2015

One step ahead: Investigating the influence of prior knowledge on the perception of others' actions

Nicholson, Toby

http://hdl.handle.net/10026.1/3660

http://dx.doi.org/10.24382/1462
Plymouth University

One step ahead: Investigating the influence of prior knowledge on the perception of others' actions

Toby Nicholson

School of Psychology

Plymouth University

A thesis submitted to Plymouth University in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

March 2015

Abstract

One step ahead: Investigating the influence of prior knowledge on the perception of others' actions by

Toby Nicholson

Historically, a dominant view has been that we understand others by directly matching their actions to our own motor system, emphasising the importance of bottom-up processes during social perception. However, more recent theories suggest that instead we actively anticipate others actions based upon intentions inferred outside of the motor system, from social cues such as language, eye gaze and object information. Across 13 experiments, the established representational momentum paradigm, as well as a cross-modal visuotactile paradigm were employed to test the hypothesis that people's perceptual processes while observing the actions of others would be affected by such top-down cues about the actor's intentions.

We found, first, that people overestimate other people's actions in the direction of motion. Importantly, these overestimations were directly influenced by social cues. Saying or hearing a word congruent with a subsequently observed action resulted in the action being perceived as further along its trajectory. Second, we found that people anticipate the tactile outcomes of other people's actions with their own sensory tactile systems but that the mechanisms differed for bottom-up and top-down driven predictions. In a task in which people had to detect tactile stimulation while watching others, seeing impending hand-object contact increased the bias to perceive tactile stimulation, even when there was none, while impending contact that could not be seen but only inferred increased tactile sensitivity.

These findings are discussed in the context of recent theories of top-down predictive processing during social perception and from the perspective of multisensory integration.

Table of Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Acknowledgements

I would firstly like to thank my family for their never-ending love, support and belief in me. Without them I would not have been able to embark on my masters, which is what brought me to Plymouth and laid the foundations for this PhD. So I am eternally grateful to them for allowing me to embark on this journey and helping me achieve it.

I would especially like to thank Trudi for her love and support. They have kept me afloat in many times of trouble and reminded me how lucky I am. I also am very grateful to her for the level of tolerance and patience she has shown, not only when the thesis stole away so much of my time, energy and enthusiasm, but also when it drove me into despair and she was left to pick up the pieces. Her love and compassion have been invaluable.

I would also like to thank my friends for helping to take my mind off my PhD on many occasions, and for reminding me that there was a world outside of it.

I would also like to thank all the lovely fellow PhD people who I have met at Plymouth throughout the three years. They have made it a lovely place to work and created a warm and supportive environment which made the three years so worthwhile.

Above all I would like to thank Dr. Patric Bach for encouraging me to apply for the PhD, and for his unwavering support and encouragement in me, which have been invaluable throughout the three years and kept me going. Through times of stress, worry and many crises of confidence Patric has continually offered me support and encouragement, and helped me to believe in myself much more. In addition to this he has provided many stimulating discussions which I have greatly enjoyed and which have been a constant reminder of why I am here. His intelligence and passion for research has been an inspiration to me, and has been a defining influence in opting for a career in research. Lastly, his constantly smiling face, jovial nature and his laid back personality have all made supervision under him a pleasure, and reminded me to try and worry less. Thank you to him for giving me this opportunity and for showing so much faith in me.

Author's Declarations

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee.

This study was financed with the aid of a studentship from Plymouth University School of Psychology. Relevant scientific seminars and conferences were attended at which work was often presented.

Publications:

Nicholson, T., Roser, M. and Bach, P. (submitted) *The right tool for the job: Action goal understanding is primarily driven by object, not motor, information.*

Nicholson, T., Solbrig, L., Hudson, M., Tipper, S.P. & Bach, P. (submitted) *Seeing or knowing: dissociating bottom-up and top-down predictions of others' sensory action outcomes.*

Hudson, M., Nicholson, T. and Bach, P. (submitted) *I see what you say: Prior knowledge of other's goals is automatically integrated into the perception of their actions.*

Hudson, M., Nicholson, T. and Bach, P. (under review) *One step ahead: the perceived kinematics of others' actions are biased towards expected goals.*

Bach, P., Nicholson, T. and Hudson, M. (2014) *The affordance-matching hypothesis: how objects guide action understanding and prediction*. Front. Hum. Neurosci. 8:254. doi: 10.3389/fnhum.2014.00254

Oral and Poster presentations:

Nicholson, T., Solbrig, L., Hudson, M., Tipper, S.P. & Bach, P. (2014, September) *I feel what you are doing: differential effects of observed and predicted touches.* Talk presented at the British Association for Cognitive Neuroscience Conference and Annual meeting, York, UK.

Hudson, M., Nicholson, T. & Bach, P. (2014, September) *I expect you to do as I say! Prior intentional attributions bias the perceived kinematics of other's actions.* Poster presented at British Association for Cognitive Neuroscience Conference and Annual meeting, York, UK.

Nicholson, T., Bach, P. & Roser, M. (2014, September) *The understanding of everyday object direct actions relies primarily on object not movement information.* Poster presented at British Association for Cognitive Neuroscience Conference and Annual meeting, York, UK.

Nicholson, T., Solbrig, L., Hudson, M., Tipper, S.P. & Bach, P. (2014, June) *I feel what you are doing: differential effects of observed and predicted touches.* Talk presented at 'Mindfield' the 6th School of Psychology Postgraduate and Staff Conference, Plymouth, UK.

Nicholson, T., Solbrig, L., Hudson, M., Tipper, S.P. & Bach, P. (2014, May) *I feel what you are doing: differential effects of observed and predicted touches.* Poster presented at Workshop on Concepts, Actions and Objects (CAOS), Rovereto, Italy.

Hudson, M., Nicholson, T. & Bach, P. (2014, May) *I expect you to do as I say! Prior intentional attributions bias the perceived kinematics of other's actions.* Poster presented at Workshop on Concepts, Actions and Objects (CAOS), Rovereto, Italy.

Nicholson, T., Bach, P. & Roser, M. (2013, June) *Action goal understanding is primarily driven by object, not motor, information.* Poster presented at the 3rd European Workshop on Social Neuroscience, Gent, Belgium.

Nicholson, T., Bach, P. & Roser, M. (2013, May) *Action goal understanding is primarily driven by object, not motor, information.* Poster presented at Concepts, Actions & Objects, Centre for Mind and Brain Sciences, University of Trento, Rovereto, Italy.

Nicholson, T., Bach, P. & Roser, M. (2012, June) *Where is the goal? Identifying goal processing regions in the action observation network.* Talk presented at 'Beyond the Science of the mind' the 4th School of Psychology Postgraduate and Staff Conference, Plymouth, UK.

Word Count for the main body of the thesis: 44,330

Signed…………………………………………………………………….

Date………..................

Chapter 1 - Social perception and its mechanisms

The importance of social perception

We live in a rich social environment that shapes how we understand the world around us. How we perceive other people and their actions has a direct influence on our own subsequent behaviour, and in turn how others perceive and interact with us. This perceptual loop is vital for social interactions. Social perception facilitates anticipating others' actions, which in turn aids planning and coordinating one's own actions, enhancing the fluency of social interactions (Bekkering et al., 2009; Sebanz, Bekkering & Knoblich, 2006). For example, when being a passed a ticket from the bus driver we meet her hand at the end of the action rather than intercept it on its way. From studying someone's gaze, we can predict which item they will choose on the shelf, and, when dancing, we can fluently anticipate our partners' movements. In addition, social perception strongly influences learning, as evidenced by our tendency to imitate the actions of others (Brass, Bekkering, Wohlschläger, & Prinz, 2000; Meltzoff & Decety, 2003; Meltzoff & Moore, 1977). For instance, when learning a new skill we tend to observe and copy the actions of someone who already has the skill, such as how children learn to lace their shoes or button their coat for the first time by watching a parent. Likewise, when in a foreign country one learns the conventions of the culture by observing how the locals behave, such as validating ones train ticket prior to travel. This tendency to imitate happens in social situations even when not intended and occurring outside of direct awareness (Bargh, Chen & Burrows, 1996; Chartrand & Bargh, 1999).

Social perception is also important when merely passively observing other people. Our ability to read other people's intentions appears to come to us naturally (Allison, Puce & McCarthy, 2000; Tomasello, Carpenter, Call, Behne, & Moll, 2005). For instance, when noticing a friend sweating we are not surprised if he subsequently removes his jumper. Moreover, this tendency to apply intentionality to other people also applies to seemingly non-social stimuli, such as Heider and Simmel's abstract figures (Heider & Simmel, 1944; Hubbard, 2004). Such effects can also work against us though. For example, when asked in isolation to match the length of a line with three other lines, people show high accuracy (Asch, 1956; Bond & Smith, 1996). However, when the same judgement is required in a group situation with seven other confederates who give an incorrect answer on purpose prior to the participant's judgement, people are far more likely to copy the incorrect answer. This demonstrates that the behaviour of others can strongly influence one's own decision making. Similarly, people show implicit biases towards one's own racial group, emphasising the potential dangers of automatized social perception (Greenwald, McGhee & Schwartz, 1998; Smith-McLallen, Johnson, Dovidio, & Pearson, 2004). These studies show that we are both consciously and unconsciously affected by social perception and this can change our behaviour in both useful ways, but also in ways that could be harmful.

Social perception is clearly then of central importance for understanding how we integrate our own thoughts, feelings and actions with those of others, and therefore has profound implications for human and societal development. One difficulty at the core of understanding social perception is how we make sense of other people's actions. With no direct access to their internal mental states, we are left only with inferences generated by what we see others say and do. Nevertheless, humans typically take an "intentional stance" (Dennett, 1996) and interpret others' behaviour in terms of their goals and desires. However, sometimes these inferences are wrong, and it is not always clear how we adjust our understanding of the situation fast enough to respond efficiently when required. These problems are compounded by the fact that people are masters of deception, and social interactions are riddled with strategies that aim to control the perception we present to others. Despite this we still seem relatively efficient at anticipating others' behaviour, a skill likely to have been vital evolutionarily, and one that seems to mark us out from many other species (Penn & Povinelli, 2007; Saxe, 2006). What remains less clear are the mechanics of how we achieve this, making social perception of central importance in the quest for a more complete understanding of the human mind in its social environment.

A case of simulation? Humans see, humans do?

One dominant view of how people make sense of the actions of others has risen to prominence in recent decades, and has sparked a plethora of research buoyed by its implications for how social perception works. In this view social perception strongly involves one's own motor system (Rizzolatti & Craighero, 2004). According to this perspective, during action observation visual information is transformed in one's own motor system, and an internal simulation of the action is re-enacted. This is thought to provide a gateway to social perception, and understanding the mind of the observed through one's own motor experience (Rizzolatti & Sinigaglia, 2010).

The catalyst for these ideas was an accidental finding by a group of Italian researchers investigating the motor cortex of the Macaque monkey. Measuring single cell activity, these researchers stumbled upon the discovery that a certain group of neurons in the premotor cortex fired both when the monkey performed an action, but also when the monkey passively observed an experimenter performing the same action (di Pelligrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992). This unique group of neurons became known as 'mirror neurons', and was subsequently also found to exist in the parietal lobe of the macaque (Gallese, Fogassi, Fadiga, & Rizzolatti, 2002). The authors interpreted the dual activation during observation and execution as reflecting a motor representation of the action that they believed was crucial for action understanding, because it allowed the other person's action to be understood through one's own motor experience (Rizzolatti et al., 1996).

Subsequent research using indirect imaging techniques found support for a similar 'mirror system' within human premotor and parietal regions, the assumed human homologs of the regions where mirror neurons were found, with overlapping activations during both action performance and action observation (Decety et al., 1997; Iacoboni et al., 1999; Rizzolatti et al., 1996). A dominant explanation for the function of such an overlap is the 'direct matching hypothesis', which argues observed actions are internally re-run in the motor system to provide the observer with an understanding of the action's goal (Rizzolatti, Fogassi & Gallese, 2001). This view, therefore, sees the motor system as key to understanding other people's actions. Social cognition is conceptualised predominantly as a bottom-up matching process of observed action to one's own motor representation. Such a viewpoint extended previous theories emphasising the importance of simulation for social perception (Gordon, 1992), by providing the first neurophysiological evidence.

Further support for such motor views of action observation comes from the finding that different body parts, such as the hand, mouth or foot, results in somatotopic activation of the premotor cortex during action observation that matches the somatopy when acting with the same body parts (Buccino et al., 2001). In addition, studies have shown that expert dancers show greater activation of mirror areas when observing dance moves from their own repertoire, compared to those they do not perform and one would assume have less comprehension of (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006). Likewise, studies have found an increase in activation within mirror regions when viewing possible, but not impossible, biological movements, again suggesting that these are more 'understandable' (Stevens, Fonlupt, Shiffrar, & Decety, 2000). All these studies imply that the motor system is integral to social perception and supports the direct matching hypothesis.

Behavioural studies have also revealed an involvement of the motor system during social perception. For example, action observation can facilitate the simultaneous performance of a similar action but interfere with the performance of a different action (Kilner, Paulignan & Blakemore, 2003). Likewise, the action of lifting a box can distort the simultaneous perception of the weight of a lifted box that is only observed (Hamilton, Wolpert & Frith, 2004). These studies demonstrate that both action execution and action observation can affect one another when performed simultaneously, providing support for the idea that perception and action share a common code (Hommel, Müsseler, Aschersleben & Prinz, 2001; Prinz, 1997). Other studies have demonstrated that the prediction of an observed action improves when the action more closely resembles one's own action (Knoblich & Flach, 2001; Knoblich, Seigerschmidt, Flach, & Prinz, 2002). It has also been shown that, compared to controls, individuals who have experienced de-afferentation (a lack of working nerve fibers to communicate sensory information to the brain) demonstrate difficulties in accurately perceiving whether an observed actor's expectation of the weight of a lifted box is correct (Bosbach, Cole, Prinz, & Knoblich, 2005). This provides support that action observation involves a simulation of the observed action involving one's own motor and sensory system. From this perspective, any simulation resulting from action observation would be easier to understand and predict when it mirrors one's own motor experience.

Studies showing the specific muscles involved in an observed action can become activated during passive viewing also demonstrate an involvement of the motor system during action observation (Candidi, Vicario, Abreu, & Aglioti, 2010; Gueugneau, McCabe, Villalta, Grafton, & Della-Maggiore, 2015; Urgesi, Candidi, Fabbro, Romani, & Aglioti, 2006). Moreover, disruption of the motor cortex using transcranial magnetic stimulation has been shown to delay predictive gaze shifts during action observation providing support for its importance in guiding action perception (Elsner, D'Ausilio, Gredebäck, Falck-Ytter, & Fadiga, 2013). These findings all point towards an overlap between perception and action strengthening the case that the motor system is important for social perception.

The growing body of research focused on the mirror system has led to a raft of different cognitive processes being associated with it. These include imitation (Iacoboni et al., 1999), empathy (Avenanti, Bueti, Galati, & Aglioti, 2005; Iacoboni, 2005), theory of mind (Gallese & Goldman, 1998) and language (Arbib, 2005; Rizzolatti & Arbib, 1998), to name but a few, all of which are key aspects of social perception. For example, a growing number of studies have demonstrated that observing touch (Bufalari, Aprile, Avenanti, Di Russo, & Aglioti, 2007; Schaefer, Xu, Flor, & Cohen, 2009) and pain (Lamm, Decety, & Singer, 2011; Morrison, Tipper, Fenton-Adams, & Bach, 2013; Voison, Marcoux, Canizales, Mercier, & Jackson, 2011) can engage brain regions associated with directly experiencing the phenomena oneself. Similar effects have also been found when observing emotional expressions, with brain activations overlapping with those involved in performing those expressions (Bastiaansen, Thioux, & Keysers, 2009). Its apparent importance for social perception has also led to the suggestion that abnormalities in the system may be responsible for disorders such as Autism (Iacoboni & Dapretto, 2006; Oberman et al., 2005) and Schizophrenia (Arbib & Mundhenk, 2005).

Overall, the evidence provides a strong case that action observation involves, at least in part, many of the same processes involved in action execution, making the case for a direct matching explanation of social perception compelling. However, what remains less clear is the exact role that one's own motor system plays in social perception. From the direct matching perspective, these studies demonstrate the integral part of the motor system in understanding the actions of others, but others have proposed alternative explanations within which a top-down prediction precedes and generates any motor activity (Csibra, 2007; Kilner, Friston, Frith, 2007). Moreover, several recent findings cannot be explained within a purely bottom-up direct matching account, further challenging the view that motor simulation is the key to social perception.

The limits of a bottom-up explanation of social perception

As previously mentioned, the initial interpretation saw mirror neurons as crucial for action understanding, and therefore central to social perception (Rizzolatti, Fogassi & Gallese, 2001). From this perspective an internal simulation of what one perceived when observing a conspecific was necessary in order to make sense of the goal of the action and therefore their intention.

This explanation is attractive primarily because it emphasises the bottom-up nature of social perception, where sensory information activates matching motor representations in the observer. As the main information we have about other people seems to be what we observe in a bottom-up manner, it is easy to be convinced that the simulation of other people's actions could be involved in our ability to decode them. However, there are many reasons why such an interpretation seems to extrapolate further than the evidence really allows (Caramazza, Anzellotti, Strnad, & Lingnau, 2014; Hickock, 2009).

The initial inspiration for the theory of direct matching was motivated by the early monkey studies, which suggested that mirror neurons predominantly responded only to actions towards an object, thought to emphasise their goal-driven nature (Rizzolatti & Craighero, 2004). However, subsequent research has found that many macaque mirror neurons also respond to actions where no object is present (Kraskov, Dancause, Quallo, Shepherd & Lemon, 2009) or are activated by intransitive actions using the mouth, for example protrusion of the tongue or lip and lip smacking where it is hard to derive any straightforward intention (Ferrari, Gallese, Rizzolatti, & Fogassi, 2003). Moreover, other studies have shown that mirror neurons also respond when the object to which an observed grasp is directed is hidden from view (Umilta et al., 2001), and a proportion of mirror neurons have been shown to also fire when an action is withheld (Maranesi, Livi, Fogassi, Rizzolatti, & Bonini, 2014). All these findings are hard to reconcile with a predominantly bottom-up, direct matching explanation of mirror neurons, making their exact role in cognition far from clear.

In humans, similar issues arise, as lesions to key areas of the mirror system do not lead to clear-cut impairments in action understanding, which would be expected from a direct matching explanation. For example, damage to the Inferior Frontal Gyrus (IFG), thought to be the human homolog of region F5 in the monkey, does not affect patients' comprehension of actions (Kalenine, Buxbaum, & Coslett, 2010). Likewise, several studies have shown that, at the single subject level, individuals with brain lesions often show dissociations between the ability to recognise objects and the actions associated with them, and the ability to use those objects (Negri et al., 2007; Pazzaglia, Smania, Corato, & Aglioti, 2008; Tessari, Canessa, Ukmar, & Rumiati, 2007). Such dissociations suggest that the motor system cannot be as crucial to social perception as the direct matching hypothesis argues.

This is further supported by the demonstration that individuals with a lack of motor experience can still make sense of the actions of others. For example, a patient with upper limb aplasia, who was born without upper limbs, demonstrated similar accuracy and speed as control participants when tested on the comprehension of manual actions (Vannuscorps, Andres, & Pillon, 2013). While it cannot be discounted that the patient may have developed alternative methods for comprehending manual actions based on visual information alone, it still provides further support that motor experience is not crucial for social perception, further weakening the case for the importance of the motor system in action understanding.

