werd van een taak waarbij mensen een serie afbeeldingen te zien kregen op de computer die ze vervolgens moesten onthouden. Deze afbeeldingen waren steeds afbeeldingen van verschillende stereotype situaties (donker steegje, rechtszaal) gevolgd door een woord dat een bepaald categorie-lid (crimineel) moest voorstellen die je in de situatie tegen zou kunnen komen. Vervolgens werden er objecten getoond op het scherm die te maken hadden met de persoon, maar verschillende functies hadden in elke situatie. Bijvoorbeeld, een pistool is geassocieerd met een crimineel, maar heeft een andere functie in een rechtszaal dan in een donker steegje. De vraag was of participanten meer aandacht zouden geven aan objecten als deze beter pasten bij de situatie waarin ze een persoon zagen. Het zou namelijk ook kunnen dat participanten alleen zouden denken aan het gedrag van de persoon in beeld ongeacht de situatie. Er werd tijdens de taak gebruik gemaakt van een gaze-tracker waarmee we konden bepalen hoeveel aandacht iemand had voor bepaalde objecten. De resultaten lieten inderdaad zien dat mensen een voorkeur hadden voor objecten die pasten bij een persoon afhankelijk van de situatie waaraan deze werd gekoppeld (donkere steeg – crimineel – pistool). Dit suggereert dat wanneer mensen geen contextuele informatie hebben, ze zich laten leiden door gedrag van een ander, maar wanneer ze deze informatie wel hebben, ze aandacht hebben voor objecten die voor hen relevant zijn in deze situatie.

In het eerste hoofdstuk was visuele aandacht de hoofdmaat. We waren ook benieuwd naar fysiek gedrag om een directer beeld te krijgen van de grenzen van fysieke imitatie. In het tweede hoofdstuk werd gebruik gemaakt van een computertaak waarbij participanten een serie afbeeldingen zagen van een persoon die zijn linker- of rechterhand uitstak. Deze hand was of een gesloten handbeweging of een open handbeweging. De persoon in beeld zat steeds dichtbij, aan de overkant van de korte kant van een tafel, of ver weg, aan de overkant van de lange kant van een tafel. Afhankelijk van de kleur van de hand (deze was in de helft van de gevallen groen) moesten participanten de hand in beeld nadoen in spiegelbeeld of met hun tegenovergestelde hand. De gedachte was dat bij het zien van een open hand, mensen automatisch geneigd zouden zijn hun tegenovergestelde (complementaire) hand te gebruiken in plaats van hun hand in spiegelbeeld. Echter, omdat je pas iemand pas de hand kunt schudden als deze binnen je persoonlijke ruimte is, werd dit niet verwacht als de persoon verder weg zat. Door de tijd te meten hoe lang participanten erover deden om een handbeweging te maken, konden we zien of ze sneller waren bij het maken van een open hand of gesloten hand en of ze sneller waren als de persoon in beeld veraf of dichtbij zat. Vijf studies moesten hierover uitsluitsel geven en lieten vrij consistent geen effect zien van afstand. Wel vonden we dat mensen consequent sneller waren met het vertonen van een complementaire hand vergeleken met hun hand in spiegelbeeld als een open hand werd vertoond. Voor gesloten handen was dit effect er niet en leek het erop dat mensen juist sneller waren in het maken van een handbeweging in spiegelbeeld, wat duidt op een imitatie effect. Deze set van studies laat zien dat complementair gedrag automatisch vertoond kan worden net als imitatiegedrag. In tegenstelling tot het eerste hoofdstuk lijkt het erop dat dit automatisme zo sterk is dat afstand geen rol speelt. Dit kan mogelijk verklaren waarom je het gevoel kan hebben mee te bewegen als je volledig in een wedstrijd zit of een computerspel aan het spelen bent, zonder dat je fysiek in de situaties aanwezig bent.

Het derde hoofdstuk gaat verder in op de rol van perspectief nemen zoals beschreven in het begin van deze samenvatting. Imitatie wordt vaak gekoppeld aan perspectief nemen omdat het kopiëren van andermans gedrag vaak gepaard gaat met het bepalen hoe het voor een ander is om bepaald gedrag te vertonen. Zoals eerder aangegeven, is imitatie op deze manier voornamelijk een attribuut van passieve situaties. In interactie met anderen en bij het vertonen van complementair gedrag is het nemen van een eigen perspectief (1^{ste} persoon perspectief) echter meer van belang. Je moet ten slotte weten hoe je een bal vangt vanuit je eigen perspectief, het perspectief van degene die de balt gooit is dan minder relevant. Toch heeft een aantal studies laten zien dat het perspectief innemen van een ander belangrijk is wanneer gedrag van anderen mogelijk relevant is voor jezelf. Dit is aannemelijk, het zou namelijk vermoeiend zijn om consequent ieders perspectief in te moeten nemen of willekeurig iedereen te imiteren die je tegenkomt. Om dit punt te ondersteunen hebben dezelfde onderzoekers echter vaak gebruik gemaakt van passieve, in plaats van interactieve situaties. In het derde hoofdstuk bespreken we drie studies met eenzelfde type taak, maar waarbij een interactieve situatie is toegevoegd. Participanten zagen een afbeelding van een persoon recht tegenover in beeld met twee objecten (boek en glas) voor hem liggen op een tafel. Gevraagd werd om te bepalen waar het boek lag ten opzichte van het glas. Vanuit je eigen perspectief zou dit bijvoorbeeld links kunnen zijn, maar vanuit het perspectief van de persoon is dit tegenovergesteld (rechts). Door kleine veranderingen aan te brengen in deze scène konden we een passieve scène, waarbij de persoon vooruit keek zonder de objecten aan te raken, vergelijken met een interactieve scène waarin de persoon één van de objecten oppakte en naar voren reikte. In drie studies keken we vervolgens hoe snel participanten in konden schatten waar het boek lag ten opzichte van het glas en hoe vaak ze spontaan het perspectief innamen van de persoon in beeld afhankelijk van de type scène. De hypothese was dat in plaats van dat interactieve scènes ervoor zouden zorgen dat mensen vaker het perspectief zouden innemen van een ander, dit juist lastiger zou zijn in interactieve scènes omdat deze mensen dwingen een eerste-persoonsperspectief in te nemen. Dit zou betekenen dat als participanten gevraagd werd om het perspectief in te nemen van een ander in een interactieve scène, dit meer moeite zou kosten en dus tot meer fouten en langere reactietijden zou leiden vergeleken met passieve scènes. Twee van de drie studies lieten dit zien, al leek het erop dat perspectief nemen in interactieve scènes lastiger was ongeacht of mensen hun eigen perspectief of die van de persoon in de scène moesten innemen.