Finally, many of the studies cited as support for a human mirror system have come from research involving indirect measures of cellular activity, such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG) and transcranical magnetic stimulation (TMS), where any activity or disruption of brain regions is coarsely measured. Due to the differences in species and resolution of neuronal measurement, any results from these methods are not straightforwardly equatable to those from the single cell studies with monkeys. Moreover, many methodological techniques used between the studies differ, making any extrapolation from one body of research to the other somewhat speculative (Turella, Pierno, Tubaldi, & Castiello, 2009). To date, only one study has presented evidence for human mirror neurons at the cellular level (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010). Interestingly, this study found evidence for human mirror neurons in areas outside the classical premotor and parietal regions, suggesting that the mirror system may in fact be much wider than previously thought, and may not be solely motor in nature (see, Caramazza et al., 2014, for similar claims). However, as this study was based on a patient sample the results should be interpreted with caution.

When taken together, these findings provide strong counter evidence that the direct matching hypothesis would not be sufficient to explain the mechanics of social perception. They show that mirror neurons fire in a number of different situations where bottom-up input levels vary, that lesions in key regions dissociate action perception from action execution abilities, and that a lack of motor experience does not result in a reduction of action comprehension. Therefore, social perception must rely on more than merely the internal simulation of observed actions. Indeed, more recent models assume that the brain is a prediction device and that all incoming stimulation, both social and non-social, is processed in the light of prior expectations (Bar, 2009; Brown & Brune, 2012; Bubic, Von Cramon, & Schubotz, 2010; Clark, 2013; Friston, 2011; Howhy, 2013).

Is the brain a prediction machine?

Ever since the early days of psychology, researchers have theorised about the importance of inference during perception (Helmholz, 1925). But in the last few decades a number of theories have risen to prominence, which argue that predictions are more important for the brain and behaviour as a whole than previously thought. One of the most influential is the *free-energy theory* of the brain (Friston, 2010). This argues that the central goal of biological systems is to maintain order against entropy, and that the brain is no different. Therefore the chief goal of the brain is to minimise free energy, which is achieved through the implementation of top-down generative models. These models predict sensory inputs based on prior knowledge freeing up cognitive resources for processing new or unexpected stimuli. This process, which Friston calls *active inference*, has the goal of minimising the surprise encountered during perceptual events (Clark, 2013). Therefore, the aim of these predictions is to minimise the prediction error, allowing cognitive resources to be streamlined and attention to focus on the errors alone. In this way prediction errors are important for learning through their potential to influence future predictions (Wills, Lavric, Croft & Hodgson, 2007).

From this standpoint, top-down predictions are central to the interaction between perception and action, facilitating effective behaviour through the anticipation of sensory information. These ideas fit with current models of action control, which claim that motor behaviour includes an efference copy of the action in order to anticipate the sensory outcomes of actions and respond quickly to events in the world, particularly unexpected outcomes (Wolpert & Flanagan, 2001). Evidence to support this comes from studies identifying *sensory attenuation*, the finding that the sensory effects of self-produced actions are perceived as weaker than those produced by an unpredictable external influence (Bays, Wolpert & Flanagan, 2005; Blakemore, Frith & Wolpert, 1999; Blakemore, Wolpert & Frith, 1998; Shergill, Bays, Frith, & Wolpert, 2003). This is thought to allow the discrimination between self-produced and externally produced sensations. Moreover, evidence suggests that this phenomenon is due to predictive, rather than postdictive, mechanisms (Bays, Flanagan & Wolpert, 2006).

A similar view focuses on the role of memory in forming these top-down predictions. According to Bar (2009), learning produces memory scripts. These scripts produce the predictions we generate about what to expect in different environments. From this view memories can generate simulations even in the absence of any observation. Bar suggests that these predictions can even be derived by analogy. For example, when one comes across an object for the first time, for instance a new fruit, its colour, shape and texture will all be linked to previous fruits will similar features, drawing an analogy between the novel object and previously experienced objects. In effect, Friston's generative models are labelled memories here, but the effect is the same.

Both these theories share the idea that cognition is essentially predictive, with perception not merely reflecting the receipt of bottom-up sensations, but instead resulting from the interaction between top-down predictions, based on prior knowledge, and bottom-up sensory input. These prior experiences produce top-down predictions based on the current context, which result in some aspects of the incoming sensory data becoming more salient than others. Moreover, they also imply that these predictions inform perception suggesting that they may have the power not just to guide but also to directly bias perception towards expectations (Feldman & Friston, 2010; Panichello, Cheung, & Bar, 2012). While this has the benefit of guiding effective actions and speedier reactions to changes in sensory input, it also provides a model to interpret other people, as the same forward models used for one's own actions could also be applied to the actions of others to anticipate their future actions (Brown & Brune, 2012).

Is social perception predictive?

Phenomena such as sensory attenuation (Hughes, Desantis, & Waszak, 2013) support the idea that people internally anticipate the consequences of their own actions, based on prior experience, which allows the discrimination between self-generated and externally generated actions. But recent theories explaining social perception (Csibra, 2007; Kilner et al., 2007) have suggested that such top-down predictions also provide the ability to anticipate the consequences of other people's actions through the tight link between perception and action (Hommel et al., 2001). These theories suggest that any motor involvement during social perception does not represent a simulation that deciphers the goal of the action. Instead, they argue it represents an action prediction, based on an inferred intention, allowing anticipation of, and comparison with, incoming sensory information. These ideas produce some interesting predictions about how such influences may directly influence social perception.

One such alternative theory is the *action reconstruction account*, within which simulation of another person's action is like a predictive apparatus akin to a tool used for verification (Csibra, 2007). According to this view, rather than simulation producing an understanding of observed actions in a retrospective fashion based on the receipt and subsequent interpretation of bottom-up sensory data, instead the interpretation of an action comes first from outside the motor system. Cues such as eye gaze (Macdonald & Tatler, 2015), objects (Bach, Nicholson & Hudson, 2014) and intention statements (Macdonald & Tatler, 2013) provide rich information about an action's goal that can be used to make such initial action interpretations. This shifts the emphasis of social perception from relying on the interpretation of bottom-up signals to the *prediction* of bottom-up signals based on inferences generated by prior experience. Such predictions can guide one's own actions, in pursuit of one's own goals during competition and coordination, but also aids collective goals, for example joint action (Sebanz & Knoblich, 2009), and can therefore aid the fluency of social interactions (Sebanz et al., 2006).

Csibra (2007) distinguishes his account from the direct-matching account by referring to the distinction between imitation and emulation. In the direct-matching account observed actions are mirrored within the motor system to identify their intention. In contrast, in the action reconstruction account this sequence is reversed. A possible intention is inferred outside the motor system, which produces a top-down emulation of the actions required to achieve the predicted intention. Here then the emulation differs from a direct imitation by being produced not by the visual input of sensory data but instead by top-down prior knowledge. This action hypothesis can then be used to check whether the predicted action is indeed occurring, to fill in and compensate for ambiguous sensory stimulation, and to predict ahead what is going to happen next. From this perspective, social perception is predictive in nature and the mirror system is involved in gaining an understanding of the action, but its role is changed from producer of understanding to a *checker* of understanding.

Similar to this view is the *predictive coding account*, which is central to the *free-energy theory* of the brain (Friston, 2010). While it shares with the action reconstruction theory an emphasis on top-down predictions it is rather more specific in regards to its mechanisms (Kilner et al., 2007). According to predictive coding, the brain is organised hierarchically with multiple interacting signals communicating in a bidirectional manner from higher cortical levels down to lower ones and vice versa (Friston, 2010). The chief processing goal of such an organisation is reducing *prediction error* at each level of the hierarchy. The prediction error is the difference between a prediction and the subsequent observation of the thing predicted. When a prediction is correct, the prediction error will be minimal. In contrast, when a prediction is incorrect, the error will be high. Importantly, prediction errors themselves have perceptual consequences, as they render the source of the discrepancy from the initial prediction more salient. In this way, top-down predictions aid perception but also have the potential to bias perception when a prediction is incorrectly accepted or rejected. Top-down predictions are therefore not merely a passive tool to aid comprehension. Instead, according to this view they actively guide and affect perception. Therefore, while sharing the emphasis on top-down predictions with the action reconstruction view, the predictive coding account goes beyond this to make very specific predictions about how such a system should directly affect perception.

According to the predictive coding account such predictions are strengthened through experience, due to the mechanism's aim of reducing the size of prediction errors. This means that as prediction errors change in magnitude over time through learning, stimuli are perceived differently as a result (Cheung & Bar, 2012). Top-down predictions are beneficial to perception as they allow the disambiguation of stimuli, allowing sensory data to be processed more fluently (Bar, 2003, 2004; Fenske, Aminoff, Gronau, & Bar, 2006; Kok, Jehee, de Lange, 2012). However, such effects can also result in faulty predictions being accepted when sensory inputs either heavily correlate with a predicted outcome (Tsakiris & Haggard, 2005) or are limited due to context (Summerfield, Egner, Mangels, & Hirsch, 2006). This implies that top-down predictions are very useful during perception but can also bias people to certain interpretations, based on prior experience, which may not be veridical (Kok, Brouwer, van Gervan, & de Lange, 2013). For example, deficits in producing prediction errors can result in hallucinations (Horga, Schatz, Abi-Dargham, & Peterson, 2014).

The predictive coding theory suggests that the genesis for these predictions is prior knowledge. In the case of social perception, this refers to previous experiences of other people's behaviour, such as their emotional expression, verbal statements of intent or the direction of their gaze, and how such cues predict subsequent actions. In this way, these cues become associated with certain intentions, for example when somebody looks towards an object this is likely to elicit the prediction that they will move towards or pick up the object (Castiello, 2003; Pierno, Becchio, Tubaldi, Turella, & Castiello, 2008; Pierno et al., 2006). Therefore, the context of the situation, in the form of cues picked up either from other people (Teufel, Fletcher & Davis, 2010), or the objects they have access to (Bach et al., 2014), directly shape the top-down predictions a person will make in a given social situation. As in the action reconstruction account, the mirror system is seen as the core node that develops these predictions – what the other person is likely to do – from these assumed intentions (Kilner et al., 2007). However, here more than just checking the assumed intention, top-down predictions appear to exhibit a level of influence that implies a direct effect on perception not specified by Csibra.

The predictive coding account and the action reconstruction account are highly compatible, both suggesting that top-down predictive processing is central to social perception. However, the predictive coding account resides in a larger context of the free-energy theory (Friston, 2010), and subsequently goes beyond the action reconstruction theory in its reach. Moreover it posits that such a predictive mechanism should directly affect perception, with predictions enhancing or distorting what is perceived depending upon their accuracy. This can disambiguate stimuli, facilitating a comprehension of other people's actions that were not clear from the bottom-up input alone. Moreover, as prediction is seen as a central feature of perception, top-down predictions should happen automatically, whether people want them to or not. This suggests that our perception of other people's actions is not as veridical as previously thought, and is instead strongly shaped by what we expect to perceive due to prior experience.

Predictive perception in a social context

Top-down theories suggest that social perception itself is predictive and that prior knowledge directly shapes how stimuli are perceived. Therefore, rather than internally simulating an observed action to understand its intention, contextual cues in the form of gaze direction, emotional expressions and verbal statements have the potential to generate top-down predictions based on our prior experiences of the intentions associated with these cues (Teufel et al., 2010). This facilitates the anticipation of potential future actions of others and informs our own action planning in a complex and dynamic environment. A number of different lines of research provide evidence that we indeed predict the actions of others, often based on the intentional cues present, and suggests that this tendency begins early in life.

From an early age, children engage in anticipatory eye movements during action observation that takes into account both action and object information (Falck-Ytter, Gredeback & von Hoften, 2006). Moreover, when children observe others interacting with objects, anticipatory gaze shifts reflect a prediction of the goal of the action. For example, observing a person with a phone produces gaze shifts to the ear (Henrichs, Elsner, Elsner, & Gredebäck, 2012; Hunnius & Bekkering, 2010). It has also been shown that a child's ability to predict the goal of an observed action based on the type of grip and its match to available objects correlates with their own ability to perform such a grip (Ambrosini et al., 2013). This suggests that the tendency to anticipate events is an early developmental process, which seems to rely on one's own action and object knowledge, in line with predictive coding (Kilner et al., 2007) and action reconstruction views (Csibra, 2007).

Studies in adults further support these conclusions. A seminal study showed that the pattern of eye movements when stacking a pile of blocks was very similar to when the same task was observed, but completed by another person (Flanagan & Johansson, 2003). Moreover, the coordination between gaze and hand was predictive rather than reactive. Gaze was anticipatory for one's own actions, but also when observing others' actions, revealing some continuity between both domains. Research also suggests that anticipatory gaze shifts during action observation are present when both movement and contextual cues are available, but not when only movement information is available, supporting a guidance by top-down information rather than bottom-up information alone (Eshuis, Coventry & Vulchanova, 2009). Observing another person looking at an object has also been shown to prime actions towards that object, as if one were predicting, with one's own motor system, the action the other person will do in line with Csibra's model (Castiello, 2003). Likewise, observing object-directed gaze activates similar brain regions as observing actions with the object, in line with the idea that gaze at an object predicts a (simulated) reach towards it (Pierno et al., 2006, 2008). In addition, during action observation hand pre-shaping (whether it was precision or power grip) has been shown to elicit more accurate and proactive eye saccades to the target object of the reach (a large or a small object, Ambrosini, Costantini & Sinigaglia, 2011). All these studies demonstrate the predictive nature of eye movements also during adult action observation, supporting the idea that social perception is anticipatory and that intentional cues are a key driver of such predictions.

Other research has shown that some of the classical mirror neuron regions are also involved in action prediction. For example, one study has shown that the dorsal premotor cortex is important for predicting how an observed action will continue behind an occluder and that disruption to the region impairs this process (Stadler et al., 2011; 2012). Similarly, prior knowledge of object presence behind an occluder during action observation leads to activation within the somatosensory cortex, even though the touch is not directly perceived, implying that the consequence of the reach was predicted (Turella, Tubaldi, Erb, Grodd, & Castiello, 2011). This finding is line with previous research showing overlapping activation for touch and the perception of touch by another (Bufalari et al., 2007; Schaefer et al, 2009), but, because it occurred when contact was not directly perceived, is also evidence that such an overlap resulted from top-down predictive processes rather than bottom-up processes alone. This suggests that motor activity elicited during social perception could in fact reflect predictive processing as opposed to a direct matching process.

Research has also started to show that the integration of kinematic and object information is central to action observation. For example, action predictions are dependent upon the context within which they happen and which objects are available, and minute kinematic differences determine their generation and accuracy (Manera, Becchio, Cavallo, Sartori, & Castiello, 2011; Stapel, Hunnius & Bekkering, 2012). Likewise, in a recent study by Jacquet and colleagues (2012) participants identified, in a condition of visual uncertainty, complete and incomplete object-directed actions. For each object, an optimal (low biomechanical cost) and sub-optimal (high biomechanical cost) movement was presented. In line with the idea that predictions are derived from the object's affordances and bias visual perception towards them, participants more easily identified the movements optimally suited to reach a given object. Moreover, other studies have shown that observing somebody next to an object activates the most effective grip for interaction as if the observer was in the other person's position (Costantini, Ambrosini, Scorolli, & Borghi, 2011; Cardellicchio, Sinigaglia, & Costantini, 2013). In addition, when observing reaches to objects, the spatial alignment of hand and object leads to more automatic imitation even when these aspects are task irrelevant, suggesting that these social predictions are automatic (Bach, Bayliss & Tipper, 2011). These studies emphasise the importance of intentional cues in generating expectations during action observation, which can facilitate accurate predictions.

A number of other studies have probed social perception by using point light displays of biological actions interrupted by periods of occlusion to test whether people predict the actions of others accurately. One study found that over three different time intervals (100, 400 & 700ms) participants consistently showed fewer errors when judging a static test pose that was temporally in line with the duration of occlusion, providing evidence that people simulate in real-time (Graf et al., 2007). Subsequent research has supported this and found that these simulations are present even when the length of the action prior to occlusion is as short as 20ms (Parkinson, Springer & Prinz, 2012; Springer, Brandstader & Prinz, 2013). However, a different study which included a wider variety of test poses found that action simulation tended to lag the real time of the action (Sparenberg, Springer & Prinz, 2012), while another found that participants tend to view actions as slightly further forward in time (Jarraya, Amorin & Brady, 2005). Taken together these studies provide evidence that people can accurately simulate the actions of others during action observation, but provide mixed conclusions on whether such a process is anticipatory, real time or slightly slower.

Taken together all these studies provide support for the existence of predictions during social perception. However, in many of the studies, prediction was the actual task required, making it difficult to assess the automaticity of such predictions. Similarly, while all these studies support the idea that people anticipate other people's actions during social perception, none of them directly test whether these predictions affect perception itself. Therefore, in order to investigate whether perception is predictive during action observation perception itself needs to be tested, rather than the ability to predict or the enhanced processing of predictable actions.

Summary & Overview

Social perception is of vital importance both for social interactions and for action observation. Previous views conceptualising social perception as strictly reliant on the bottom-up simulation of an observed action using one's own motor system have recently been undermined by evidence and interpretations to the contrary (Kalenine et al., 2010; Carramazza et al., 2014). This has shifted the focus towards top-down predictions during social perception (Kilner et al., 2007; Csibra, 2007). Findings from a wide range of studies, across a number of different aspects of social perception, are beginning to provide support for the claim that we anticipate the actions of others and such predictions are heavily reliant on social cues of intention. If this is true it would flip around the traditional conceptualisation of social perception relying predominantly on bottom-up cues (Rizzolatti & Sinigaglia, 2010), and instead shift the focus of research to include the role of top-down predictions based on social cues to intention.

The implication of these top-down theories is that perception itself is directly affected by predictions, suggesting that the impact of any predictions during social perception should be automatic. However, up until now the research supporting the role of predictions during social perception have not directly measured such perceptual effects but have instead inferred its impact based upon people's explicit ability to predict and their tendency to process predictable actions faster and more fluently.

The current thesis had the aim to test whether predictions have such an automatic effect on perceptual processes during action observation. Moreover, if such predictions are generated by prior knowledge of the intentional nature of certain cues (eye gaze, statements of intent, objects in the scenes), the manipulation of these cues should in turn modulate the top-down predictions and produce different effects on perception. To achieve this, an established experimental effect from the non-social domain, representational momentum, was utilised in order to test whether our perception of other people's actions really is predictive and how intentional cues affect it (Chapter 2 to 4). In addition to this, in order to investigate the role of other sensory systems during social perception, a cross modal paradigm was also utilised (Chapter 5). It measured people's tactile perception during action observation in order to test how predictions of contact affect tactile perception on the observer's own body, and whether these predictions are derived from bottom-up or top-down information.

Chapter 2 - Methodology

As laid out in Chapter 1, the aim of the current thesis is to investigate the claims of top-down theories of social perception (Csibra, 2007; Kilner et al., 2007), which suggest that social cues that imply an intention should result in top-down predictions of what will be perceived, influencing the perception of others' actions. While intentions can be processed at multiple different levels of abstraction the current thesis focuses on low-level short term intentions relating to actions, rather than higher level intentions relating to attitudes and beliefs. In order to investigate this question, a paradigm was required which tested people's visual perception during action observation, and which would make such forward predictions measurable. Representational momentum was chosen due to its robust demonstration of perceptual modulation based on prior expectations (Freyd & Finke, 1984; Hubbard, 2005).

Representational Momentum as a manifestation of predictive coding

The term "representational momentum" was first coined by Jennifer Freyd in a 1983 paper as a possible explanation for differences in response times when testing people's memory of frozen-action photographs (Freyd, 1983). Participants were shown one frozen action photograph of a person jumping from a wall, followed by another very similar photograph from the same scene, but at a slightly later point in time. However, both photographs were close enough in location to not be easily distinguishable. Their order was counterbalanced, so that half the time the order mirrored the real-world temporal order of the images, while the other half the temporal order of the images was reversed, as if the movement was going backwards in time. Participants were asked to judge whether the second photograph was the same or different as the first. As the stimuli were never the same, the correct answer to all stimuli was that they were different. Freyd found that response times to make such a "different" response were significantly longer when the order of the images followed the real world temporal order. The author suggested that this delay could have resulted from a conflict generated by an internal representation of the implied movement, making it harder to report differences.

In a subsequent experiment this hypothesis was tested using a different experimental set-up (Freyd & Finke, 1984). Here, rather than photographs, participants were shown a sequence of rectangles presented at different angles to produce the appearance of an ongoing rotation. They were asked to judge whether a following static probe rectangle was in the same or different angle as the last seen image (see Figure 1). The probe could either have the same angle, or if it was different could be either rotated further forward or further backward. They found that both errors and reaction times increased when judging the forward probes, as if participants were continuing to internally represent the movement once it disappeared, making forward probes harder to distinguish from this forward rotated mental image. This confirmed the interpretation of the previous experiment and solidified the term, representational momentum, as a description of the effect. Since these initial studies, a raft of research has investigated the representational momentum effect confirming it to be a highly robust finding across a wide range of stimuli and methodological arrangements (see Hubbard, 2005; 2014, for reviews).

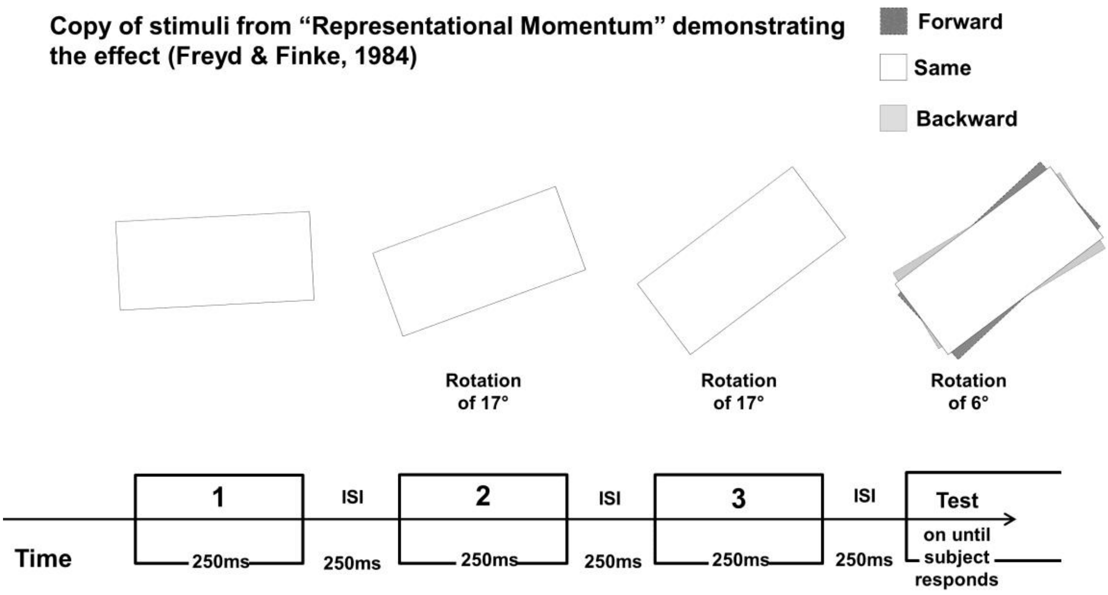


Figure 1. Adapted from Freyd & Finke's (1984) paper demonstrating the effect. Participants viewed a series of rectangles interspersed with inter-stimulus intervals (ISI). The fourth image could either be in the same location as the third image or in one of two possible different locations rotated either 6° forward (dark grey, large dashes) or backward (light grey, small dashes) from the position of the third image. (note: the stimuli depicted is not the exact stimuli used but a recreation for illustrative purposes)

Because participants are asked to accurately judge the object's final position, the representational momentum effect reflects an involuntary prediction generated by the perceived motion of the stimulus, which interferes with the participant's ability to accurately judge its final position. It is directly in line with theories of predictive coding (Friston & Kiebel, 2009). Expected stimuli (forward displacements) match the predictions and they are perceived as identical with the object's last seen position. Unexpected stimuli (backward displacements), however, elicit salient prediction errors and are readily detected. The effect therefore implies that people automatically simulate how a movement will continue and this changes how the probe stimuli are perceived: relative to predictions rather than objective reality.

Recent research is very much in line with such an interpretation. Historically, the representational momentum effect has been interpreted in a number of different ways (Hubbard, 2010). These range from low-level explanations relating to eye movements (Kerzel, 2000; Kerzel, Jordan, & Müsseler, 2001) to higher-level explanations related to mental representations (Freyd, 1987), beliefs (Hubbard, 2004), and a representational change in memory (Hubbard, 1995). However, more recent work suggests that it has a perceptual locus. While one study has implicated a fronto-parietal network in the effect linking it to working memory (Amorim et al., 2000), other research has provided evidence that area MT is integral to the effect, suggesting it may be more perceptual (Senior, Ward & David, 2002). Indeed, more recent research has revealed that the perceived result of a forward motion can be mathematically described by a superposition of observed stimuli and expected stimuli (Kimura & Takeda, 2015), and others have revealed that perceived motion already induces forward-directed perceptual changes while the motion is perceived, in line with a forward prediction of motion that happens not only in the gap between the disappearance of movement and presentation of the probe stimulus (Roach, McGraw & Johnson, 2011). Moreover, predictable movement at the leading edge of a stimulus suppresses BOLD responses, while an unpreceded movement at the trailing edge increases the BOLD response, as would be expected by predictive coding (Schellekens, van Wezel, Petridou, Ramsey, & Raemaekers, 2014). Finally, recent studies have revealed that unexpected probe stimuli in a visual sequence similar to representational momentum elicit visual mismatch negativities in the event related potentials (Kimura, Kondo, Ohira, & Schröger, 2011), a component directly related to the perceptual processing of unexpected stimuli, and interpreted as a prediction error response in the visual domain (Kimura, Schroeger & Czigler, 2011).

Moreover, in line with the assumptions of predictive coding, the representational momentum effect is directly affected by one's top-down knowledge about the stimulus, and the forces affecting it. For example, the representational momentum effect has been found to be susceptible to physical factors. For instance, movements in the direction of gravity show a greater representational momentum effect than movements against gravity (Hubbard & Bharucha, 1988; Hubbard, 1995, 1997). Other physical factors shown to modulate the effect include speed (Freyd & Finke, 1985), acceleration (Finke, Freyd & Shyi, 1986), friction (Hubbard, 1995, 1998; Kerzel, 2002), centripetal force (Hubbard, 1996; Kerzel, 2003), shape (Nagai & Yagi, 2001) and depth (Bertamini, 1993).

Similarly, representational momentum has also been shown to be modulated by ones' prior knowledge of the specific stimulus, providing further evidence for top-down influences on representational momentum. In one study researchers compared the effect of different objects on representational momentum, as well as the effect of the same object being presented with different conceptual labels (Reed & Vinson, 1996). In the first experiment, an identical target object was either labelled a rocket or a steeple, and representational momentum was measured when the object moved either up, down, left or right. The conceptual difference generated by the label here was that while a rocket is known to move in the real world a steeple is not. The results showed that across all directions the rocket elicited greater representational momentum, even though the object itself was identical and the only difference was what the participants were told what the object was. In later experiments they compared target objects of different shapes, and found that a 'rocket' target showed a greater upward representational momentum effect than a weight, box or church, while the weight showed the greatest downward representational momentum across all object types. These results showed that the conceptual processing of the object altered the level of forward extrapolation, providing evidence that more than the mere physical properties of the stimuli modulate the effect.

In a later study the same researchers investigated how far such effects relied on the concept of the objects alone, specified by its label, or by its appearance (Vinson & Reed, 2002). Again, stimuli depicting a rocket were used as a stimulus, but this time an atypical rocket was also used, which, although labelled as a rocket, did not match the prototypical appearance of one. This allowed a test of whether the conceptual effects emerged from the conceptual label alone or on an interaction of appearance and label. They found that the atypical rocket did not produce a similar increase in representational momentum as the regular rocket. This suggests that any conceptual effects on representational momentum are not produced by the label alone, but rely on the interaction between label and visual appearance. These results show that modulation of representational momentum in relation to conceptual information relies not only on prior object knowledge, but is directly integrated with the visual appearance, in line with the assumption that top-down predictions constantly interact and are verified by bottom-up information (Friston & Kiebel, 2009).

These studies demonstrate that while the representational momentum effect is automatic and represents a low level extrapolation of motion, it is affected not just by physical properties of the stimuli, but also by conceptual properties. They therefore demonstrate that the effect results, at least in part, from top-down predictions generated by prior knowledge.

Representational momentum during social perception

More recently, the representational momentum paradigm has started to be applied to less abstract stimuli, such as biological motion. However, even though these studies reveal the same representational momentum effect as for non-social motion, it is less clear to what extent these social predictions are influenced by top-down information about other individuals.

For example, evidence of a representational momentum effect has been found when participants watched short video clips of complex natural scenes, such as a high school, town square, department store and railway station, involving multiple different people, and therefore the processing of multiple actions at once in a holistic and naturalistic manner (Thornton & Hayes, 2004). Similar representational momentum effects have been found when participants are asked to judge the development in movement of an animated face from a neutral expression to the full extent of certain emotional expressions, showing the tendency to report a more extreme expression than was present (Yoshikawa & Sato, 2008). These findings highlight that the representational momentum effect is transferable to more naturalistic scenes involving real, as opposed to animated, biological motion and where a complex range of contextual cues are present. They therefore offer initial support for the hypothesis that social perception is indeed influenced by predictions, and that this tendency extends to more complex perceptual environments, but leave open whether top-down information affects these processes.

In another study, participants – sign language experts or novices – observed whole body stimuli of real actors performing sign language and revealed evidence for representational momentum when observing manual actions (Wilson, Lancaster & Emmorey, 2010). Interestingly, however, the study showed that experienced signers showed a decrease in representational momentum compared to non-experienced signers, rather than an increase. Based on the assumption that more experience with the actions should strengthen any top-down predictions of how the action would continue, and therefore increase representational momentum, this result seems to contradict the idea that action predictions are derived from prior top-down knowledge. The study did, however, demonstrate that easy actions resulted in more representational momentum than awkward actions, supporting the influence of action information on the effect. However, the finding that experience does not increase the effect challenges the view that expectations based on prior knowledge can affect representational momentum and, therefore, perceptual prediction.

A key question for the present purposes is whether representational momentum of others' actions reflects the intentions attributed to them. Recall that top-down models of social perception argue that any intention one attributes to another individual should have a top-down effect on the perception of their subsequent behaviour, biasing it towards these goals. Two studies attempted to address this issue. In one study an animated head turning to face the participant was used to measure representational momentum in conditions where the gaze of the actor was either ahead of, in line with, or lagging behind the head rotation (Hudson, Liu & Jellema, 2009). It was assumed that leading gaze would suggest a goal in this direction. While representational momentum was consistently found for each condition, it was indeed enhanced when the eye gaze led the direction of motion. This is line with the idea that people use the eye gaze of others as a cue to anticipate their future actions.

In a subsequent study the same authors investigated the effect that two cues, eye gaze and emotional expression, occurring in tandem had on representational momentum (Hudson & Jellema, 2011). Again, the gaze of the actor could either be ahead of, in line with or lagging behind, the direction of the head's motion, but this time the actor's emotional expression was also varied. Emotional expressions were classed as either congruent (i.e. joy and anger) or incongruent (i.e. fear and disgust) with the head's rotation towards the viewer. The assumption was that congruent expressions would further increase representational momentum when eye gaze led the direction of motion, as they would provide a further suggestion for an intentional head rotation towards the viewer. However, the findings revealed the opposite: representational momentum was greater when leading eye gaze was accompanied by fear or disgust, and the cues were therefore incongruent with one another.

Together, the above studies provide evidence that social cues such as gaze and emotional expression can produce similar representational momentum effects as non-social stimuli. While they provide some initial evidence that such effects can be modulated by social cues, the effect of intention on this modulation was not straightforward. While leading eye gaze seemed to increase the representational momentum effect of a turning head in one study, aversive emotional expressions seemed to increase the representational momentum of a turning head in the gaze direction. If representational momentum reflects a perceptual prediction based on top-down processes, cues suggesting an action goal – such as leading gaze and positive emotional expressions – would be expected to produce greater perceptual prediction and therefore more representational momentum. While these results provide evidence for the modulation of the effect in response to changes in social information, they provide unclear evidence concerning the influence of prior intentions on the perception of others' actions.

Summary and conclusions

The above studies involving biological motion illustrate the extension of the representational momentum effect to social perception, and provide support for the idea that the effect is modulated, to some extent, by the content of the social cues. However, the results of these studies also provide contradictory findings when viewed from the perspective of top-down predictive explanations of social perception. According to these theories intentional cues should increase perceptual prediction. However, while some studies provided evidence that unexpected social stimuli decrease representational momentum, others find the opposite, suggesting that unexpected social stimuli increase representational momentum. These studies therefore provide general support for the influence of top-down information on perceptual prediction, but fail to provide clear-cut evidence about how prior knowledge of others' intentions affects predictive processing.

The aim of the current thesis was to investigate this influence of intentional cues on perceptual prediction during action observation. This allowed top-down theories of social perception to be tested, specifically whether cues that signal an intention will increase perceptual prediction compared to those viewed as non-intentional. Importantly, intentions can be conceptualised at a variety of different levels, from short-term low-level intentions, to high-level long term intentions. While there is likely to be a link between high-level beliefs and subsequent action prediction during action observation, the current thesis focuses on low-level intentions at the level of actions as opposed to high-level intentions relating to cognition.

The established representational momentum paradigm was used to address this question. While the paradigm has already successfully been applied to social stimuli to demonstrate top-down effects on social perception, the influence that social cues have on these effects has produced inconclusive results, making the validity of top-down theories of social perception difficult to confirm on the basis of this research.

One limitation of the studies was that the intentions of the actions observed may not have been as clear-cut as hoped. For example, while facial expressions expressing fear and disgust were classed as incongruent with head rotations towards the participant, such a classification is not a given. Both these expressions can equally be viewed as valid expressions directed to a conspecific. Similarly, the so called congruent emotional expressions, particularly anger, in some situations may produce an expectation the head will turn away from, rather than toward, a conspecific. Therefore, in order to better answer the question of whether intentional cues increase perceptual prediction, and therefore representational momentum, the stimuli need to present a more clear distinction between intentional and non-intentional actions. To achieve this the representational momentum paradigm was used to display object-directed actions to address the influence of intentional cues on social perception more directly.

Chapter 3 - Do cues to intention increase perceptual prediction during action observation?

The aim of the initial studies was to establish a working representational momentum paradigm and test the hypothesis that cues, which signal an action's intention, generate predictions during action observation. One key intentional cue accessible during action observation is the kinematics of an action and how they relate to the available goal objects. Therefore the congruency between kinematics, in the form of hand shape and object orientation, was used to investigate whether actions perceived as intentional, those where hand and object were congruent, resulted in larger representational momentum than actions not perceived as intentional, when hand and object were incongruent. This would provide evidence that top-down information about other people's intentions produces perceptual predictions of their forthcoming actions, therefore supporting a predictive coding explanation of social perception (Kilner et al., 2007).

The kinematics and grip type of an action depend upon an action's goal. For instance, when reaching to pick up a cup, there is a particular hand shape (a power grip) required to successfully grasp it, and this differs from the hand shape required to pick up a pen (a precision grip). Similarly, the kinematics of a reach and the shape of one's hand depend upon the orientation and position of the object to be grasped. For example, when one wants to grasp a book, the book's position (upright or on its side) determines the required hand orientation. The failure to apply the appropriate kinematics and hand shape will prevent the successful achievement of one's goal of grasping the object. This is evident in the case of patients with optic ataxia, where a lesion to the posterior parietal lobe can result in impairment in the ability to successfully reach and grip an object (Rosetti, Pisella & Vighetto, 2003).

The process which helps us to select appropriate grips appears to be highly automatic. Research has shown that in the general population, object recognition can engage grip information even when this is irrelevant to the task performed. For example, seeing an object in a graspable orientation speeds up reaction times with the hand most suitable to perform this grasp (Tucker & Ellis, 1998; Symes, Ellis & Tucker, 2007). Similarly, responses using a power or precision grip are faster when this grip matches the type of object observed (e.g. small and large objects, Ellis & Tucker, 2000). These studies demonstrate that the mere observation of an object can automatically prime the appropriate grip.

This direct link between grip and object makes it an ideal tool to manipulate the perceived intentionality of an action. Indeed, research has demonstrated that, from an early age, children link the actions of others to the target objects associated with those actions (Gergely, Nádasdy, Csibra, & Biro, 1995; Reid, Csibra, Belsky, & Johnson, 2007; Kochukhova & Gredback, 2010; Hunnius & Bekkering, 2010). For example, infants as young as 6-9 months have been shown to be sensitive to the link between grip types and object sizes, showing dis-habituation when observing grasps that are incongruent with an initial grip (Daum, Vuori, Prinz, & Aschersleben, 2009). Moreover, 20-month-old toddlers can use grip cues to anticipate which one of two objects an action is directed to (Paulus, Hunnius & Bekkering, 2011).

In adults, merely showing a particular grip type triggers anticipatory eye movements towards an object with the corresponding shape (Fischer, Prinz, & Lotz, 2008). Another study measured eye movements when observing a reach that could be directed to one of two possible objects, one matching and one mismatching (Ambrosini et al., 2011). They found that, compared to reaches with the aim of merely touching an object, reaches with the goal of grasping an object resulted in earlier eye movements towards the object. This suggests that the kinematic cues inherent in the grasp, but not the touch, allowed observers to infer the intention of the action. This in turn demonstrates that kinematic cues signal to observers the intention of an action well in advance of its completion, and allow its further course to be predicted.

Grip-object matches also affect explicit action and object judgements. For example, viewing different grip types facilitate the subsequent recognition of objects matching this grip (Helbig, Steinwender, Graf, & Kiefer, 2010), and viewing objects gripped congruently helps identifying the goal of the action (Yoon & Humphreys, 2005). Conversely, the application of an incorrect grip type to a target object interferes with the ability to judge the correctness of an action goal (van Elk, van Schie & Bekkering, 2008) and a spatial mismatch between grip type and object impairs observers' ability to judge the semantic appropriateness of actions (Bach, 2004). These studies show that viewing a grip type facilitates recognition of the associated goal object, and similarly that processing an object facilitates the recognition of a grip type suitable for acting upon it. This shows the integrative nature of object and grip information, and confirms that object-grip matches are processed to some extent automatically also during action observation.

Together, these studies demonstrate that kinematic cues and their link to objects are an important signal of intention during action observation. If such intentions provide top-down information about forthcoming actions, then observers should specifically create predictions for congruent actions but not incongruent actions. To test this, we applied the representational momentum paradigm to reaches to a target object, which could either be spatially congruent or incongruent to the hand grip. If an action's intention triggers a prediction in the observer, representational momentum should be larger when the observed reach is congruent with the object compared to when it is incongruent.

Experiment 1 – Hand-object matches

Experiment 1 utilized the match of grip type and the orientation of the target object to convey the action's intention. Observing a hand reach towards an object with a congruent – but not with an incongruent – grip suggests the intention to reach for the object. If congruent actions signal the intention of the action, according to a predictive coding account of perception, they should generate forward predictions of how the observed action will continue to a greater degree than incongruent actions (Kilner et al., 2007). To the extent that representational momentum is a measure of these prediction processes, it should therefore be increased for such congruent (intentional) actions.