Het laatste hoofdstuk gaat iets verder en kijkt naar wat er precies gebeurt in de hersenen bij het zien van gedrag dat tot imitatie kan leiden. Door herhaaldelijk gedrag te zien van anderen net na of voordat je zelf bepaald gedrag vertoond, kan dit leiden tot automatisch imitatie of complementair gedrag. Op hersenniveau betekent dit dat het zien van bepaald gedrag leidt tot activiteit in dezelfde motorgebieden die ook betrokken zijn bij het produceren van dit type gedrag. Een aantal experimenten suggereert dat bepaalde neuronen, zogenaamde spiegelneuronen, dit mechanisme faciliteren. Een consequentie van het herhaaldelijk koppelen van visuele informatie en fysiek gedrag is dat deze koppeling over tijd een voorspellende relatie wordt. Als je voldoende informatie hebt of weet wat er gaat gebeuren, is het mogelijk om gedrag te simuleren al voordat je visuele informatie binnen krijgt. Bijvoorbeeld, als je een dierentuin inloopt, ben je meer ontspannen wanneer je langs de tijgerkooi loopt dan als je op safari bent. Omgevingsinformatie helpt bij het voorspellen van wat er komen gaat en dit kan op neuraal niveau effect hebben in motorgebieden nog voor er een verandering plaats heeft plaatsgevonden. Het vierde hoofdstuk beschrijft een studie die gebruik maakt van EEG (elektro-encefalogram) om te kijken naar de rol van motorgebieden in de hersenen bij het zien van simpele handbewegingen. Participanten zagen een serie van korte filmpjes die in verschillende mate voorspelbaar waren. Afhankelijk van de kleur van de hand aan het begin van het filmpje konden ze bijvoorbeeld voorspellen of de hand een precisie beweging (met een duim en wijsvinger) zou maken naar een object toe of een open handbeweging. Voor sommige kleuren was de opkomende handeling sterk voorspelbaar en voor andere kleuren was de uitkomst niet voorspelbaar. Onze verwachting was dat mensen al zouden simuleren wat de uitkomst van een handeling zou zijn voordat er enige beweging had plaatsgevonden, maar alleen als ze vrijwel zeker zouden weten wat er zou komen. Daarnaast verwachtten we dat handelingen die niet voorspeld konden worden juist zouden leiden tot sterkere activiteit in motorgebieden tijdens het zien van de handeling. Met behulp van een time-frequency analyse, waarbij er wordt gekeken naar de mate van synchronisatie waarin verschillende clusters van neuronen een signaal afgeven, vergeleken we video's die beter voorspelbaar en minder voorspelbaar waren met elkaar. We vonden dat motorgebieden sterker actief waren wanneer handelingen sterk voorspelbaar waren vergeleken met niet-voorspelbare handelingen nog voordat er een handbeweging te zien was. Tijdens het zien van de handbeweging was er geen verschil in motor activatie tussen voorspelbare en niet-voorspelbare bewegingen. Wel konden we een distinctie maken tussen situaties waarin voorspelbare handelingen werden gevolgd door een onverwachte en een verwachte handeling. Ook was het zien van een onverwachte uitkomst verschillend van het zien van een niet-voorspelbare beweging, wat suggereert dat participanten waarschijnlijk niet gokten wat de uitkomst zou zijn voordat de beweging werd getoond.