Participants were shown reaches to objects, which were either positioned standing upright or on their side. The hand orientation could either match the orientation of the object, representing a normal reach to grasp, or it could suggest a mismatch with the object's orientation, suggesting a reach that was not obviously bound to grasp the object. As in the standard representational momentum paradigm, participants observed the hand at three different sequential points along its reach. Participants were then asked to judge whether a fourth static image of the hand, which followed a brief interruption, was in the same location as its last previous position or in a different location. This fourth probe image could be in one of three different locations, the same position as prior to the interruption, slightly further forward towards the object or slightly backward, away from the object.

The first hypothesis was that this experiment would, overall, replicate the classical representational momentum effect. When the probe hand is displaced further forward along the movement's trajectory, participants should be more likely to mistake it for the hand's last position and respond 'same', compared to when the probe hand was displaced further backward along its trajectory.

The second hypothesis was that this representational momentum effect would be larger when the hand orientation was congruent, rather than incongruent, with the orientation of the object. This was based on the assumption that a spatial match between hand and object would allow people to infer that the reach was indeed directed at the goal object, and the hand intended to pick it up, compared to mismatches, which would disrupt such goal inferences. This would mean that in cases where a reach appears designed to grasp an object (intentional action) this should increase the expectation that the hand is aimed for the object and perceptual prediction should be stronger (larger representational momentum). Conversely, when reaches are more ambiguous (non-intentional action) and it is not clear where the hand is headed, these predictions may be restricted (smaller representational momentum).

Method

Participants. 35 students (14 male, mean age = 22.5, SD = 6.1) at Plymouth University or members of the wider Plymouth community took part in exchange for participation points or payment (£8 p/h). All had normal or corrected-to-normal vision and were native English speakers. All provided written informed consent prior to participation and were debriefed once the experiment had finished. The study was approved by the University of Plymouth's ethics committee.

Stimuli and Apparatus. The stimuli were filmed using a Canon Legria HFM36 HD video camera and edited using Moviedek and Corel PaintShop Photo Pro x3. Stimuli consisted of natural reaches of a man's right hand towards one of seven target objects (plastic cup, book, paint tin, cleaning spray, wine bottle, water bottle and tea cup) placed in either a horizontal or vertical orientation (see Figure 2). Individual frames were extracted from the videos, and for each reach 12 images were chosen which spanned the trajectory, from the start of the reach to approximately two thirds of the way through the reach. These images were then separated into 4 groups of 3 images, each representing a slightly different stage of the action, such that each covered equidistant points in time. For each of these 4 groups the last of the 3 images in the set was used to create the probe image. This involved digitally shifting the hand forward ('forward' probe) or backward ('backward' probe) along the trajectory by 20 pixels. This led to the addition of 2 images to each set. This resulted in 4 sets of 5 images for each reach. For each of these sets an additional set of images was digitally created which involved replacing the target object with the same object but positioned in the different orientation. This resulted in a total of 4 sets of images for each reach, 2 sets showing the original (congruent) reaches to a horizontally or vertically orientated object, and 2 sets showing the same reach but to the modified (incongruent) object orientation. Therefore half of the sequences showed reaches to a horizontally positioned object and half to a vertically positioned object, and within each of these, half showed a reach that matched the orientation of the target object, while in the other half the reach did not match the orientation of the object. For all images the object and hand were superimposed onto a black background to eradicate any other details from the scene. The complete stimulus set consisted of 28 (14 congruent, 14 incongruent) image sequences, each consisting of 20 images (4 groups of 5).

The stimuli were presented on a 20-inch monitor (60 hz) using Presentation (Neurobehavioural Systems, Inc.).

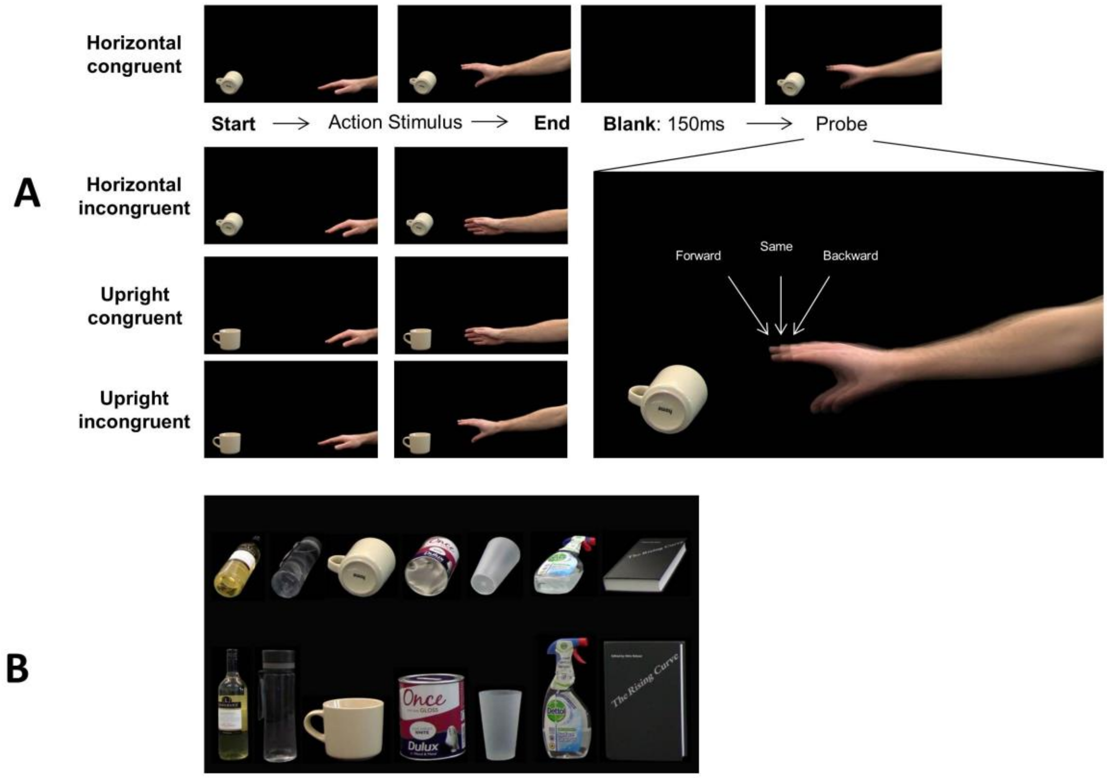


Figure 2. A: Trial sequence of reaches for each condition (horizontal congruent, horizontal incongruent, upright congruent, upright incongruent), followed by a blank screen, and then the probe stimulus. The larger image is a magnification of the probe image which shows the different probes ('forward', 'same' 'backward') overlapping one another for illustrative purposes. B: The objects in their horizontal orientation (upper row) and their upright orientation (lower row). Sizes have been modified for presentation purposes. In the experiment they were presented at their actual size.

Procedure. Participants were seated in a cubicle approximately 60 cm from a colour monitor, and were given written and verbal instructions. The experiment was split into two parts, part one included the representational momentum task while part two included a questionnaire explicitly asking participants how they thought each action would continue. In part one, each trial (see Figure 2) began with a fixation cross, presented for 500ms, followed by a blank screen also presented for 500ms. The first frame of the action sequence was then presented for 1000ms, followed by two subsequent frames, each presented for 150ms. A blank screen, also presented for 150ms, separated each of the 3 frames. Following the last frame a blank screen was presented for 350ms, before a probe stimulus was presented, which could either be exactly the same image as the last of the action sequence, or could show the arm slightly further towards the target object or slightly further away from the target object. Participants were asked to judge, as quickly and accurately as possible, whether the probe stimulus showed the arm in the 'same' or 'different' position compared to its last position prior to the probe. They made their judgment by pressing one of two possible response keys ('x', right, and 'z', left, counterbalanced between participants). If no response was made within 3 seconds, a prompt was displayed that they were 'too slow' along with a reminder of the response button assignment.

Part one of the experiment consisted of 12 training trials (not analysed), followed by 336 experimental trials, split by two break periods. Across part one, each stimulus combination was shown once (4 starting points, 2 hand orientations, 2 object orientations, 3 probe locations, and 7 objects), and their order was randomised. In part two, participants were shown all the actions again in the same way as part one, but instead of a probe image, participants were asked a question regarding how they thought the action would proceed. For each action, participants were instructed to choose between one of three possible answers: ‘the hand would pass the object or knock it over’ (press left arrow key), ‘the hand would grasp the object’ (press down arrow key) or ‘the hand would come to a stop before the object’ (press right arrow key). If no response was made within 5 seconds, a prompt was displayed that they were ‘too slow’ along with a reminder of the response button assignment. Across part two, each stimulus combination was shown once (4 starting points, 2 hand orientations, 2 object orientations and 7 objects) resulting in 112 trials. The experiment in total lasted approximately 30 minutes.

Statistical analyses. To ensure that participants were able to differentiate between probe types (same and different), the difference between the percentages of probes correctly identified as ‘same’ and the average of the percentages of probes incorrectly identified as ‘same’ (‘forward’ or ‘backward’) was computed for each participant. This produced a measure of participants’ *sensitivity* to differences between the probes. To calculate the level of representational momentum, the percentage of incorrect ‘same’ responses when the probe was actually ‘backward’ was subtracted from the percentage of incorrect ‘same’ responses when the probe was ‘forward’ for each condition of intention, congruent and incongruent.

Results

Exclusions. One participant's data were removed from the analysis because they failed to respond in more than 20% of trials. Inclusion or exclusion of this participant did not affect the results. Trials where reaction times were longer than 2000ms were also excluded from the analysis (<2% of all trials).

Sensitivity. Analysis of perceptual sensitivities confirmed that participants were able to distinguish 'same' from 'different' probes. They identified more probes as 'same' when they really were at the same locations as the last image, compared to the average of 'backward' or 'forward' probes, $t(33) = 3.03$, $p = .005$. Perceptual sensitivities in the congruent and incongruent conditions did not differ, $t(33) = .90$, $p = .376$.

Representational Momentum. The percentage of 'same' responses for the 'forward' and 'backward' probes was entered into a 2x2 repeated measures ANOVA with Probe direction ('forward' vs. 'backward') and Intention (congruent/incongruent) as factors. The analysis revealed a significant effect of Probe direction, $F(1,33) = 10.34$, $p < .005$, $\eta_p^2 = .239$, 95% CI [3, 13], with participants more likely to judge 'forward' probes as 'same' compared to 'backward' probes, confirming the classical representational momentum effect. Importantly, as predicted, there was also a significant interaction between Probe direction and Intention, $F(1,33) = 6.54$, $p < .02$, $\eta_p^2 = .165$, 95% CI [1, 7], demonstrating that this tendency to mistake more 'forward' than 'backward' probes as 'same' was larger for congruent grips than for incongruent grips (see

Figure 3). The average percentage of 'different' responses to same probes was 31%, and did not differ depending upon the intention of the action, $t(33) = 1.33$, $p = .192$.

Reaction times. The same 2x2 used for the representational momentum analysis was applied to the reaction times for correct responses, and revealed that neither main effect nor interaction were significant, all $F < 1.61$, all $p > .213$.

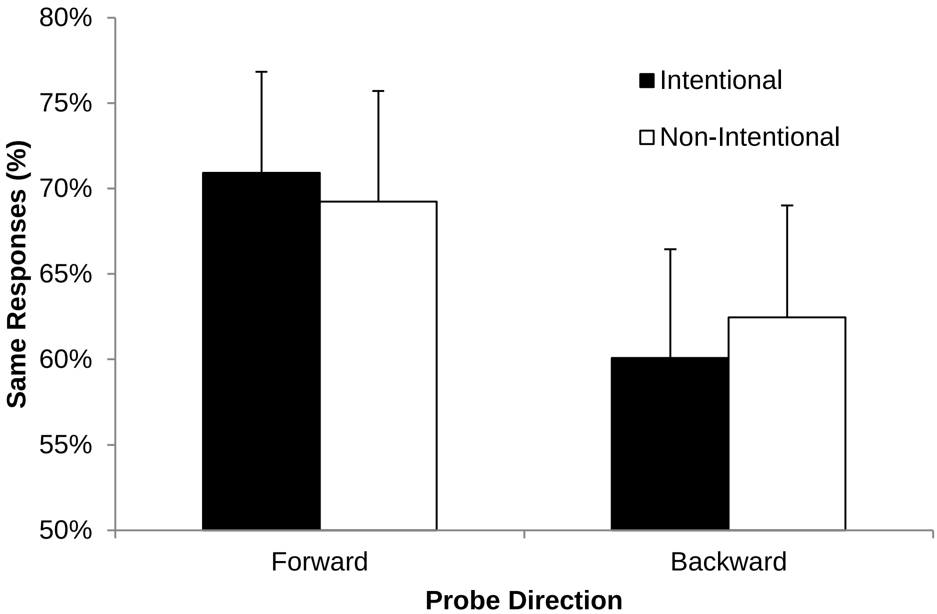


Figure 3. Mean percentages of 'same' responses to 'forward' and 'backward' probes, for both intentional (congruent grip, black bars) and non-intentional actions (incongruent grip, white bars) in Experiment 1. Error bars represent 95% confidence intervals.

Questionnaire. One participant did not complete part 2 of the experiment, resulting in 33 participants' data being analysed. Analysis of the questionnaire responses confirmed that participants did differentiate the intentionality of congruent and incongruent actions. Congruent actions were identified more often than incongruent actions as destined to grasp the object, $t(32) = 4.68$, $p < .001$ (see Figure 4).

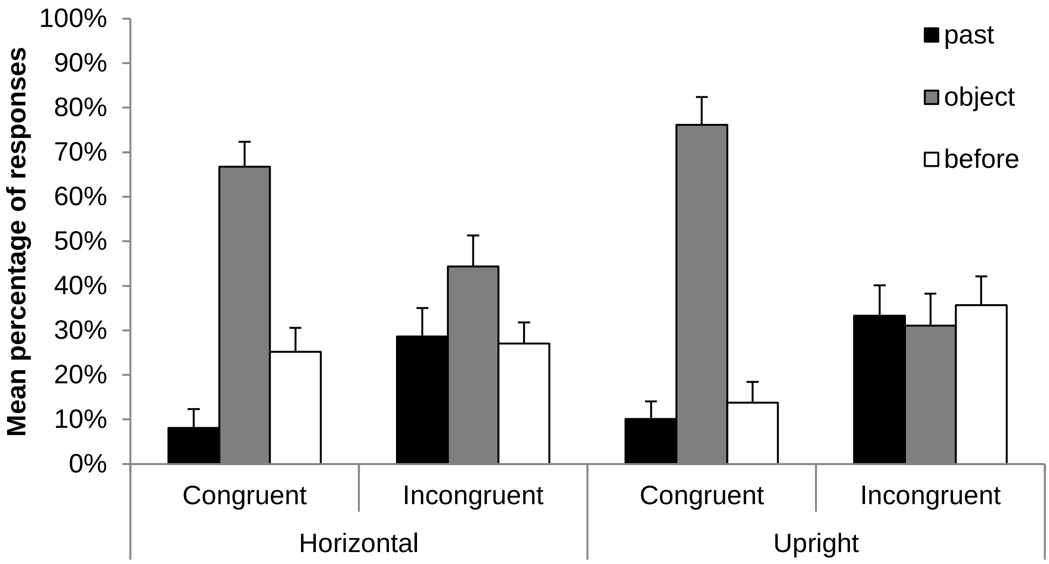


Figure 4. Mean percentages of responses to judgements of how actions were expected to end, by knocking over or going past the object (black bars), by grasping the object (grey bars) or by stopping before the object (white bars), for both congruent and incongruent actions when the object was either in a horizontal or upright orientation in Experiment 1. Error bars represent 95% confidence intervals.

Discussion

The results of Experiment 1 demonstrated that the classical representational momentum effect also applies to the observation of object-directed reaches. Displacements in the direction of motion were more easily mistaken for the hand's final position than displacements in the other direction. This supports prior research, which has provided evidence for representational momentum using biological stimuli (Hudson et al., 2009; Wilson et al., 2010; Hudson & Jellema, 2011), and extends it to intentional object-directed actions.

Crucially, the findings also confirmed that representational momentum was greater when viewing reaches with hand orientations that matched the orientation of the object. This provides the first evidence that representational momentum, when observing the actions of others, can be modulated by the action's perceived intention. As the intentional link between object and hand was task irrelevant, the fact that representational momentum was greater for intentional actions implies that the relationship between hand and object was processed automatically. The effect of intention can therefore be seen as resulting from an automatic perceptual prediction based on the congruency between hand grip and object. This is supported by the questionnaire, which demonstrated that congruent actions were viewed as more likely to grasp the object than incongruent actions. However, it is notable that over a third of incongruent actions were seen as likely to grasp the object, suggesting that the effect of intention may have been reduced by the perceived ambiguity of some of the incongruent stimuli.

The results of Experiment 1 provide initial evidence for top-down theories of social cognition. The spatial congruency between hand and object orientation modulated the representational momentum effect, with congruent actions increasing the tendency to misperceive 'forward' probes as further forward than they really were. However, as this was the first demonstration of an effect of intention on action perception, a replication of the study was run to ensure the conclusions of Experiment 1 were reliable.

Experiment 2 – Replication

In order to test the robustness of the findings of Experiment 1, a direct replication of the study was run. A power analysis was applied to existing data to estimate the number of participants required to confidently replicate the effects (Cohen, 1992), and it was calculated that 64 participants would be required to do so.

Method

64 students (9 male, mean age = 19.5, SD = 2.6) of Plymouth University took part in exchange for participation points or payment (£4). All other aspects were identical to Experiment 1.

Results

Exclusions. Five participants' data were removed from the analysis because these participants failed to respond in more than 20% of trials. Their inclusion or exclusion did not affect the results. Trials where reaction times were longer than 2000ms were also excluded from the analysis (5% of all trials). All analyses mirrored those of Experiment 1.

Sensitivity. Analysis of perceptual sensitivities confirmed once again that participants identified more probes as 'same' when they really were at the same locations as the last image, compared to the average of 'backward' or 'forward' probes, $t(58) = 4.99$, $p < .001$. Perceptual sensitivities in the intentional and non-intentional conditions did not differ, $t(58) = .074$, $p = .941$.

Representational Momentum. The analysis revealed a significant effect of Probe direction, $F(1,58) = 11.35$, $p < .005$, $\eta_p^2 = .164$, 95% CI [2, 8], replicating the classical representational momentum effect demonstrated in Experiment 1, with participants more likely to judge forward displacements as 'same' compared to backward displacements (see Figure 5). However, the interaction between Probe direction and Intention was not significant, $F(1,58) = 1.262$, $p = .266$, $\eta_p^2 = .021$, 95% CI [-1, 5], although numerically congruent actions did show a larger representational momentum effect. The average percentage of 'different' responses to same probes was 42%, and did not differ depending upon the Intention of the action, $t(58) = .604$, $p = .548$.

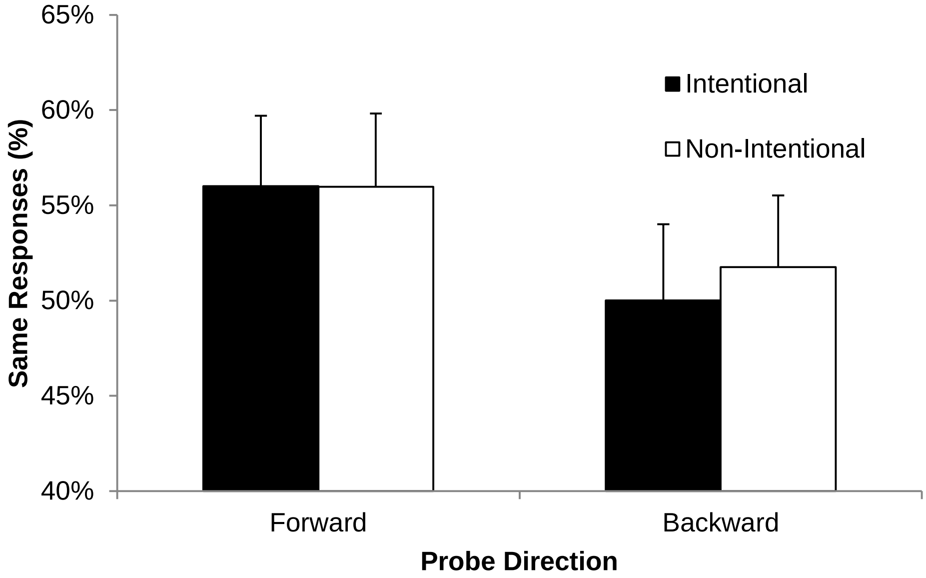


Figure 5. Mean percentages of 'same' responses to forward' and 'backward' probes, for both intentional (congruent grip, black bars) and non-intentional actions (incongruent grip, white bars) in Experiment 2. Error bars represent 95% confidence intervals.

Reaction times. Analysis of reaction times of correct responses revealed a significant effect of Probe direction, $F(1,58) = 12.61$, $p < .005$, $\eta_p^2 = .179$, 95% CI [17, 58], with 'backward' probes eliciting faster reaction times than 'forward' probes (see Figure 6). This supports the representational momentum finding in relation to incorrect response judgements, showing that when responding correctly, participants responded faster to backward displacements than to forward displacements. There was also a marginally significant effect of intention, $F(1,58) = 3.82$, $p = .056$, $\eta_p^2 = .062$, 95% CI [0, 28], with non-intentional actions responded to faster than intentional actions. The interaction between Probe direction and Intention was not significant, $F < 1$. The average reaction time when correctly responding to 'same' probes was 835ms and did not differ depending upon the intention of the action, $t(58) = .840$, $p = .404$.

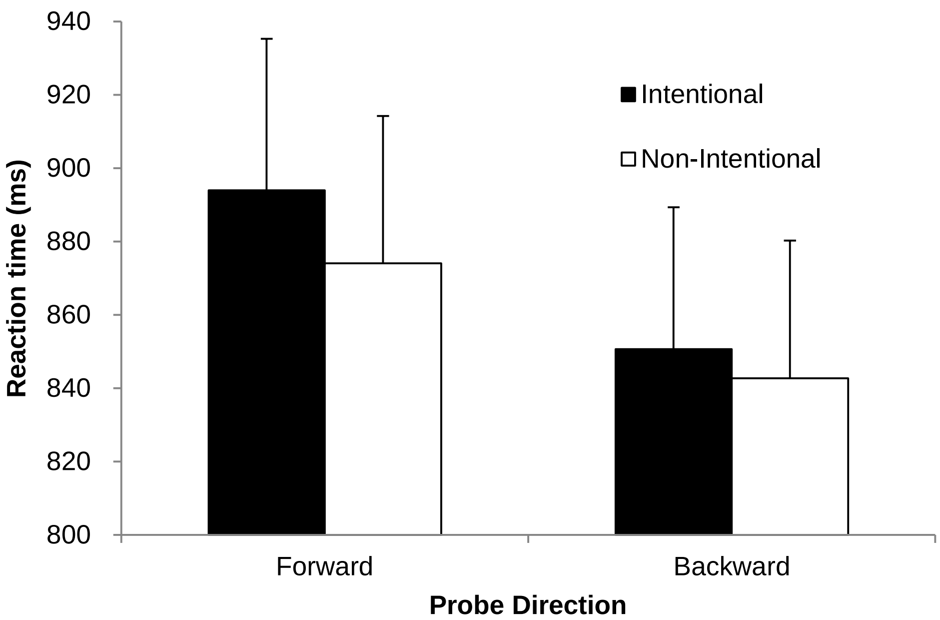


Figure 6. Mean reaction times to correct responses to 'forward' and 'backward' probes, for both intentional (congruent grip, black bars) and non-intentional actions (incongruent grip, white bars) in Experiment 2. Error bars represent 95% confidence intervals.

Questionnaire. Analysis of the questionnaire responses confirmed that participants did differentiate the intentionality of congruent and incongruent actions. Congruent actions were identified more often than incongruent actions as destined to grasp the object, $t(58) = 10.02$, $p < .001$ (see Figure 7).

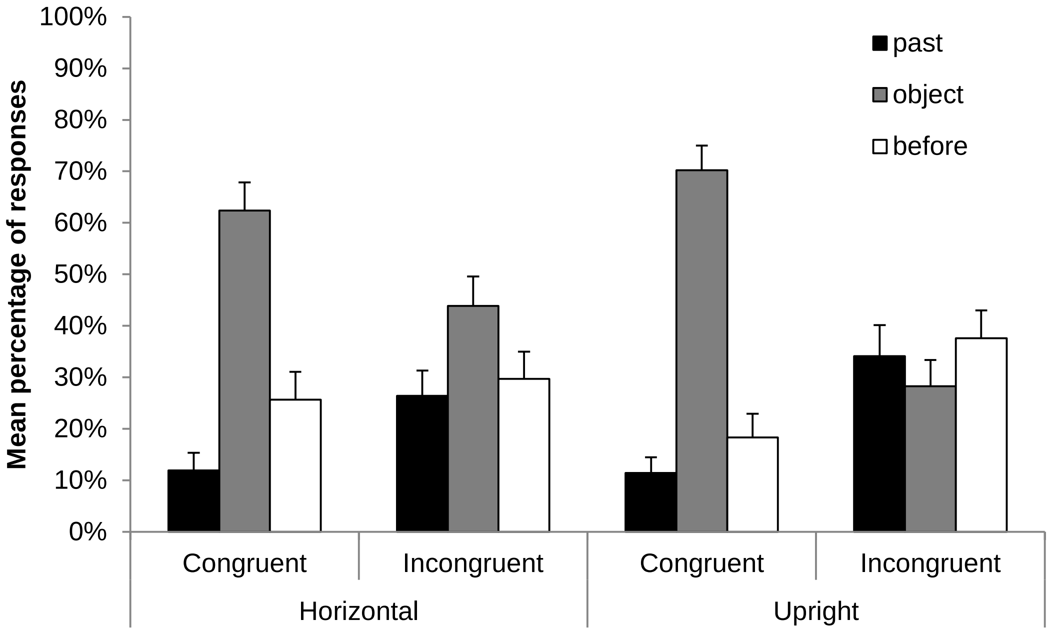


Figure 7. Mean percentages of responses to judgements of how actions were expected to end, by knocking over or going past the object (black bars), by grasping the object (grey bars) or by stopping before the object (white bars), for both congruent and incongruent actions when the object was either in a horizontal or upright orientation in Experiment 2. Error bars represent 95% confidence intervals.

Discussion

The results of Experiment 2 replicate the representational momentum effect shown in Experiment 1, with participants being more likely to mis-identify forward displaced probe hands with the hand's final position than backward displaced probe hands. However, Experiment 2 failed to replicate the effect of intention on action observation found in Experiment 1. While object-congruent actions again elicited more representational momentum numerically than incongruent actions, there was no statistical difference between them. Moreover, as in Experiment 1, over a third of incongruent actions were rated in the questionnaire as likely to grasp the object, again suggesting that some of the incongruent actions were, to some extent, seen as intentional as well.

Reaction times did differ when responding to 'forward' and 'backward' probes, with 'backward' probes eliciting faster correct responses than 'forward' probes, in line with prior research (Freyd & Finke, 1984). This supports the 'same' judgement response data in suggesting that, even when participants correctly distinguish a 'different' probe, this process is slower for 'forward' probes that conform to the direction of the reach.

Overall, the results are mixed. Experiment 2 again demonstrated a representational momentum effect. However, the effect of intention could not be replicated, calling the original finding into question. While a non-significant finding does not mean that no effect of intention exists, the larger sample size would be expected to strengthen, rather than weaken, any association between the observed reach and its perceived intention. While the questionnaire data show that intentional actions were identified as such, they also suggest that the non-intentional actions were more ambiguous than expected. Over a third of incongruent actions were classified as reaching the goal object, suggesting that the spatial congruency between hand and object may have been too subtle to produce a robust effect of intention.

Experiment 3 – Verbal cue

Due to the failure to find a robust effect of intention on representational momentum, a series of new studies were run on a modified paradigm that used language as an intentional cue. An important source of intentional information is the utterances of the observed individual. Language is humans' central form of communication, and is therefore a far more explicit indicator of their intention. In many cases, people state their intentions before acting, especially in social situations (Searle, 1969). Moreover, research has demonstrated that when an action is performed in response to an action sentence, people find it harder to respond correctly when the response required contradicts the action implied by the sentence (Glenberg & Kaschak, 2002). This is in line with the notion that language comprehension is situated within the context of action and elicits visuomotor imagery (Stanfield and Zwaan, 2001) and may rely, at least in part, on the same system involved in action selection and execution.

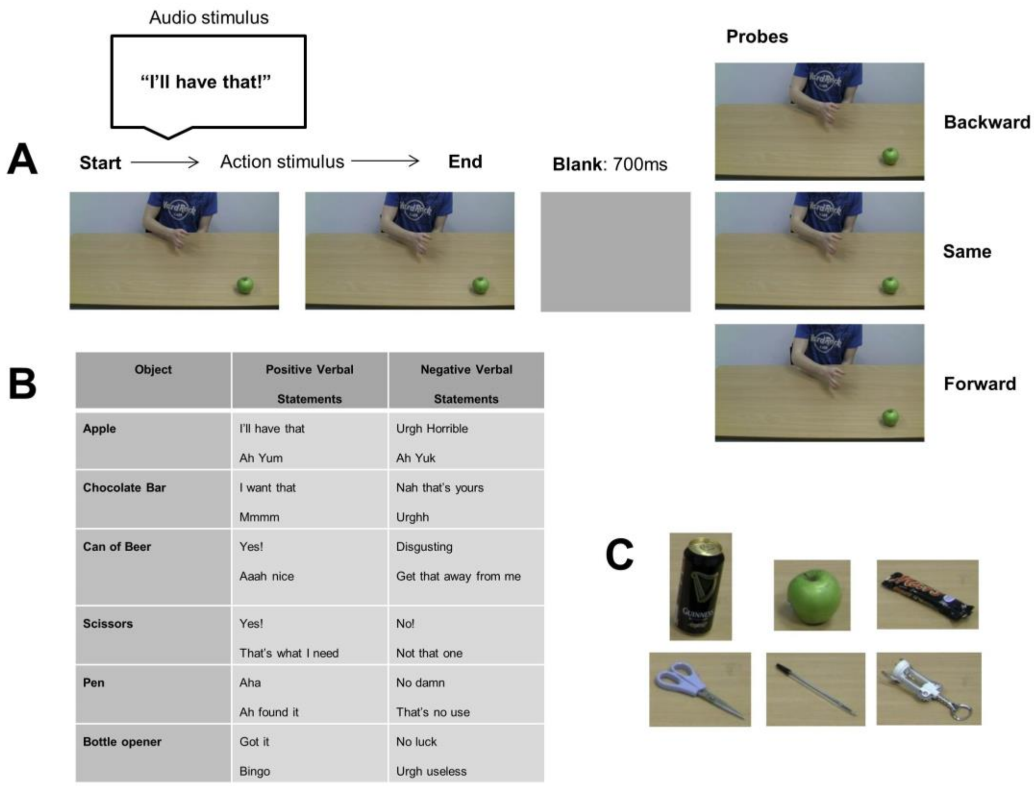
Experiment 3 therefore tested the impact of verbal intention statements on representational momentum. The stimuli depicted an actor reaching to a number of different objects on the table in front of them. At the onset of the action, participants were played (via headphones) a verbal statement that highlighted the intention of the actor. Some of the verbal statements declared a desire for the target object ('I'll have that'). Other, non-intentional verbal statements declared a lack of desire for the target object ('not that one'). To strengthen the intentional link between the verbal statements and the actions, the statements were spoken with an extra emphasis, highlighting the emotion being expressed. Therefore, intentional verbal statements were spoken in a positive tone while non-intentional verbal statements were spoken in a negative tone.

The assumption was that if intentional verbal cues facilitate top-down predictions, participants should show larger representational momentum after hearing intentional verbal statements suggesting a goal to reach for the object compared to non-intentional statements. Participants should therefore be more likely to miss forward displacements after having heard a positive intentional statement compared to hearing a negative non-intentional statement.

Method

Participants. 41 students (14 male, mean age = 21.2, SD = 4.1) at Plymouth University or members of the wider Plymouth community took part in exchange for participation points or payment (£8 p/h). All had normal or corrected-to-normal vision and were native English speakers. All provided written informed consent prior to participation and were debriefed once the experiment had finished. The study was approved by the University of Plymouth's ethics committee.

Stimuli and Apparatus. The stimuli were filmed using a Canon Legria HFM36 HD video camera. Stimuli consisted of an actor sitting opposite the camera and reaching with their right arm diagonally across the table to one of six objects (beer can, chocolate bar, apple, pen, bottle opener and scissors) located on the actor's left and the viewer's right (see Figure 8). All of the actor's upper body was visible to the camera with the exception of their head. For each reach 12 temporally equidistant steps through the reach were extracted. Each action sequence consisted of 4 images, and 4 different start points were selected from different stages of the reach so that some sequences showed the start of the reach while others showed the reach further towards the target object. In addition 24 verbal statements were created, 12 positive statements and 12 negative statements (see figure 7). Each object was matched with its own unique verbal statements, 2 positive and 2 negative statements which were varied across the experiment. The stimuli were presented on a 15-inch monitor using Presentation (Neurobehavioural Systems, Inc.).



Object	Positive Verbal Statements	Negative Verbal Statements
Apple	I'll have that Ah Yum	Urgh Horrible Ah Yuk
Chocolate Bar	I want that Mmmm	Nah that's yours Urghh
Can of Beer	Yes! Aaah nice	Disgusting Get that away from me
Scissors	Yes! That's what I need	No! Not that one
Pen	Aha Ah found it	No damn That's no use
Bottle opener	Got it Bingo	No luck Urgh useless

Figure 8. A: Trial sequence of reaches toward a target object. The audio stimulus began at the same time as the action stimulus and ended before it ended. A grey blank screen followed, before the probe stimulus was presented. An example of each is shown ('backward', 'same' and 'forward'). B: The table displays all the verbal statements used for each object. Each object had 2 unique positive and 2 unique negative statements. C: The six target objects used in the experiment (beer can, apple, chocolate bar, scissors, pen and bottle opener). Sizes have been modified for presentation purposes. In the experiment they were presented at their normal size.

Procedure. Participants were seated in a cubicle approximately 60 cm from a colour monitor, and were given written and verbal instructions. Each trial (see Figure 8) began with a fixation cross presented for 500ms, followed by a blank screen also presented for 500ms. The first frame of the action sequence was then presented for 200ms, followed by the next three frames of the action sequence, presented sequentially, each for 200ms. In contrast to the first experiments, the different steps were not separated by blank screens, to achieve a more natural presentation. At the same time as the first frame was presented, either a positive or negative verbal statement was played via headphones. Each verbal statement lasted longer than 200ms but less than 800ms, so that the verbal statements were presented at the same time as the action sequence but always finished prior the end of the action sequence. Participants were asked to listen to the verbal statement and not to ignore it, even though it was not specifically task-relevant. After the final action of the sequence, a neutral image of roughly equal luminance (a grey box matching the dimensions of the frame and presented in the same location) was presented for 700ms, before the probe image was presented. The probe stimulus could either be the same as the final frame of the action sequence ('same' probe), further along the observed trajectory ('forward' probe), or in a position just prior to the final frame ('backward' probe). Participants were asked to judge, as quickly and accurately as possible, whether the probe stimulus showed the arm in the 'same' or 'different' position as its last position prior to the probe. They made their judgment by pressing one of two possible response keys ('x', right, and 'z', left, counterbalanced between participants). If no response was made within 3 seconds, a prompt was displayed that they were 'too slow' along with a reminder of the response button assignment. The experiment consisted of 12 training trials (not analysed) and 288 experimental trials. Each stimulus combination was presented once (4 starting points, 2 positive verbal statements, 2 negative verbal statements, 3 probe locations and 6 objects), and their order was randomised. The experiment lasted approximately 30 minutes.

Results

Exclusions. One participant's data were removed from the analysis because the participant failed to respond in more than 20% of trials. Inclusion or exclusion of this participant did not affect the results. Trials where reaction times were longer than 2000ms were also excluded from the analysis (<2% of all trials).

Sensitivity. Analysis of perceptual sensitivities confirmed that participants were able to distinguish 'same' from 'different' probes. They identified more probes as 'same' when they really were at the same locations as the last image, compared to the average of 'backward' or 'forward' probes, $t(39) = 10.18$, $p < .001$. Perceptual sensitivities in the intentional (positive verbal statement) and non-intentional (negative verbal statement) conditions did not differ, $t(39) = .837$, $p = .408$.

Representational Momentum. The percentage of 'same' responses for the 'forward' and 'backward' probes was entered into a 2x2 repeated measures ANOVA with Probe direction ('forward' vs. 'backward') and Intention (positive verbal statement/negative verbal statement) as factors. The analysis revealed a significant effect of Probe direction, $F(1,39) = 22.96$, $p < .001$, $\eta_p^2 = .371$, 95% CI [7, 17], with participants more likely to judge 'forward' probes as 'same' compared to backward probes, confirming the classical representational momentum effect (see Figure 9). There were no other significant effects, all *F's*< 0.3.

Reaction times. The same 2x2 used for the representational momentum analysis was applied to the reaction times for correct responses. One participant was removed from the analysis as they failed to make a correct response in some of the conditions. The analysis revealed a significant effect of Probe direction, $F(1,38) = 9.36$, $p < .005$, $\eta_p^2 = .198$, 95% CI [22, 100], with correct responses to 'backward' probes significantly faster than correct responses to 'forward' probes. There were no other significant effects, all *F's*< 2.94.

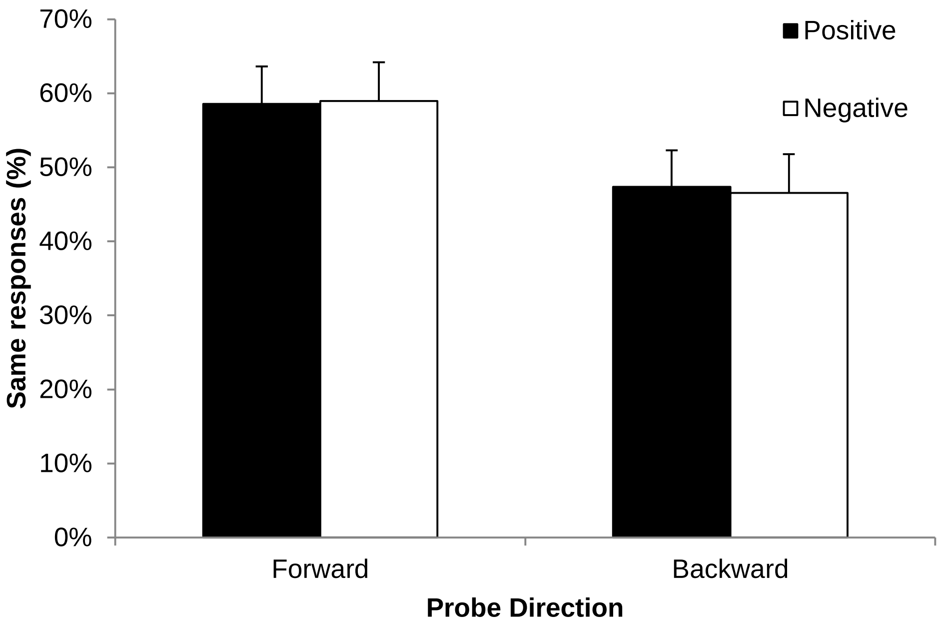


Figure 9. Mean percentages of 'same' responses to forward' and 'backward' probes, for both intentional (positive verbal statement, black bars) and non-

intentional actions (negative verbal statement, white bars) in Experiment 3. Error bars represent 95% confidence intervals.