De bevindingen in deze dissertatie hebben een aantal dingen aangetoond: a) contextuele informatie bepaald aandacht voor objecten die relevant zijn voor complementair gedrag; b) complementair gedrag kan automatisch worden vertoond net als imitatie en is onafhankelijk van interpersoonlijke afstand, c) interactieve situaties maken perspectief nemen lastiger, en d) motorprocessen hebben voorspellende eigenschappen en maken het mogelijk voor mensen om gedrag van anderen te anticiperen. Alhoewel de bevindingen gevarieerd zijn en niet allemaal even goed aansluiten op de hoofdvraag, laten ze zien dat het model waarbij het zien van gedrag automatisch hetzelfde gedrag stimuleert (fysiek of in de hersenen) te simplistisch is. Afhankelijk van contextuele informatie, de mate waarin gedrag van anderen passief wordt bekeken of actief wordt gebruikt en de mate waarin gedrag van anderen voorspelbaar is, is complementair gedrag vaak nuttiger dan imitatiegedrag. Dit heeft gevolgen voor een grote selectie aan theorieën die zich beperken tot de koppeling tussen het zien van gedrag en het automatisch kopiëren van hetzelfde gedrag. Ook heeft dit consequenties voor de rol van spiegelneuronen die in eerdere opvattingen geobserveerd gedrag vertalen in hetzelfde gedrag. Deze vertaling wordt gezien als een voorwaarde voor het begrijpen van andermans gedrag en zou de bron zijn van imitatie en zelfs empathie. Nieuwe bevindingen laten zien dat spiegelneuronen reageren op een bijzonder breed palet aan type handelingen, ook wanneer deze geen bepaald doel hebben.

Onderzoekers uit verschillende labs die hebben gekeken naar dezelfde type processen hebben vergelijkbare bevindingen gedaan en hebben aangetoond hoe imitatie en complementair gedrag elkaar afwisselen over tijd in interactieve situaties. Deze dissertatie heeft daar (hopelijk) deels aan bijgedragen.

When imitation falls short: The case of complementary actions

Imitation is seen by many researchers as the driving force of human evolution and as a primary factor controlling the development of culture (Legare & Nielsen, 2015). Imitation allows you to learn to ride a bike and to stay in line at the supermarket but it also affects your choice of clothing and can even determine what type of work you do. Imitation, or copying behavior displayed by others, might look simple at first glance but is in fact quite complex. One precondition for imitation is that you have to learn that what others do is similar to what you yourself are doing or are able to do (e.g., seeing a person's right hand move – moving your own right hand). Specifically in the first few years of infancy this is problematic. When you see behavior displayed by somebody else (e.g., your mother smiling at you), you will see this behavior from a flipped perspective (as displayed by the mother) and you will only see the visual effects of the displayed behavior. At the same time, producing behavior yourself (e.g., smiling; moving your hand), is determined by a motor command which directs muscle movements that subsequently provide visual feedback from a first person perspective (or no feedback at all). This discrepancy provides the following question: How do you know what muscles to use if you only observe the visual effects of behavior performed by others, seen from a mirror perspective? This problem is called the *correspondence problem*. What complicates this even more is that until a certain age infants are not able to self-identify or take the perspective of others. But even under these conditions infants are able to show imitative behavior from an early age.

There is an ongoing debate about the mechanisms that solves the correspondence problem which is strongly connected to the nature-nurture debate. Where some researchers argue that people are born with a module that 'translates' observed behavior into the production of the same behavior (Meltzoff, 1988), a different group of researchers argues that people are born with a general purpose mechanism that can produce imitation as well as different types of behavior as a consequence of associative learning processes (Heyes, 2016). These learning processes are based on a simple rule that specifies that repeated co-occurrence of observed and performed behavior creates associative links. Recent empirical research and a re-analysis of 20 years of baby research support the latter view and shows that there is no strong evidence for an inborn module that facilitates matching (imitative) connections. The findings do support the idea that baby's learn by being imitated by others (e.g., parents), by looking in the mirror or by observing their own hands while moving them. These experiences create associative links between neural regions that code the visual effects of behavior and regions involved in motor control which can stimulate imitative behavior over time. For example, if for every time a baby smiles, a parent smiles back, the baby will over time associate the production of a smile to the observation of a smile (e.g., seeing the corners of the mouth move upwards) which accordingly can produce imitative behavior. The intensity and frequency of these experiences strengthen associative connections. For example, you can feel that you move along when observing a ballet performance if you are a ballet-dancer yourself or as a basketball player watching a basketball game but maybe to a lesser degree when observing a ballet performance as a basketball player.

Even though imitation is important in a lot of situations, it is primarily a passive process. More importantly, imitation is in the majority of social situations not a fitting type of behavior. For example, when someone throws a ball you catch it, if somebody pours you a drink you bring your glass forward, when somebody cries, comforting him or her is a more fitting type of response. The beauty of associative learning processes is that this type of behavior (complementary actions) is learned using the same mechanism that produces imitation. Because neural modules involved in shaping associative links are not specifically designed for imitation, linking behavior that is either congruent (similar) to observed behavior (imitation) or incongruent to observed behavior (complementary actions) can be a result of the same learning process. Even though the repeated coupling of seeing a laughing mother when a baby smiles constitute a congruent association (imitation), there are a multitude of situations in which people learn incongruent associations (e.g., throwing a ball – catching it). Nonetheless, research within (social) psychology is dominated by imitation and specifically the automatic coupling between observing and performing congruent (identical) behavior.

In the current dissertation we have looked at the distinction between imitation and complementary actions and we argue that in a host of situations imitation is not a product of repeated learning and is not functional in predominantly social (interactive) settings. The first step was to use classic studies in the field and to look at the boundary effects of imitation behavior within each study. An example is a line of research that has shown how pictures of stereotypical groups trigger associative behavior related to this group (e.g., criminal – aggressive behavior). The first chapter builds on this example and describes two studies in which we used a task where participants saw pictures of different stereotypical situations (dark alley, court room), followed by a word representing a social category member (criminal) that was associated with the situation. Subsequently, we showed pictures of objects associated with the social category that had different functions depending on the situation. For example, a gun is associated with a criminal but has different functions in a court room compared to a dark alley. Our question was whether participants would pay attention to objects only if the object was seen as fitting given the situation and social category. Alternatively, participants would primarily be driven by the association between the social category and the object irrespective of the situation. During the task we used gaze-tracking to track the amount of attention devoted to each of the objects. The results showed that participants indeed paid more attention to objects associated with the social category but differently so depending on the situation (dark alley – criminal – gun). This suggests that when people do not have any situational information to start with, they might be driven by behavior displayed by others but when they do have this information they pay attention to objects that are useful as a means of response in that specific situation.

In the first chapter visual attention was the main dependent variable. We were also interested in physical, overt behavior to reach a better understanding of the boundaries of overt imitation. In the second chapter we used a computer task in which participants saw a series of pictures of a person extending either his right or left hand. This hand was either a closed hand (fist) or an open hand movement. The person extending his hand was either sitting close by, across from the short side of a table, or far away, across from the far side of a table. Depending on the color of the hand (this was green in fifty percent of the images), participants had to copy the exact hand movement in the image in mirror perspective (e.g., seeing a right closed hand and making a left closed hand movement) or using their opposite hand. The idea was that when seeing an open hand movement, people would automatically feel the urge to use their opposite (complementary) hand instead of their hand in mirror perspective. However, because you can only shake somebody's hand if this person is within reach, this effect was not expected if the person was out of reach. By measuring how long it takes to produce a hand movement, we could see if participants were faster when making an open hand movement or a closed movement and if they were faster if the person on the screen was sitting close by or out of reach. We ran five studies that consistently showed no effect of distance. We did find that participants were faster in making a complementary (opposite) hand movement compared to performing a hand movement in mirror image only when an open hand movement was observed. For closed hand movements this effect flipped, participants were faster performing a hand movement in mirror image compared to using their opposite hand, indicating an imitation effect. This set of studies shows that complementary actions are performed automatically, similar to earlier studies looking at imitation behavior.

Contrary to the first chapter it seems that this automaticity is so strong that it is not affected by perceived distance. This might explain why you can have the feeling to be part of a sports game while watching television or when playing a computer game, without being physically there.

The third chapter concerns the role of perspective taking in imitation as explained in the first part of this summary. Imitation is often coupled to perspective taking because copying someone’s behavior is linked to being able to judge how it ‘feels’ to perform behavior from that person’s perspective. As argued before, imitation is primarily used in passive situations. However, when interacting with others and when performing complementary actions, taking your own perspective (first person perspective) is more important. For example, you have to know how to catch a ball from your own perspective, taking the perspective of the thrower will be less relevant. Nonetheless, a set of studies has shown that perspective taking is important specifically when behavior displayed by others is important or relevant to you. This makes sense, it would be tiring to constantly compute the perspective of those around you or to involuntary imitate everyone you see. To make this point, however, researchers often have used passive instead of active social situations. In the third chapter we discuss three studies each using the same task, in which we added an active rather than passive social scene. Participants observed pictures of a person on a computer screen in front of whom two objects were displayed (a book and a glass) on a table. Subsequently, participants had to indicate where the book was positioned with respect to the glass. From your own perspective this could be left but when taking the perspective of the person in the image this would be the opposite (right). By making small changes to the scene, we could transform a passive scene, where a person was looking straight forward without touching any of the objects, to an interactive scene where the person grasped one of the objects and held it straight forward (towards the screen). In three studies we looked at the time it took for participants to judge the correct location of the objects and how often they took the perspective of the person in the image depending on the type of scene. We hypothesized that contrary to the idea that interactive scenes would lead to increased perspective taking, they would actually inhibit perspective taking because these scenes typically require taking a first-person perspective. This means that if participants were asked to take the perspective of the person in the image in an interactive scene, this would lead to more errors and longer response times (it would take longer to choose the correct location) compared to more passive scenes when judging the object’s location. Two of the three studies supported this hypothesis, although it seemed that perspective taking in interactive settings was more difficult irrespective of the participant's task to take the perspective of the person in the picture or their own.

The final chapter takes a different approach to the first three in that it looks at what happens in the brain when observing behavior that can lead to imitation. By repeatedly observing behavior displayed by others just after or before performing behavior yourself, this facilitates automatic imitation or complementary behavior. On a neural level, this means that when simply observing specific behavior this activates the same motor regions that are involved when performing the behavior yourself. A number of experiments suggest that specific neurons, so called mirror neurons, facilitate this mechanism. A consequence of the repeated coupling of visual information and overt behavior is that over time this association becomes a predictive one. If you have sufficient information about what is going to happen in the near future, it is possible to simulate behavior before you actually observe it. For example, if you are at the zoo, you are more relaxed when walking past the tiger cage compared to being on a safari. Environmental information helps you with predicting upcoming events and this can be measured on a neural level in motor regions before any event has taken place. The fourth chapter describes a study in which we used EEG (electroencephalogram) to observe changes in motor regions in the brain when observing simple hand movements. Participants were shown short clips which were predictable to a varying degree. Depending on the color of the hand at the start of the clip participants could predict if they would see a precision movement (using a thumb and index finger) towards an object or a whole hand grasp. For some of the colors, the hand movement was highly predictable (e.g., 70 % of the trials they would see a precision movement) but for other colors the hand movement was not predictable (50 % precision, 50 % whole hand). We expected that participants would simulate the hand movement before any movement would take place but only when they were very sure about the outcome. Also, we expected that movements that could not be predicted would lead to stronger involved of motor regions but only during the observation of the movement. Using time-frequency analysis, where we looked at the extent to which clusters of neurons fire in synchrony, we compared clips that were highly predictable with those that were not predictable. The results showed that motor regions in the brain were more active when participants anticipated a highly predictable movement compared to anticipating a not predictable movement prior to the onset of the movement. While observing the hand movement, there was no difference in neural involvement between highly and not predictable movements. However, we were able to distinguish instances were highly predictable movements were followed by a predicted and an unpredicted movement. Also, observing an unpredicted movement was different from observing a not predictable movement, which suggests that participants were probably not guessing the specific movement before observing it.