Discussion

Experiment 3 followed the results of the first two experiments by successfully finding a representational momentum effect. Using different stimuli, participants were once again consistently more likely to misperceive the position of an actor's arm when it was further forward in the action sequence than when it was further backward. The representational momentum effect was also evident in the reaction times, with backward displacements resulting in significantly faster detections than forward displacements. However, the verbal cues conveying the intention of the action did not affect representational momentum, with both intentional and non-intentional verbal statements resulting in virtually identical levels of representational momentum. While language is a clear tool in the communication of intentionality more generally, it did not translate to the perception of low-level actions here.

Overall, while Experiment 3 revealed a representational momentum effect both for judgements and reaction times, supporting the findings of the first two experiments, it failed to show that intentional verbal cues modulated the effect. However, the lack of any explicit requirement to process the verbal cues or any method to measure the degree to which they were processed, makes it difficult to know whether the lack of an effect of intention was due to limitations in the experimental design or because verbal cues to intention do not modulate top-down predictions.

Experiment 4 – Gaze cue

While the first three experiments all demonstrated a representational momentum effect, they failed to demonstrate a reliable effect of intention based on either kinematic or verbal cues. Consequently, the next experiment applied the paradigm to a new intentional cue, eye gaze.

Eyes are not just the window to the soul but are also the window to other people's actions and intentions (Frischen, Bayliss & Tipper, 2007). Before performing an action, actors typically look at the endpoint of the action, rather than at the movement's path (Land, Mennie & Rusted, 1999; Flanagan & Johannson, 2003), and observers use these cues to infer the action's goal (Rotman, Troje, Johansson, & Flanagan, 2006). For example, studies have demonstrated that merely seeing somebody else look at an object can prime actions towards that object, as if one were observing the action (Castiello, 2003). In addition, seeing somebody else looking at an object has been shown to result in similar brain activity to that produced when observing actions directed to the same object (Pierno et al., 2006, 2008). These studies suggest that eye gaze does not just alert us to the intentions of other people, but it also allows us to anticipate the particular actions that the person may initiate. This further suggests that eye gaze is an effective cue in conveying the intentions of other people. Just as we use gaze as a key tool to guide our own actions, we process other people's gaze, linking its direction to potential subsequent actions.

In order to test the influence of gaze as a cue to an actor's intention, the representational momentum paradigm was applied to actions showing an actor performing congruent reaches towards an object. Crucially, participants were shown an actor gaze either towards or away (towards the other side of the table) from an object before reaching towards the same object. While gaze towards an object signals a potential intention to grasp the object, gazing away from an object does not. Once again based on the assumption that intentional actions should generate a top-down prediction of the action, it was hypothesised that representational momentum would be greater when reaches followed a gaze towards the object, compared to when the gaze was directed away from the object prior to action onset.

Method

Participants. 39 students (6 male, mean age = 19.9, SD = 3.2) at Plymouth University or members of the wider Plymouth community took part in exchange for participation points or payment (£8 p/h). All had normal or corrected-to-normal vision and were native English speakers. All provided written informed consent prior to participation and were debriefed once the experiment had finished. The study was approved by the University of Plymouth's ethics committee.

Stimuli and Apparatus. The stimuli were filmed using a Canon Legria HFM36 HD video camera and edited using Moviedek and Corel PaintShop Photo Pro x3. Stimuli consisted of an actor sitting opposite the camera and reaching with their right arm diagonally across the table to one of seven target objects (book, cup, bottle, pen, mobile phone, tape measure, and wristwatch) located on the actor's left and the viewer's right (see Figure 10, panel C). For each object two separate actions were filmed, one where the actor looks towards the object before reaching for the object (gaze toward/intentional) and one where the actor looks away from the object, towards the other side of the table, before reaching for the object (gaze away/non-intentional). This resulted in 16 different action sequences, half where the actors gaze was toward the object and half where it was directed away from the object (See figure 9, panel A). Individual frames were extracted from each action, and for each reach 12 images were chosen which spanned the trajectory, from the start of the reach to approximately two-thirds of the way through the reach.

Procedure. Participants were seated in a cubicle approximately 60 cm from a colour monitor, and were given written and verbal instructions. Each trial began with a fixation cross, presented for 500ms, followed by a blank screen also presented for 500ms. The first frame of the action sequence was then presented for 200ms, followed by three subsequent frames in the action sequence, each presented for 200ms. Following the last frame of the sequence a grey image of equal overall luminance was presented for 700ms, before a probe stimulus was then presented. The probe stimulus could either be the same as the final frame of the action sequence ('same' probe), further along the observed trajectory ('forward' probe), or in a position just prior to the final frame ('backward' probe). Participants were asked to judge, as quickly and accurately as possible, if the probe stimulus showed the arm in the 'same' or 'different' position as its last position prior to the probe. They made their judgment by pressing one of two possible response keys ('s' for 'same' and 'd' for 'different'). If no response was made within 3 seconds, a prompt was displayed that they were 'too slow' along with a reminder of the response button assignment. The experiment consisted of 16 training trials (not analysed), followed by 294 experimental trials, split by two break periods. Each stimulus combination was shown once (7 starting points, 2 gaze directions, 3 probe locations, and 7 objects), and their order was randomised. The experiment lasted approximately 30 minutes.

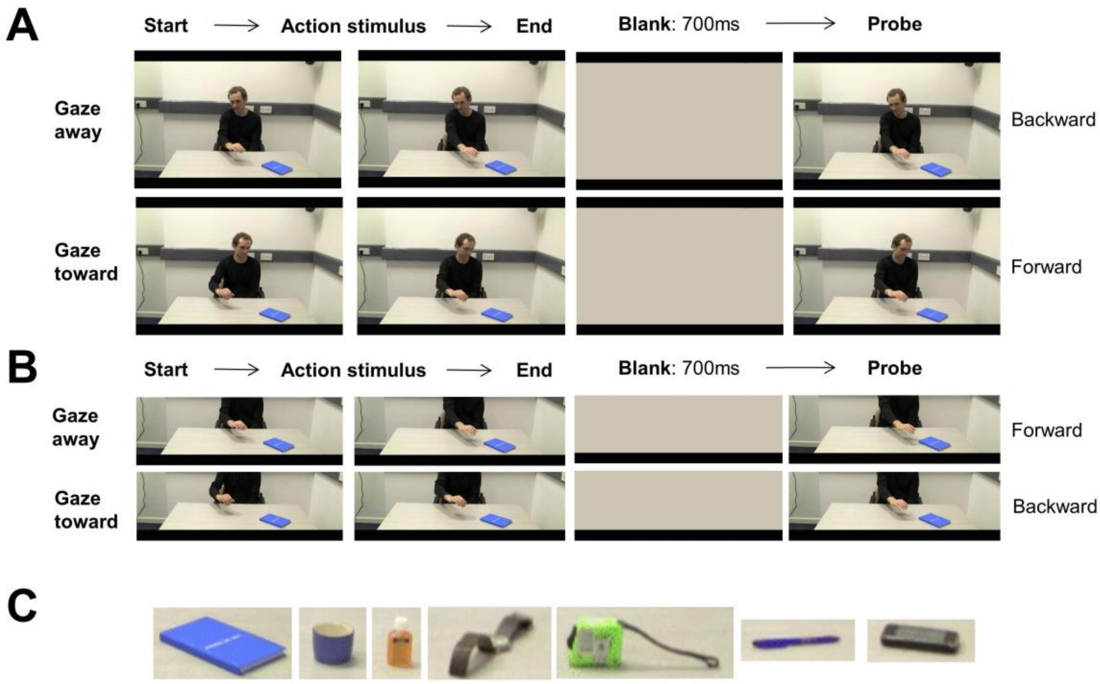


Figure 10. A: Trial sequence from Experiment 4 of reaches for each condition (gaze away and gaze toward). A grey blank screen followed, before the probe stimulus was presented. An example of a 'backward' and 'forward' probe are presented. B: Trial sequence from Experiment 5 of reaches for each condition when the top half of the image has been removed (gaze away and gaze toward). A grey blank screen followed, before the probe stimulus was presented. An example of a 'backward' and 'forward' probe are presented. C: The seven target objects used in both Experiment 4 and 5 (book, cup, bottle, wristwatch, tape measure, pen and mobile phone). Sizes have been modified for presentation purposes. In the experiment they were presented at their actual size.

Results

Exclusions. Out of the 294 experimental trials, only the trials where the probe was a frame that across the experiment could occur in each of the three probe positions ('same', 'forward' and 'backward') were included. This resulted in 168 trials being analysed, but the results did not differ when the full dataset was analysed. Four participants' data were removed from the analysis because these participants failed to respond in more than 20% of trials. Inclusion or exclusion of these participants did not affect the results. Trials where reaction times were longer than 2000ms were also excluded from the analysis (<2% of all trials).

Sensitivity. Analysis of perceptual sensitivities confirmed that participants were able to distinguish 'same' from 'different' probes. They identified more probes as 'same' when they really were at the same locations as the last image, compared to the average of 'backward' or 'forward' probes, $t(34) = 8.41$, $p < .001$. Perceptual sensitivities were also shown to differ significantly depending upon the intention of the action, $t(34) = 2.06$, $p = .047$. This showed that perceptual sensitivity was higher when viewing intentional actions compared to non-intentional actions.

Representational Momentum. The percentage of 'same' responses for the 'forward' and 'backward' probes was entered into a 2x2 repeated measures ANOVA Intention (intentional/gaze toward or non-intentional/gaze away), and Probe direction ('forward' vs. 'backward') as factors. The analysis revealed a significant effect of Probe direction, $F(1,34) = 19.28$, $p < .001$, $\eta_p^2 = .362$, 95% CI [8, 21], with participants more likely to judge 'forward' probes as 'same' compared to 'backward' probes, confirming the classical representational momentum effect. In addition, the analysis revealed a significant effect of Intention, $F(1,34) = 20.28$, $p < .001$, $\eta_p^2 = .374$, 95% CI [5, 14], with participants more likely to judge 'different' probes in the non-intentional condition as 'same' than those in the intentional condition. Importantly, as predicted, there was also a significant interaction between Probe direction and Intention, $F(1,34) = 10.03$, $p < .005$, $\eta_p^2 = .228$, 95% CI [3, 12], demonstrating that the tendency to mistake more 'forward' than 'backward' probes as 'same' was larger for intentional than non-intentional actions (see Figure 11).

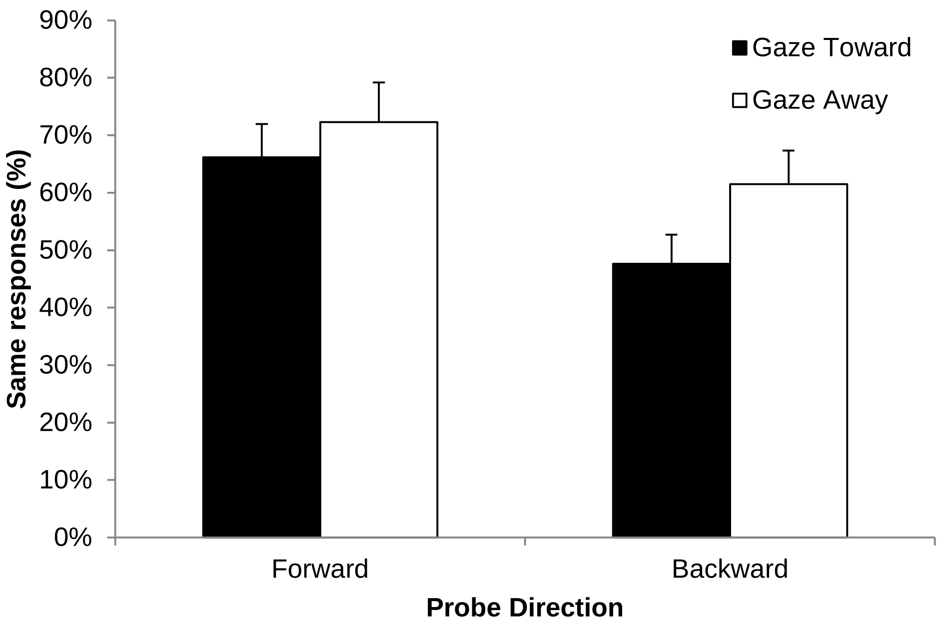


Figure 11. Mean percentages of 'same' responses to forward' and 'backward' probes, for both intentional (gaze toward, black bars) and non-

intentional actions (gaze away, white bars) in Experiment 4. Error bars represent 95% confidence intervals.

Reaction times. The same 2x2 used for the representational momentum analysis was applied to the reaction times for correct responses. One participant was removed from the analysis as they failed to make a correct response in some of the conditions. The analysis revealed a significant effect of Probe direction, $F(1,33) = 6.30$, $p = .017$, $\eta_p^2 = .160$, 95% CI [11, 89], with correct responses to 'backward' probes significantly faster than correct responses to 'forward' probes. There were no other significant effects, all *F's*< 1.4.

Discussion

The results of Experiment 4 once again demonstrate a representational momentum effect when observing the actions of others. Participants were more likely to misperceive the position of an actor's hand when the probe stimulus was further forward along its trajectory than when it was further backward. This tendency was found across gaze conditions demonstrating the bias to predict the future position of an actor's hand regardless of gaze information. This supports the earlier studies in this chapter, and prior research demonstrating people's tendency to predict the actions of others ahead of real time (Graf et al., 2007).

Most importantly, this tendency was greater when the actor gazed towards the object, compared to when they gazed away from it, prior to action onset. This demonstrates the importance of gaze when observing the actions of others (Flanagan & Johannson, 2003; Frischen et al., 2007) and provides support for its role in generating top-down predictions of other people's intentions and therefore future actions. It supports top-down theories of social perception by illustrating that when social cues such as eye gaze imply an action's intention the tendency to misperceive the position of an actor's hand further forward in time increases. It would provide key evidence that perception can be modulated by intentional information in line with the idea that cues to the intention of an observed action induce top-down predictions and bias participants' perception.

While the findings of Experiment 4 showed that perceptual prediction increases when the actor's eye gaze signals their intention, the results could be explained by differences in action kinematics alone. During action execution, eye gaze informs action control, so that the actions in the intentional condition may not have been identical to those within the non-intentional condition, even though the model was instructed to perform both equally fluently. When performing the gaze away actions, the model could not see the object, which is likely to have had an effect on the fluency, accuracy and kinematics of the action compared to when the gaze was directed to the object and the hand was fully visible. It is therefore possible that the increase in representational momentum resulted from subtle kinematic differences between the actions, rather than the gaze cues. In order to address this concern, a control experiment was run showing the same actions with all gaze cues removed.

Experiment 5 – Gaze cue control

In order to test whether the results of Experiment 4 resulted from differences in the actor's gaze rather than their kinematics a control experiment was run. Experiment 5 was identical to Experiment 4 with the exception of the stimuli presented. The top halves of all images presented in Experiment 4 were removed, removing all head and gaze cues, and leaving only the arms and chest of the actor present. This allowed a test of whether kinematic differences between the arm movements in the gaze towards and gaze away conditions were responsible for the differences in representational momentum.

If representational momentum does not significantly differ between the gaze toward or away actions, this would support the assumption that the direction of gaze was responsible for the results of Experiment 4. If, however, the kinematics belonging to the gaze toward action results in a significant increase in representational momentum compared to the kinematics belonging to the gaze away action, as in Experiment 4, then this would show that the differences between the conditions are due to the kinematics of the actions rather than the direction of gaze.

Method

Participants. 31 students (5 male, mean age = 20.3, SD = 5.2) at Plymouth University or members of the wider Plymouth community took part in exchange for participation points or payment (£8 p/h). All had normal or corrected-to-normal vision and were native English speakers. All provided written informed consent prior to participation and were debriefed once the experiment had finished. The study was approved by the University of Plymouth's ethics committee.

Stimuli and Apparatus. The same stimuli used in Experiment 4 were digitally edited by removing the top half of each image, resulting in identical action sequences but without access to the direction of the actor's gaze (see Figure 10, panel B).

All other aspects of the method were identical to Experiment 4.

Results

Exclusions. Out of the 294 experimental trials, only the trials where the probe was a frame which across the experiment could be classed as all three of the probe positions ('same', 'forward' and 'backward') were included. This resulted in 168 trials being analysed. Two participants' data were removed from the analysis because these participants failed to respond in more than 20% of trials. Inclusion or exclusion of these participants did not affect the results. Trials where reaction times were longer than 2000ms were also excluded from the analysis (<2% of all trials).

Sensitivity. Analysis of perceptual sensitivities confirmed that participants were able to distinguish 'same' from 'different' probes. They identified more probes as 'same' when they really were at the same locations as the last image, compared to the average of 'backward' or 'forward' probes, $t(28) = 12.18$, $p < .001$. Perceptual sensitivities were also shown to differ significantly depending upon the intention of the action, $t(28) = 2.105$, $p = .044$. This showed that perceptual sensitivity was higher when viewing intentional actions compared to non-intentional actions.

Representational Momentum. The percentage of 'same' responses for the 'forward' and 'backward' probes was entered into a 2x2 repeated measures ANOVA Intention (intentional/gaze toward or non-intentional/gaze away), and Probe direction ('forward' vs. 'backward') as factors. The analysis revealed a significant effect of Probe direction, $F(1,28) = 31.20$, $p < .001$, $\eta_p^2 = .527$, 95% CI [9, 20], with participants more likely to judge 'forward' probes as 'same' compared to 'backward' probes, confirming the classical representational momentum effect. In addition, the analysis revealed a significant effect of Intention, $F(1, 28) = 30.30$, $p < .001$, $\eta_p^2 = .520$, 95% CI [6, 12], with participants more likely to judge 'different' probes in the non-intentional condition as 'same' than those in the intentional condition. Importantly, as predicted, there was also a significant interaction between Probe direction and Intention, $F(1,28) = 4.55$, $p < .05$, $\eta_p^2 = .140$, 95% CI [0, 11], demonstrating that the tendency to mistake more 'forward' than 'backward' probes as 'same' was larger for intentional than non-intentional actions (see Figure 12).

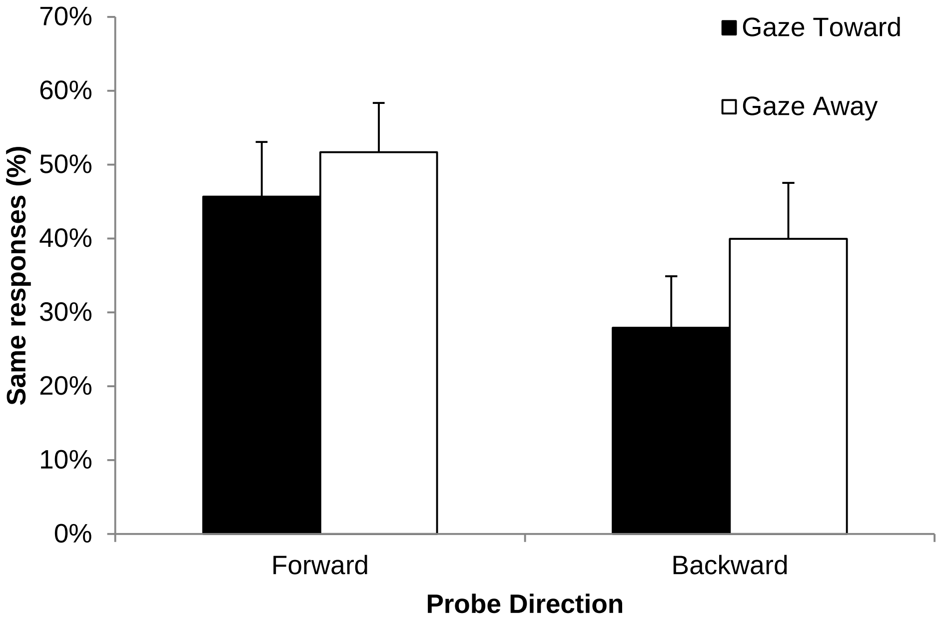


Figure 12. Mean percentages of 'same' responses to 'forward' and 'backward' probes, for both intentional (gaze toward, black bars) and non-intentional actions (gaze away, white bars) in Experiment 5. Error bars represent 95% confidence intervals.