The findings in this dissertation have shown a number of things: a) contextual information determines the visual attention for objects relevant for complementary behavior; b) complementary actions can be performed automatically, similar to imitative actions, and is independent of interpersonal distance; c) interactive settings inhibit perspective taking, and d) motor processes have predictive processes that helps people to anticipate other's actions. Even though the findings provide mixed evidence and are not fully coherent with regards to the main question, they go against a model in which observing behavior automatically triggers (covert or overt) behavior. Depending on contextual information, the degree to which behavior is displayed by others in a passive or active manner and the degree to which behavior is predictable, complementary behavior is often more fitting than imitative behavior. This has consequences for a number of theories that limit themselves to the coupling between the observation and automatic performance of similar behavior. Also, our findings have consequences for the role of mirror neurons which by some have been described as means of facilitating imitation and even empathy. Similar findings in the field have shown how mirror neurons respond to a broad array of behaviors, also when the behavior has no specific goal.

Researchers from different labs that have looked at the same processes have shown comparable findings to ours and have demonstrated how imitative and complementary actions are used interchangeably in interactive situations. This dissertation has (hopefully) contributed in part to this effort.

Dankwoord

Dit is mijn favoriete deel van het proefschrift en soms het enige gedeelte wat ik van mijn voorgangers heb gelezen (sorry, je mag het ook bij die van mij doen). Niet alleen om diegenen te bedanken die mijn promotie mogelijk hebben gemaakt maar ook om een zootje mensen eens flink in de schijnwerpers te zetten. Zonder jullie had ik die ellendige revisies, onmogelijke analyses en wekenlange testsessies niet gered.

Grappig genoeg vond ik psychologie in de bachelor jaren niet heel boeiend. De vakken waren interessant maar niet levendig en ik had niet het idee om er verder in te gaan. Toch 'solliciteerde' ik bij de research master psychology aan de UvA en kon gelijk na de bachelor beginnen. Deze master veranderde alles voor mij. De focus op onderzoek, de leuke groep, de goede docenten en de manier hoe studenten werden behandelt was top. Na een half jaar stage in Washington en een master project bij Kai was ik om. Al wist ik aan het einde van de master pas wat promoveren was, ik wist gelijk dat ik het wilde doen. Het duurde echter nog twee jaar waarin ik werkte als onderzoeksassistent en parttime bij een testuitgever voordat Kai me opbelde met het beslissende nieuws: Ik heb een promotieplek voor je. Ik vroeg of ik dan gelijk maar ontslag moest nemen bij mijn parttime baan. 'Ja, dat zou ik maar doen ja'.

Vier jaar later heb ik geen enkele spijt van deze beslissing en heb (tot nu toe) de beste tijd van mijn leven gehad. Op dag 1 in de Diamantbeurs zag ik een uitermate relaxte Evert-jan z'n proefschrift aftypen. Hij gaf nog wel en passant aan dat het een rollercoaster ride zou worden maar met die houding dacht ik dat het wel mee zou vallen. Ik kan me nog goed herinneren dat ik als een speer begon: nog even die analyses eruit knallen (ik had al veel data verzameld als assistent) en dat papertje kon er ook nog wel even uit. Gelukkig kwam ik er al snel achter dat het lang niet zo makkelijk ging en de output stagneerde en ging met vlagen heen en weer (als voorspeld). Toch heb ik in al die jaren geen dag thuis gewerkt (behalve bij ziekte) en had altijd zin om naar de UvA te gaan. Niks beter dan een plek met zoveel vrijheid, de mogelijkheid om dagenlang papers te lezen en dan weer twee weken lang analyses uit te zoeken. Hopelijk heeft het na vier jaar iets opgeleverd naast alle bloed, zweet en tranen die ik erin heb gestoken (conform de Frans de Waal standaard). Ik weet dat het in ieder geval niet was gelukt zonder de volgende mensen.

Om chronologisch te beginnen wil ik vooral mijn ResMas groep bedanken. Misschien was het wel niet eens de inhoud maar vooral de gezelligheid in deze twee jaar van de master. Het was al binnen een paar weken een grote vriendengroep, locals en internationale studenten door elkaar, in ons geclaimde computerhok op de derde verdieping van gebouw A. De echte toppers zie ik nog steeds zoals lieve Tamara, my Barcelona girl, Evin my favorite crazy girl, sweet Alba, Wei-yi my homie (leave your country and get here), mate Adam en Angelos hardworking funny man! Nog zoveel meer die vaak als AIO (en zelfs collega's) door zijn gegaan, Ozum, Effie, Kaya, Joram, Frieda, Hillie en Li-yang. Veel van jullie doen inmiddels een postdoc of zijn assistent professor, zo tof om te zien hoe goed iedereen bezig is. Tijdens mijn stage in Amerika heb ik ook ontzettend veel geleerd van Prof. Arie Kruglanski, buitengewoon intelligent, aardig en betrokken. Mooi ook om aan de hand van zijn eigen theorie te veranderen van asseser naar locomotor. Mijn andere vrienden en collega's Joce (m'n NYU Abu Dhabi Canadees), Jess, Vlad en Mike mijn Columbia heights buddy.

Na de master begon het allemaal in de Diamantbeurs, kamer 4.23. Een kamer met vele veranderingen in de eerste maanden. Evert-jan ging weg, Marleen was net weg en ik en Daniela kwamen er bij. Daniela, ik zie je als mijn academia zuster, we begonnen samen en zijn samen klaar, we zaten in de zelfde literatuur, waren even gefrustreerd en enthousiast over analyses, papers, professoren en alles rondom sociale cognitie (hoe minder sociaal hoe beter). Ik heb bijzonder veel plezier gehad om samen met je een kamer te delen en de nodige struggles van het academische gebeuren te trotseren. Tegenover me zat Liesbeth, positivo pur sang en buitengewone topmeid, gelukkig om nog wat locals op de afdeling te hebben tussen al die internationale figuren. Met gemak hebben we van 4.23 dé kamer gemaakt. En natuurlijk Xia. ik weet nog dat de ronde ging dat Gerben een chinese aio had die in onze kamer zou komen. Ik vreesde (conform sterotype) een schuchter, slecht engels sprekend meisje. Niks was minder waar, je bent echt de beste Xia (you are the best). Gezellig, hilarisch, en heel lief. Ben ook erg trots op onze signature minimal coffee cueing, die hebben we flink verfijnt binnen 3 jaar. Vind het echt te jammer dat je er niet bij bent! Hoop eigenlijk dat je nog blijft plakken als je klaar bent, China is ook weer niet zo boeiend.

Al was onze kamer de leukste van de afdeling veranderde de Diamantbeurs in gebouw G en veranderde de indeling. Lisanne en David kwamen samen met Xia, Daniela en ik in 2.34, een andere kamer met dezelfde nummers. Van bedompt systeemplafond naar een hoge design kamer met verstelbaar bureaus. De truc was om zo lang mogelijk erachter te staan totdat je niet meer kon wat begon met een half uur en eindigde in 4-5 uur staan achter elkaar. Nog nooit zo veel last van mijn rug gehad, bedankt verstelbaar bureau. Lisanne, je bent ook een toppert, lekker een beetje ouwehoeren de hele dag en voordat je het weet is je week om. Erg leuk waren de momenten buiten de UvA (al was dat te weinig), vooral je doordeweekse verjaardagsparty. Naast mijn kamergenoten was de hal gevuld met leuke mensen. Bastiaan, goeie vent en te leuk om met een paar biertjes een goed intellectueel gesprek te voeren, al was dit zeer sporadisch. Zelfde geld voor Allard, Coen, Giel, Marc en Jonas, gezellige groep jongens. Hannahtje, heb het erg gezellig met je gehad, beetje ouwehoeren en rondjes lopen rond de UvA, zet hem op! Astrid (nieuwe buur!), YongQi, Milena, Aafje, ook bedankt voor de leuke tijd. Mark je bent een goeie vent, vond het erg leuk om deels de SPO cursus te geven al was het maar om die handvol studenten die dat ook vonden. Ook anderen die ik niet alleen leuk vind maar wie ik als wetenschapper erg bewonder zoals Disa en Suzanne. Heren professoren, ook mijn dank: Bert-jan, Gerben (de speecher), Rob, Nils en Frenk en natuurlijk een flinke duim omhoog voor Annemiek, hoeksteen van de afdeling en gezelligheids team Sanne en Ran. Al is Michael nu uitbater van een trendy houtbewerkings shop in Tübingen, hij was ook mijn scriptie begeleider en allround top docent. Daarnaast de perfecte persoon om gedeelde frustraties mee te bespreken. Een grote pluim ook voor Agneta, buitengewoon gezellig, leuk maar bovenal relaxed, vergat vaak dat ik tegen mijn promotor aan het praten was en had dus misschien wat beleefder kunnen zijn. Ik ga er maar gemakshalve van uit dat mij begeleiden redelijk simpel is verlopen.

David, gouwe gozert, echt gezellig om met jou in een kamer te zitten. Samen maakten we wel even schoon schip in de academische wereld, die kon namelijk wel een likje verf gebruiken. Lekker dingen afzeiken of dingen eens flink appreciëren (!) en zonder een biertje door de week maar wel het driedubbele in het weekend. Al is het niet vaak gebeurd, als de gelegenheid zich voordeed liep de boel weer eens flink uit de hand. Met als gevolg dat het ochtend/ dagprogramma van dag 2 er vaak aan moest geloven. Natuurlijk ben ik wel een paar jaartjes ouder dus ik moest me wel een beetje gedragen maar ik kan me nog een aantal keren (vaag) herinneren dat daar weinig van terecht kwam. Logischerwijs ben jij dan ook mijn paranimf.

Ook buiten de UvA, via KLI of ASPO heb ik veel toffe mensen leren kennen. Product uit Portugal en ASPO blits was de pineapple groep, Reine, Anna, Lotte, Florien, Nic en Thijs, de nieuwe generatie (al kijk ik waarschijnlijk toe vanaf de zijlijn). Ook in Portugal: Rael, Lydia(!), are you still coming over here? Eervolle vermelding voor het dagelijks buffet in Portugal, daar kunnen ze bij een Nederlands congres nog wat van leren.