Reaction times. The same 2x2 used for the representational momentum analysis was applied to the reaction times for correct responses. The analysis revealed a significant effect of Probe direction, $F(1,28) = 9.32$, $p = .005$, $\eta_p^2 = .250$, 95% CI [27, 127], with correct responses to 'backward' probes significantly faster than correct responses to 'forward' probes. There were no other significant effects, all *F's*< 1.2.

Discussion

The results of Experiment 5 replicated the findings of Experiment 4. Participants were again more likely to misperceive probes further forward along the reach trajectory with the hand's last seen position, than probes further backward, demonstrating the classical representational momentum effect. Crucially, the kinematics of reaches, which belonged to actions when the actor gazed at the object, produced more representational momentum than the kinematics of reaches, which belonged to actions when the actor gazed away from the object, despite the actual gaze information not being present. This complete replication of the results of the previous experiment, despite the lack of gaze information, strongly suggests that the differences in representational momentum between the conditions were due to kinematic differences between the reaches rather than due to the direction of the gaze.

As the intention of the action never varied in the current experiment the difference between the two conditions cannot be explained by different top-down predictions indicating different intentions. However, as both action conditions were related to slightly different bodily positions (due to different gaze directions) they are separable in relation to the biomechanical efficiency of their reach trajectories. Kinematics belonging to gaze toward actions are likely to have been more biomechanically efficient due the actor's gaze toward the object. Conversely, kinematics belonging to gaze away actions are likely to have been less efficient due to the actor's split biomechanical tasks. Therefore the differences in representational momentum here are likely to be due to differences between the biomechanical efficiency of the two action conditions.

Overall, the results of Experiment 5 demonstrate that the findings of Experiment 4 were likely to have emerged due to differences in the kinematics of the actions observed rather than due to differences in the direction of gaze. While this does not negate the influence of gaze on kinematic efficiency, it does show that here kinematic differences were sufficient to explain the differences found in Experiment 4. This demonstrates that differences in bottom-up sensory inputs can modulate perception even when the explicit intention of the actor is the same.

General Discussion

The aim of the initial experiments was to establish a robust representational momentum effect during action observation and explore whether it was modulated by social cues provided by kinematic, verbal or gaze information that implied the action's intention. All the experiments were consistent in finding a representational momentum effect, supporting previous research demonstrating the effect during action observation (Wilson et al., 2010). Moreover, the current set of experiments extends previous research by showing that representational momentum also applies to object-directed actions and persists across a number of different social stimuli and task set-ups. These findings support the idea that we perceptually predict the future location of a moving body part when observing the actions of others in a similar manner to a moving inanimate object (Wilson & Knoblich, 2005).

While intentional grip-object matches showed increased perceptual prediction during action observation in Experiment 1, this effect could not be replicated in Experiment 2. However, the later finding that kinematic differences between actions (Experiments 4 and 5) with the same intention, modulated perceptual prediction based on biomechanical efficiency suggests that kinematic cues do effect perceptual prediction, at least on some level. .

While this could be seen as a challenge to top-down theories of social perception it is important to remember that the differences in kinematics in Experiment 5 did not specify different intentions. Therefore, the findings of Experiment 4 and 5 cannot be explained by a bottom-up explanation of social perception. So while they could be taken as support for bottom-up processing during action perception this would be independent of the intention of the action. Moreover, top-down theories of perceptual prediction would predict that optimal reaches to a goal will match top-down predictions to a greater extent than sub-optimal reaches (Jacquet et al., 2012), and lead to more perceptual prediction, and therefore representational momentum. In this way while these findings challenge top-down theories to some extent, they could also be taken as evidence for them. For these reasons they can also not be seen as evidence for a bottom-up explanation of social perception.

The finding that kinematics can modulate perceptual prediction supports the aims of Experiments 1 and 2 and suggests that

the lack of a reliable intention effect for the grip and object matches could be due to a number of different reasons. One issue with these experiments concerned the different grip types, which were somewhat ambiguous. For example, all reaches ended well in advance of where the object was located, so it could be that for some participants incongruent hand grips were not perceived as such, and seen as still able to change course and grip the object. While secondary questionnaires confirmed a distinction between the different grips in line with expectations, over a third of the non-intentional actions were still viewed as directed towards the object. Therefore it appeared possible that the grip-object matches were too subtle to elicit perceptual predictive processes. This was clearly overcome in Experiment 5 by making the intention of all actions unambiguous allowing the biomechanical efficiency and fluency of the action to inform perceptual prediction alone.

However, a number of studies have shown that participants use information relating to hand-object matching as cues to intention (Yoon & Humphreys, 2005; Bub, Masson & Cree, 2008; Fischer et al., 2008). Yet here, the experiment involved no explicit need to process the object, and as this is thought to be key to engage affordance processing, it may have reduced any effect of intention. Indeed, research has shown that information relating to the integration of grip and object as opposed to grip information alone is the primary driver of action processing (Bach, Knoblich, Gunter, Friederici, & Prinz, 2005; van Elk et al., 2008). This suggests that future experiments need to force participants' attention towards the object in order to ensure that object affordances are engaged and consequently that the intention of the action is processed.

A number of possible reasons could have been responsible for the failure to find an effect of intention in relation to the verbal statements. Firstly, it may have been that the verbal statements were not processed sufficiently. Like the grip-matching experiments, the task did not require the verbal statements to be explicitly processed. That is, as the verbal cues were not task-relevant and were presented at the same time as the action sequence, participants may have consciously ignored them, despite being instructed to process them. The assumption was that this would be difficult due to the volume of the auditory stimuli and the reduction in extraneous sounds because of the presence of the headphones. Moreover, even if participants did consciously ignore them, it was assumed that they would be processed on some level automatically (for a review, see Shtyrov, 2010). However, as there was no way to measure this, it is impossible to know to what degree these verbal statements were actually processed.

Another possibility is that the verbal cues were merely not distinct enough, and were therefore not associated with the viewed action. While this is possible, the verbal cues were purposefully created to be as emotionally salient as possible. Alternatively, it might have been merely that the repetition of the verbal cues, and actions associated with these verbal cues, across the span of the experiment led to a decrease of attention, thereby undermining the association between verbal cue and action. However, when analysing the effects over the course of different blocks no such decline was evident. Lastly, it may be that the verbal statements were too complex to affect perception. Therefore, future use of verbal statements would need to be explicitly linked to the actions they refer to in order for their processing to be assured. Likewise, it may also be that verbal statements need to be simpler in order to facilitate processing. Such changes would allow the effect of verbal information as an intentional cue to be better addressed.

As mentioned above, the experiments investigating the influence of gaze direction on action prediction seemed to show that gazing towards an object increased subsequent perceptual prediction. However, a subsequent control study identified that this was due to differences in the kinematics of the actions rather than the direction of gaze. This highlights a problem with comparing natural reaches to an object that follow different gaze directions, as there will obviously be differences in kinematic information. This could be overcome by adopting artificial stimuli that keep the kinematics constant while varying the gaze direction, although such a design would restrict the ecological validity of such a study.

Other issues concerning the current experiments relate to the presentation of the stimuli. In all the experiments, while the starting point of the action was varied, the length of the action sequence throughout an experiment was always constant (3 or 4 frames). This entailed the danger of reducing the requirement for prediction more generally. Participants were certain that, once the final frame was reached, no further movement would happen, and there was no need to predict further ahead. Another factor, which may have reduced prediction, was the predictability of the spatial dynamics of the whole scene. As all objects were positioned in the same spatial location on the screen, all actions had the same endpoint. Participants therefore did not have to predict the hand's path based on the object in the scene, but could just rely on their memory of the previous reaches. This may have further reduced the requirement for prediction. In effect, participants knew in advance where each action would stop so that there was no requirement to attend to the visual cues at all.

In conclusion, while the experiments presented here provide support for the perceptual extrapolation of biological motion during action observation, they failed to find reliable evidence that top-down predictions are elicited by cues, which signal the actor's intention. Instead, they provide some evidence that differences in kinematics can modulate perceptual prediction independently of the actor's intention. While the lack of a consistent effect of intention challenges top-down theories of social cognition, the effect of kinematics does suggest that differences in biomechanical efficiency can inform perceptual prediction. Moreover, the failure to find any reliable top-down intentional effects could be due to a number of different factors regarding the methodology across the different experiments. Therefore, based on this, numerous changes were applied to the paradigm in an attempt to better address the role of top-down social cues of intention on perceptual prediction.

Chapter 4 - The effect of verbal cues on action observation

The experiments in Chapter 3 demonstrated a reliable representational momentum effect across a range of different stimuli and experimental designs. However, these experiments failed to find consistent evidence that the effect was modulated by kinematic, verbal or gaze cues to intention. This calls into question the influence of intentional social cues on predictions implied by top-down theories of social perception (Csibra, 2007; Kilner et al., 2007). Moreover, Chapter 3 also revealed evidence that changes in the kinematics of a reach during action observation can modulate representational momentum, in the absence of intentional cues. This provides evidence that changes in bottom-up sensory inputs could drive predictive processes rather than top-down predictions. However as the kinematic differences were not clearly linked to different intentions they do not support the direct matching hypothesis of social perception either (Rizzolatti & Singaglia, 2010). Instead they suggest that more tightly controlled experiments are really required to tackle the influence of social cues to intention on perception. Indeed, a number of different methodological aspects of the experiments in Chapter 3 may have contributed to the failure to find an effect of intention on action prediction. Therefore, a number of changes were made to the paradigm in an attempt to investigate the role of top-down cues to intention more effectively.

One serious concern with the grip-matching experiments (Experiments 1 and 2) was the lack of an explicit requirement to process the objects' orientation, and therefore to derive the actor's intention. Additionally, as reaches stopped relatively far from the object, the distinction between intentional and non-intentional actions was potentially ambiguous. Therefore, here, grip-matching was replaced with language cues to convey the action's intention. While the previous verbal experiment (Experiment 3) suggested that such language cues did not increase representational momentum, it suffered from a similar issue as the grip experiments, in that there was no explicit requirement to process the language cues. Therefore, a solution to this issue was to ask participants to produce the verbal statements themselves, as if they were instructing the agent, thus ensuring that the cues were processed explicitly.

In the experiments in Chapter 4, actions now varied with respect to their movement direction. Each action could either be a reach for or a withdrawal from a goal object. To generate differential action expectations, the goal objects were either painful or safe. This produced a dichotomy regarding the intentions likely to be associated with each object that could match or mismatch with action type. While safe objects are more directly associated with the intention of reaching forward to grasp, painful objects are more likely be associated with withdrawals (Anelli, Borghi & Nicoletti, 2012). To trigger each action, participants were required to state, as if they were instructing the actor, the appropriate action for the object. For safe objects, participants had to make verbal statements implying an approach goal (Experiment 6 – "Forward", Experiment 7 – "Take it") and an avoidance goal for painful objects ("Backward" and "Leave it"). This ensured that participants explicitly processed the object and also created an expectation about which action they expected to see (e.g. reach toward for a safe object and withdrawal away for a painful object). The action could then either follow the instruction or do the opposite. By independently varying both action direction (reach, withdrawal) and the object type (painful, safe), the effect of verbal statement on representational momentum could be measured.

Another potential problem with the verbal statements used in the previous experiment was their complexity. The statements were expressed emotionally and were also in many cases several words long. Both these factors may have meant that their meaning was not processed fast enough to influence the representational momentum task. By replacing these verbal statements with categorical statements (forward/backward, or take it, leave it), which were simpler, more direct and less ambiguous, it was hoped that a more direct link with the observed actions could be created. Moreover, by linking these categorical statements to two action directions, the relationship between verbal cue and subsequent observed action was more apparent and increased the likelihood that differential expectations would be elicited.

In addition to these key changes a number of other slight adjustments were made to the representational momentum task.

Firstly, while the experiments in Chapter 3 all demonstrated a robust representational momentum effect, a potential concern was that the directionality of the response keys (e.g. 'z' left and 'x' right) could affect the responses, in a Simon-like manner (Simon, 1969). While counterbalancing addressed this issue in the previous experiments, the representational momentum task here was changed from a forced choice task to a GO/NOGO paradigm to fully rule out this potential influence. Rather than having to respond to every probe with either a 'same' or 'different' response, participants were required to make a response only when they thought that the probe hand was in a different position to its last location, and not respond if they judged it to be in the same position. The 'different' probe was chosen for GO responses based on the assumption that inputs conforming to an action prediction should be attenuated, while deviations should become more salient. Therefore, the perception of 'different' probes would reflect the 'prediction error' signal, which in theories of predictive processing is central, and thought to alert one to the disconfirmation of a previous prediction (Friston, 2010; Clark, 2013).

In addition to this, as mentioned in Chapter 3, while the action sequences included different start points, sequences always included the same number of frames, which is likely to have reduced the requirement to predict, as participants always knew in advance when the movement would terminate. Subsequent pilot testing indeed revealed that varying the length of the sequence as well as the start point increased representational momentum, thus also increasing the misperception of 'forward' probes as the same as the last seen image. Indeed, in most cases the shortest sequences produced the most representational momentum, whereas it was almost eliminated in the longest sequences, where participants could be sure that the motion would not continue. Therefore, the paradigm was altered to include 3 different sequence lengths (between 3 and 5 frames) so that participants were less able to anticipate the stopping of the sequence.

Similarly, in the experiments of Chapter 3, the actions and goal objects were always located in the same spatial position on screen, resulting in identical end points. This meant that participants could rely on their memory of previous reaches rather than predict the hand's path in relation to the object. To rectify this, the position of the scene was shifted randomly along the midline from trial to trial, and all background information was digitally removed. In this way, the hand's likely path could only be derived from the position of the object on screen, and not from one's memory of the last reach. Accurate judgement of the movements would now only be possible if participants, on a trial by trial basis, predicted the movement they would see, based on the initial positions of hand and target object.

Finally, the evidence from the experiments in Chapter 3 suggested that the representational momentum task was sometimes difficult for participants to complete. While overall sensitivities showed that participants could differentiate between the 'same' and 'different' probes, closer inspection of the single subjects revealed that, in each experiment, a large number of participants showed no differences between probe types, suggesting random response strategies. To rectify this, training blocks were added in order to a) familiarize participants with the task and b) add an objective means for participant exclusion prior to the experiment. In addition, catch trials with extreme probe displacements were interspersed into the main task to measure task performance and compliance, providing an objective means for exclusion.

The first three experiments of the chapter (6, 7 and 8) investigate how expectations of intention generated by self-performed speech affected action perception as measured by the representational momentum task. In the subsequent two experiments (9 and 10), the paradigm is extended to test whether representational momentum is affected if expectations are generated not through self-performed speech but heard speech, when actors state their own intentions.

Experiment 6 – Spoken verbal cue – “Forward” and “Backward”

The aim of Experiment 6 was to implement the raft of changes to the paradigm. As in the experiments from Chapter 3, participants observed reaches towards objects, but this time also observed withdrawals away from them. The objects could either be safe or painful to grasp. The representational momentum task was the same as in the earlier experiments, but this time the participants only responded – with a press on the space-bar – when they judged the probe hand to be ‘different’ from its last position and pressed nothing if they judged it to be the ‘same’. With this shift from forced choice to GO/NOGO decision, the response data in effect measured the ‘prediction error’ signal. As before, the expectation was that ‘backward’ probes would be easier to correctly identify as different from the hand’s last location than ‘forward’ probes. Representational momentum was measured by how much greater the ability to identify a ‘backward’ probe as ‘different’ proved to be, compared to a ‘forward’ probe.

In order to investigate the effect of intentional verbal cues, participants were required to respond verbally to the type of object by verbalizing the appropriate action with the object, which would then initiate the start of the action sequence. We hoped that this direct temporal link between verbal statement and action onset would increase the perception of causality between statement and action. If the object was safe to grasp, participants said “Forward” while if the object was painful, participants said “Backward”. This served to create a meaningful context for the participants’ utterances, but also crucially aimed to produce an expectation of which action the participant would subsequently expect to see. The task therefore fused the social cues of objects and language, with the aim of eliciting the maximal effect on representational momentum. It was expected that representational momentum would be greatest whenever the verbal response matched the subsequent action direction. Therefore, representational momentum should be larger for reaches towards safe objects and reaches away from painful objects, than for reaches towards painful objects and reaches away from safe objects.

Method

Participants. 45 participants took part in Experiment 6. They comprised students from Plymouth University or members of the public from the wider Plymouth community. They took part in exchange for participation points or payment (£8 p/h). All had normal or corrected-to-normal vision, were native English speakers and right-handed. All provided written informed consent before participating and were debriefed at the end of the experiment. The study was approved by the University of Plymouth's ethics committee. An initial training/calibration session (see procedure) took place prior to the proper experiment, and those who failed this did not proceed to the experimental session. Five participants were excluded based on the training session, leaving 40 participants (26 females, mean age = 23.3 years, SD = 8.9) to proceed to the experimental session.

Apparatus and Stimuli. A Canon Legria HFS200 video camera was used to film stimuli, and edited using Moviedek and Corel PaintShop Photo Pro x6. Stimuli were filmed at 30 frames per second and consisted of natural reaches of a man's right hand towards one of four target objects (drinking glass, wine glass, plastic bottle, knife with handle oriented toward hand; see Figure 13, panel A). Individual frames were extracted from each video, and for each reach 26 images were chosen which spanned the trajectory from the start of the reach to approximately two-thirds of the way through the reach. For all images the object and hand were superimposed onto a black background to eradicate any other details from the scene. Each object was paired with a painful object that was matched for grip type and size (broken glass, broken wine glass, cactus, knife with blade oriented towards hand; see Figure 13, panel A). Each set of images was duplicated with each original object digitally replaced with its painful object pair, resulting in 2 identical sets of reaches with only the target object different. The result was 8 sets of 26 images, each containing a different target object (4 non-painful and 4 painful) and including 4 different reaches, each duplicated once. Each displayed action sequence began at a randomly chosen frame from the middle of the action (between frames 13-17). The length of each action sequence varied between 3 and 5 frames and proceeded in 2-frame steps, either towards the object (forward reaches) or away from the object (backward withdrawals). Withdrawal sequences consisted of playing the frames in reverse order.

Like the earlier experiments, the probe image could either be the same as the last image of the action sequence ('same' probe), or different. When the probe was different, it could either be further forward along its trajectory in a future position ('forward' probe) or back along the trajectory in a past position ('backward' probe). Different probes varied between 1 and 2 frames in distance from the last image of the action sequence (varied between participants based on their calibration performance, see below) in experimental trials and 4 frames in catch trials (see Figure 13, panel C). The stimuli were presented on a 22-inch Philips Brilliance 221P3LPY monitor (resolution: 1920 X 1080, refresh rate: 60 Hz) using Presentation software (Neurobehavioural Systems, Inc.). Participants wore Logitech PC120 headphones with a microphone.