Naast mijn academische vrienden en collega’s wil ik al m’n vrienden daarbuiten bedanken voor al de gezelligheid en de nodige afleiding. De soldja boys, Niels, Coen, Daan, Alex, mijn homies en Monk-maatjes Maarten en Simon, de ManN8 crew, Hillie, Reintje, Ies, Pi, nieuwe collega yanges, Timmie. En daarbij natuurlijk Boston Strip in de eerste jaren, de maandelijkse radio show en m'n illustratie klusjes voor Marcel. Gelukkig kon ik het aardig combineren door de jaren heen (denk ik).

Michiel, ik wil je hier ook speciaal bedanken. Naast dat je een bijzonder leuke vent bent ben je een van mijn grote voorbeelden, ben altijd verbaasd hoe je zoveel solide output kunt genereren en continue met nieuwe interessante ideeën komt. Ook heb ik, wellicht is je dat niet opgevallen, je manier van redeneren/ discussiëren gekopieerd. Iets consequent beargumenteren en doordenken totdat alle flaws duidelijk zijn, heb hier veel aan gehad. Hoogtepunt was sowieso Budapest, qua inhoud iets te zweverig maar leuke mensen, beetje rondfietsen en zwemmen als het teveel werd en een leuk eindfeestje plus wel verdiende uitbrak termale baden. Was elk congres maar zo!

Kai, aan jou heb ik natuurlijk alles te danken! Zat al even na te rekenen dat we inclusief het masterproject, de twee jaar onderzoeksstage en de vier jaar promotie toch al 6.5 jaar samen onderzoek doen. Heb zoveel van jou geleerd, niet alleen met schrijven, designs uitdenken maar het rijlen en zijlen van de academische wereld, welke keuzes je wel of niet moet maken en wat een goede move zou zijn voor later. Plannen is mijn slechtste kant maar heb geleerd hoe je dingen serieuzer moet nemen en vooruit moet denken over grote beslissingen (bijvoorbeeld mijn laatste grote!). Ik ben je eeuwig dankbaar dat je me de kans hebt gegeven en ben dan ook jaren gedreven geweest om je te laten blijken dat je de goede persoon hebt gekozen. Ik hoop ten zeerste dat je je daar in kunt vinden. Daarnaast hoop ik dat we via werk/ universiteit nog samen kunnen werken.

Als laatste wil ik mijn familie bedanken, Mariette, Paul, Martin, Lot, Marcelo en ome Jan, al heb ik vaak tevergeefs proberen uit te leggen wat ik nou doe heb ik het heel gezellig gehad met jullie allemaal. Bijkomende handigheid dat de UvA, mijn huis en al jullie huizen binnen 5 min te bereiken zijn. Hoe makkelijk wil je het in een al kleine stad hebben, om het lastiger te maken ga ik nu dus zelf maar verhuizen. Ook mijn schoonfamilie plus aanhang: Ans, Paul, Jos, Anka, Eva, Mark, Geer, An en Sofie, voelt als m'n tweede thuis in het zuiden. Maar boven alles natuurlijk San, paranimf nr 2 (al was het wat met tegenzin). Na wat onhandige versierpogingen een jaar voor mijn promotie traject ging je hier wonder boven wonder op in. Nu vijf jaar later zijn we nog steeds samen en gaan we ook nog eens in één huis wonen. Je hebt mijn hele AIO tijd van dichtbij meegemaakt met veel liefde en een snufje ongeduld. Al kon ik niet urenlang over analyse issues met je praten (gelukkig maar) heb je vaak genoeg rake dingen gezet die ik nodig had. ‘Wie zit daar op te wachten?’ of ‘Ja, en dus?’ zijn dingen die ik mezelf vaak vergat te vragen en ik kwam er langzaam achter waarom ik hier wel over na moest denken. Maar bovenal maak je me al jarenlang super gelukkig, daar ben ik me elke dag heel erg bewust van. Het kan me niet schelen hoe ik die de komende jaren ga vullen als ik maar bij jou ben.

Ten slotte wil ik mijn commissieleden bijzonder danken dat zij bereid zijn geweest deel uit te maken van mijn promotie commissie. Ik kijk erg op naar jullie allemaal en ben dan ook deels angstig maar voornamelijk benieuwd naar jullie oordeel over mijn proefschrift. Al is mijn proefschrift vrij divers (lees: rommelig), dit heeft het voordeel dat er veel verschillende connecties gemaakt kunnen worden tussen jullie werk en dat van mij. Het is voor mij een grote eer dat jullie de moeite en tijd hiervoor hebben genomen.

The “Kurt Lewin Institute Dissertation Series” started in 1997. Since 2014 the following dissertations have been published in this series:

2014-01: Marijn Stok: *Eating by the Norm: The Influence of Social Norms on Young People's Eating Behavior*

2014-02: Michèlle Bal: *Making Sense of Injustice: Benign and Derogatory Reactions to Innocent Victims*

2014-03: Nicoletta Dimitrova: *Rethinking errors: How error-handling strategy affects our thoughts and others' thoughts about us*

2014-04: Namkje Koudenburg: *Conversational Flow: The Emergence and Regulation of Solidarity through social interaction*

2014-05: Thomas Sitser: *Predicting sales performance: Strengthening the personality – job performance linkage*

2014-06: Goda Perlaviciute: *Goal-driven evaluations of sustainable products*

2014-07: Said Shafa: In the eyes of others: *The role of honor concerns in explaining and preventing insult-elicited aggression*

2014-08: Felice van Nunspeet: *Neural correlates of the motivation to be moral*

2014-09: Anne Fetsje Sluis: *Towards a virtuous society: Virtues as potential instruments to enhance*

2014-10: Gerdien de Vries: *Pitfalls in the Communication about CO2 Capture and Storage*

2014-11: Thecla Brakel: *The effects of social comparison information on cancer survivors' quality of life: A field-experimental intervention approach*

2014-12: Hans Marien: *Understanding and Motivating Human Control: Outcome and Reward Information in Action*

2014-13: Daniel Alink: *Public Trust: Expectancies, Beliefs, and Behavior*

2014-14: Linda Daphne Muusses: *How Internet use may affect our relationships: Characteristics of Internet use and personal and relational wellbeing*

2014-15: Hillie Aaldering: *Parochial and universal cooperation in intergroup conflicts*

2014-16: Martijn Keizer: *Do norms matter? The role of normative considerations as predictors of pro-environmental behavior*

2015-01: Maartje Elshout: *Vengeance*

2015-02: Seval Gündemir: *The Minority Glass Ceiling Hypothesis: Exploring Reasons and Remedies for the Underrepresentation of Racial-ethnic Minorities in Leadership Positions*

2015-03: Dagmar Beudeker: *On regulatory focus and performance in organizational environments*

2015-04: Charlotte Koot: *Making up your mind about a complex technology: An investigation into factors that help or hinder the achievement of cognitive closure about CCS*

2015-05: Marco van Bommel: *The Reputable Bystander: The Role of Reputation in Activating or Deactivating Bystanders*

2015-06: Kira O. McCabe: *The Role of Personality in the Pursuit of Context-Specific Goals*

2015-07: Wiebren Jansen: *Social inclusion in diverse work settings*

2015-08: Xiaoqian Li: *As time goes by: Studies on the subjective perception of the speed by which time passes*

2015-09: Aukje Verhoeven: *Facilitating food-related planning. Applying metacognition, cue-monitoring, and implementation intentions*

2015-10: Jasper de Groot: *Chemosignaling Emotions: What a Smell can Tell*

2015-11: Hedy Greijdanus: *Intragroup Communication in Intergroup Conflict: Influences on Social Perception and Cognition*

2015-12: Bart de Vos: *Communicating Anger and Contempt in Intergroup Conflict: Exploring their Relational Functions*

2015-13: Gerdientje Danner: *Psychological Availability. How work experiences spill over into daily family interactions*

2015-14: Hannah Nohlen: *Solving ambivalence in context. The experience and resolution of attitudinal ambivalence*

2015-15: Stacey Sanders: *Unearthing the Moral Emotive Compass: Exploring the Paths to (Un)Ethical Leadership*

2015-16: Marc Heerdink: *Regulating deviance with emotions: Emotional expressions as signals of acceptance and rejection*

2015-17: Danny Taufik: *"Can you feel it" The role of feelings in explaining pro-environmental behavior*

2015-18: Sarah Elbert: *Auditory information and its parameters in health persuasion. The development of a tailored smartphone application to support behavior change*

2016-01: Anna van 't Veer: *Effortless morality — cognitive and affective processes in deception and its detection*

2016-02: Thijs Bouman: *Threat by association: How distant events can affect local intergroup relations*

2016-03: Tim Theeboom: *Workplace coaching: Processes and effects*

2016-04: Sabine Strofer: *Deceptive intent: Physiological reactions in different interpersonal contexts*

2016-05: Caspar van Lissa: *Exercising Empathy: The Role of Adolescents' Developing Empathy in Conflicts with Parents*

2016-06: Marlon Mooijman: *On the determinants and consequences of punishment goals: The role of power, distrust, and rule compliance*

2016-07: Niels van Doesum: *Social mindfulness*

2016-08: Leonie Venhoeven: *A look on the bright side of an environmentally-friendly life: Whether and why acting environmentally-friendly can contribute to well-being*

2016-09: Florien Cramwinckel: *The social dynamics of morality*

2016-10: Junhui Wu: *Understanding Human Cooperation: The Psychology of Gossip, Reputation, and Life History*

2016-11: Elise C. Seip: *Desire for vengeance. An emotion-based approach to revenge*

2016-12: Welmer E. Molenmaker: *The (un)willingness to reward cooperation and punish non-cooperation*

2016-13: Liesbeth Mann: *On Feeling Humiliated. The Experience of Humiliation in Interpersonal, Intragroup, and Intergroup Contexts*

2016-14: Angela M. Ruepert: *Working on the environment*

2016-15: Femke Hilverda: *Making sense of food risk information: The case of organic food.*

2016-16: Debora E. Purba: *Antecedents of turnover, organizational citizenship behavior, and workplace deviance: Empirical evidence from Indonesia.*

2016-17: Maja Kutlaca: *The Role of Values and Value-Identity Fit in Motivating Collective Action*

2016-18: Felicity Turner: *A New Psychological Perspective on Identity content, its Conceptualization, Measurement, and Application*

2016-19: Tim W. Faber: *When Imitation Falls Short: The Case of Complementary Actions.*