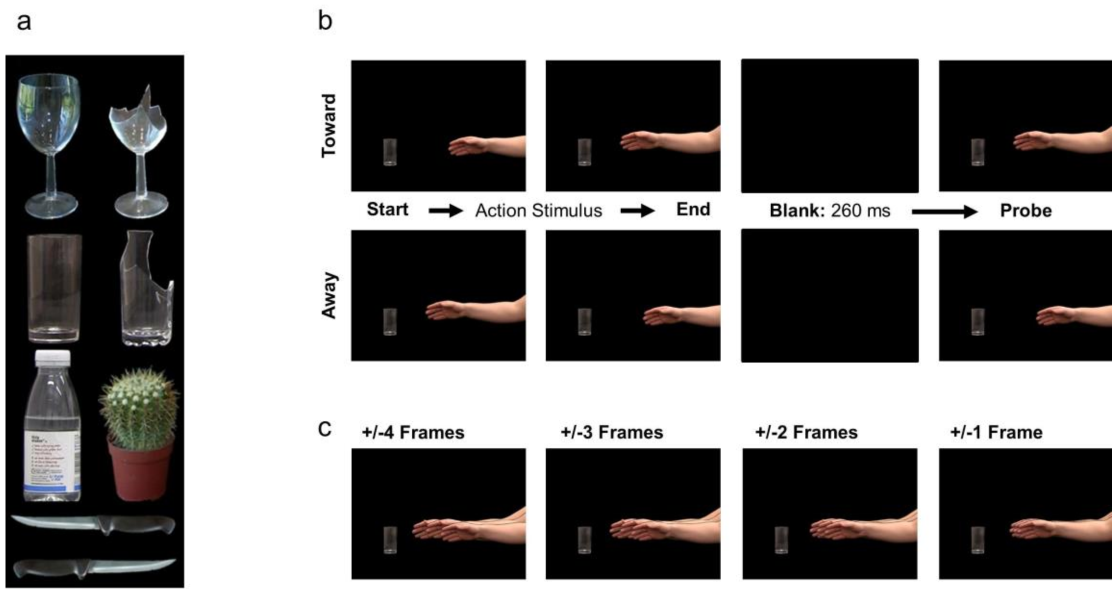


Figure 13. A: The safe objects (left column) and the paired dangerous objects (right column), and the knife oriented safely or dangerously with respect to the hand (bottom). B: Trial sequence of a reach towards (top) or away from (middle) an object (action stimulus), followed by a blank screen, and then the probe stimulus. In this example, both actions finish on the same frame, and the probe position is the same as the final action stimulus frame. C: The probe stimulus levels. In each image, the centre hand is the same as the final action stimulus frame in B, and the 'different' probe stimuli are superimposed either side of it. For reaches toward the object, the probe nearest the object was the 'forward' probe and the probe farthest from the object was the 'backward' probe. For reaches away from the object, the probe farthest from the object was the 'forward' probe and the probe nearest the object was the 'backward' probe. The difference between the 'same' and 'different' probes decreases across the images from left to right (4 frames, 3 frames, 2 frames, 1 frame).

Procedure. Participants were seated approximately 60 cm from a colour monitor, and were given written and verbal instructions. The experiment began with a training session comprising 4 training blocks (each 36 trials) that measured participants' ability to discriminate between the different types of probe. This session was added due to the finding from earlier experiments and pilots that the representational momentum task is difficult for participants. The first block contained 'different' probes that were 4 frames from last image, and with each subsequent block they became closer to the 'same' probe (+/-4, +/-3, +/-2, +/-1, see Fig. 13). This meant that in the first block 'different' probes were easier to judge, and with each block the task became harder.

Each trial began with a fixation cross displayed for 500ms, followed by a blank screen presented for a randomly chosen time between 500 and 1000ms. Following this, the first image of the action sequence was presented for 1500ms. All action sequences were presented in the same vertical position, but the position along the X-axis varied across trials, thereby increasing variation. Each subsequent frame was displayed for 80ms and depicted either a reach towards the object or a withdrawal away from the object. After this a blank screen was presented for 260ms before a probe image was displayed (distance of 'different' probes varying as specified above).

The participants' task was to press the space-bar with their left hand if they thought the probe was in a different position from its last position, and not respond if they thought it was in the same position. They had 4000ms to respond.

Participants who did not meet the inclusion criteria did not proceed to the main experiment (see Results – Exclusions). The main experimental session was identical to the training session, apart from one important change: the requirement of a verbal response to initiate the task. The first image of the action sequence would stay on screen until a verbal response was registered (detected via Presentation's sound threshold logic). Participants were instructed to verbally respond depending on whether the object was safe or painful to grasp, and were asked to say "Forward" to initiate the action sequence for safe objects and "Backward" for painful objects.

1000ms after the verbal response was registered, the action sequence, either reaching towards or withdrawing away from the object, would begin, and participants would complete the representational momentum task. The experiment consisted of 144 experimental trials with each stimulus combination shown once (4 objects X 2 object types X 2 action directions X 3 movie lengths X 3 probe types). In addition 24 catch trials (CTs) where the probe type was either +4 or -4 frames from the final position were included. The experiment was split into 3 blocks of 56 trials with self-terminated breaks between blocks.

Inclusion criteria. In the training session, participants were rated on two measures, accuracy and sensitivity. Accuracy was merely the average of correct responses across all probe types. Sensitivity was calculated by subtracting the percentage of 'same' responses when the probe was 'different' from the percentage of 'same' responses when the probe was the 'same'. If participant performance within a block dropped below chance on either accuracy (50%) or sensitivity (0%) the block was failed. Participants only proceeded to the main experiment if they passed at least block +/-2. If a participant passed block +/-2, but not +/-1, then their experimental session was set at +/-2 probe distances. However, if block +/-1 was also passed then their experimental session was set at +/-1. In the experimental session, if participants' catch trial errors exceeded the group mean error rate by + 1SD they were excluded from analysis. However, because perfect catch trial performance could be reached by pressing the space-bar on every trial, participants were also excluded if their detection of displacements in catch trials did not show at least a minimum improvement of 10% compared to experimental trials. These exclusion criteria were deliberately conservative to focus particularly on those participants that engaged with all aspects of the task. Moreover, the prior calibration session effectively ensured that participants were at detection threshold, and it was important to ensure that those were excluded for whom this threshold shifted or was measured incorrectly. The same exclusion criteria were applied to all subsequent experiments.

Results

Exclusions. Based on the inclusion criteria, nine participants were excluded after the experimental session. Of the remaining 31 participants, trials where responses were either faster than 200ms (anticipations) or slower than 3000ms were also excluded from the analysis (0.3% of trials).

Main analyses. Participants detected 94% (*SD* = 11%) of displacements in the catch trials, and 55% (SD = 16%) in the experimental trials. Responses were entered into a repeated measures ANOVA with Intention ("Forward" vs. "Backward"), Action Direction (reach vs. withdrawal) and Probe Direction ('forward' vs. 'backward') as factors. The 'same' probes were not analysed. There was a significant effect of Probe Direction, $F(1,30) = 99.15$, $p < .001$, $\eta_p^2 = .768$, 95% CI [23, 35]. As expected, backward displacements were detected more often than forward displacements, replicating the representational momentum effect. Crucially there was a three-way interaction between Intention, Action Direction and Probe Direction, $F(1,30) = 4.89$, $p = .035$, $\eta_p^2 = .140$, 95% CI [1, 26], showing that representational momentum increased when action and intention matched (see Figure 14). There were no further main effects or interactions (all $F < 2.29$, all $p > .14$).

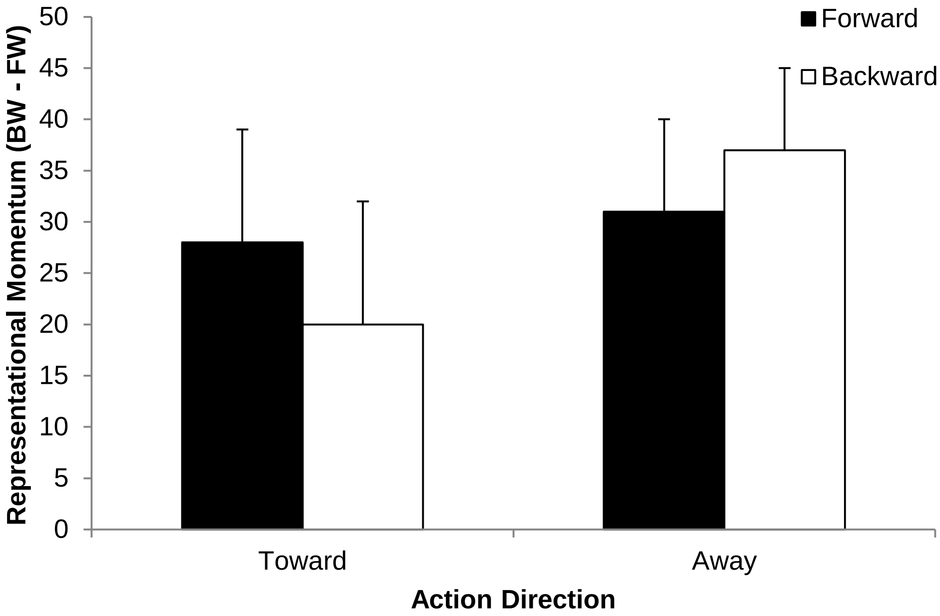


Figure 14. The interaction between prior expectation and action direction on the size of the representational momentum effect ('backward' probe detections – 'forward' probe detections) in Experiment 6. Participants said "Forward" if the object was safe and "Backward" if the object was dangerous. Error bars represent 95% confidence intervals.

Discussion

The results from Experiment 6 revealed a robust representational momentum effect. Participants more readily detected displacements that went backward, against the trajectory of motion, compared to those that went forward. This supports the idea that humans process the actions of others by generating predictions of how the action will continue, resulting in the modulation of perception forward in line with these expectations. It is also in line with the representational momentum effect found in the experiments in

Chapter 3, but here with the new stimuli and changes in methodology the effect is far more robust. It also supports previous work demonstrating the representational momentum effect when observing other people's actions (Wilson et al., 2010; Hudson et al., 2009; Thornton & Hayes, 2005).

Importantly, the current results demonstrate that this change in perception can also be affected by prior expectations created by the participant. Expectations generated by the specific verbal instructions about the kinematics of the action enhanced representational momentum. Here, if the verbal instruction of the participant matched the subsequent observed movement, representational momentum increased compared to when the verbal instruction mismatched the movement. These findings suggest that the verbal cue uttered by the participant altered their perceptual judgement of the subsequent action. Moreover, the verbal statements ("Forward", "Backward") corresponded to the direction of the reach from the actor's viewpoint, and not the viewpoint of the participant, which would have been 'left' or 'right'. Therefore, the alteration of the participant's perception as a result of "Forward" and "Backward" must have resulted from the participant taking the perspective of the actor when making the verbal statements. While prior evidence has shown evidence for representational momentum using social stimuli, this is the first study to show that social representational momentum can be affected by prior expectations.

Experiment 7 – Spoken Verbal cue – "Take it" or "Leave it"

While the results of Experiment 6 were encouraging, the failure to replicate results within Chapter 3 highlighted the need for replication. Experiment 6 showed that saying words that denoted a potential movement altered the subsequent perceptual judgement of a hand's position in relation to the target object. This indicated that lower-level verbal cues related on the kinematic level to the action's intention can alter predictive processing. The question now was whether the same would happen when the verbal utterances were of a higher level of intention.

Experiment 7 was identical to Experiment 6 apart from the verbal utterances required. This time, when a safe object was present, participants were required to say "Take it" and when a painful object was present they were required to say "Leave it". This allowed the replication of the previous experiment, and testing whether the effect also applied to intentions at a higher-level.

Method

Participants. 42 participants took part in Experiment 7. Participants were recruited and rewarded in the same manner as Experiment 6, adhering to the same ethical guidelines and participant requirements. Ten participants failed the training session, leaving Experiment 7 with 32 participants (22 females, mean age = 23.3 years, SD = 6.9) to take part in the experimental session.

Apparatus and Stimuli. These were identical to Experiment 6.

Procedure. Experiment 7 was identical to Experiment 6 with the sole exception of the difference regarding the verbal response required. This time participants were asked to say "Take it" in response to safe objects and "Leave it" in response to painful objects.

Inclusion criteria. These were the same as Experiment 6.

Results

Exclusions. In line with the inclusion criteria, eight participants were excluded after the experimental session. 1.4% of trials from the remaining 24 participants' responses were excluded based on reaction time measures.

Main analyses. The mean reported displacements were 99% (SD = 2%) in the catch trials and 61% (SD = 11%) in the experimental trials. Responses were entered into a repeated measures ANOVA with Intention ("Take it" vs. "Leave it"), Action Direction (reach vs. withdrawal) and Probe Direction ('forward' vs. 'backward') as factors. The 'same' probes were not analysed. There was a significant effect of Probe Direction, $F(1,23) = 17.60$, $p < .001$, $\eta_p^2 = .433$, 95% CI [11, 32]. As expected, backward displacements were detected more often than forward displacements, replicating the representational momentum effect. There was also a significant effect of Action Direction, $F(1,23) = 18.51$, $p < .001$, $\eta_p^2 = .446$, 95% CI [7, 19], with displacements for withdrawal more readily detected than displacements for reaches. Crucially there was a three-way interaction between Intention, Action Direction and Probe Direction, $F(1,23) = 4.67$, $p = .041$, $\eta_p^2 = .169$, 95% CI [1, 26], showing that representational momentum increased when action and intention matched (see Figure 15). There were no further main effects or interactions (all $F < 0.47$, all $p > 0.5$).

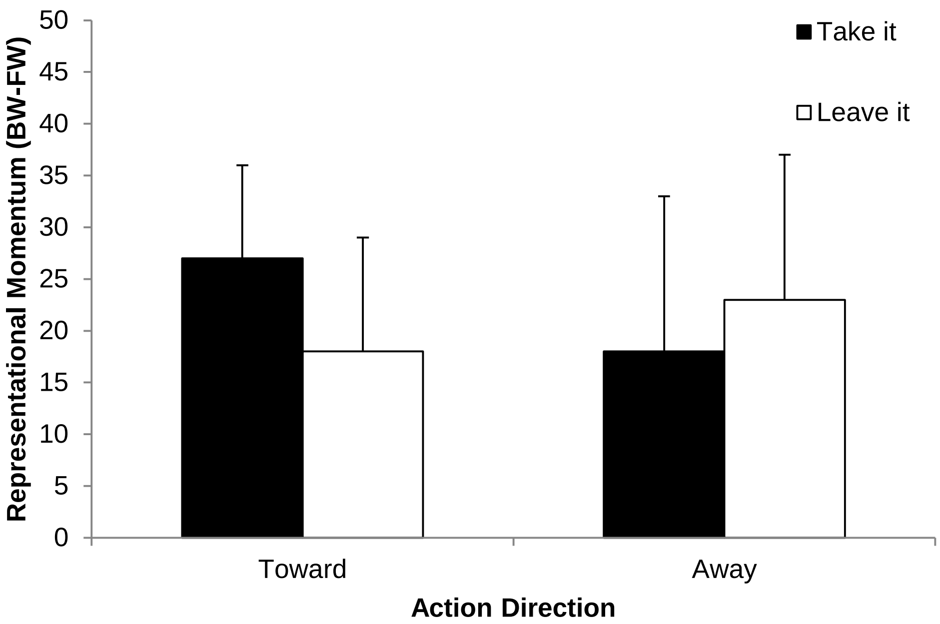


Figure 15. The interaction between prior expectation and action direction on the size of the representational momentum effect ('backward' probe detections – 'forward' probe detections) in Experiment 7. Participants said "Take it" if the

object was safe and "Leave it" if the object was dangerous. Error bars represent 95% confidence intervals.

Discussion

The results of Experiment 7 fully replicate those of Experiment 6. Once again a significant representational momentum effect revealed participants' tendency to more easily detect 'backward' probes than 'forward' probes, further confirming the robustness of the representational momentum effect.

Crucially, again the verbal utterances affected participants' subsequent perception of a hand's position, but these results show that this effect extends to expectations about the action's goal in addition to movement kinematics. As in the previous experiment, when prior expectations created by the verbal statements were met, representational momentum was larger, suggesting more forward prediction. In contrast, when prior expectations were not met, representational momentum decreased, suggesting a decline in forward prediction. This demonstrates that prior expectations generated on a range of different levels alters people's subsequent perception. The amount of predicted forward motion seems to be a combination of expectations from prior motion (bottom-up input from the hand's motion) and expectations generated from the verbal statements. The largest predictive effects emerged, for both reaches and withdrawals, when both were aligned.

However, one issue within Experiments 6 and 7 was that the verbal response was dependent on the type of object (safe/painful), making the two always confounded. This means that it is difficult to know whether the effects emerge from the intentional cues or on object information. In order to address this, another experiment was run which eliminated the role of object knowledge.

Experiment 8 – Spoken Verbal cue – Colour control

The findings from Experiments 6 and 7 have shown that expectations generated by verbal utterances in response to different types of objects altered the perception of subsequent actions. However, in these experiments, the verbal utterance for each object category was constant, making the two confounded, and making it hard to know how far the effect relied on expectations derived from the verbal statements or from object information. For instance, a decrease in representational momentum when observing reaches towards a painful object could be a result of the object's painfulness or because the verbal utterance denoted avoidance. Experiment 8 was run to dissociate the two factors from one another. The experiment was the same as Experiment 6 except that the colour of the object was randomly manipulated across object types. If the object was green, participants said "Forward" to initiate the action, and if it was red they said "Backward". This meant that verbal responses were now independent of object type allowing the influence of movement expectancies alone, regardless of object type, to be investigated.

Method

All aspects of the experiment were identical to Experiment 6 apart from the following.

Participants. 36 participants (28 females, mean age = 21.6 years, SD = 6.0) took part in Experiment 8. All participants passed the training session.

Stimuli. The existing stimuli were modified to produce two sets of stimuli, one where the object was overlaid with a green (R: 55, G: 225, B: 1) filter and one where the overlay was red (R: 255, G: 14, B: 3). The filter was set to an opacity of 30%, making the object type still clearly visible.

Procedure. The training session was identical to the previous 2 experiments apart from training block +/-4 not being administered. Throughout the experiment, object colour was chosen randomly. The frequency of object colour across the factors action direction, object type and probe was not significantly different from chance ($X2$ = 3.73, df = 11, p = .98). If participants saw a green object they responded "Forward" to initiate the action sequence, while if the object was red, they responded "Backward" to start the action.

Results

Exclusions. Five participants were excluded based on their performance on the experimental session. 1.1% of trials of the remaining 31 participants' responses was excluded based on reaction time cut-offs.

Main analyses. Participants detected 93% (SD = 13%) of displacements in the catch trials and 62% (SD = 15%) in the experimental trials. Responses were entered into a repeated measures ANOVA with Intention ("Forward" vs. "Backward"), Action Direction (reach vs. withdrawal), Object (safe vs. painful) and Probe Direction ('forward' vs. 'backward') as factors. The 'same' probes were not analysed. There was a significant effect of Probe Direction, $F(1,30)$ = 11.49, $p < .001$, $\eta_p^2 = .558$, 95% CI [21, 40]. As expected, backward displacements were detected more often than forward displacements, replicating the representational momentum effect. There was also a significant effect of Action Direction, $F(1,30) = 12.02$, $p = .002$, $\eta_p^2 = .286$, 95% CI [3, 11], with displacements for withdrawals more readily detected than displacements for reaches. There was also a significant effect of Object, $F(1,30) = 5.72$, $p = .023$, $\eta_p^2 = .160$, 95% CI [1, 6], with displacements for safe objects more readily detected than displacements for painful objects. Crucially there was a three-way interaction between Intention, Action Direction and Probe Direction, $F(1,30) = 4.31$, $p = .047$, $\eta_p^2 = .126$, 95% CI [1, 22], showing that representational momentum increased when action and intention matched (see Figure 16). There was also a three-way interaction between Intention, Object and Probe Direction, $F(1,30) = 10.88$, $p = .003$, $\eta_p^2 = .266$, 95% CI [6, 23]. This showed that when observing actions involving painful objects, saying "Forward" increased representational momentum (irrespective of whether the action was a reach or a withdrawal) compared to saying "Backward", but the reverse was true when observing actions involving safe objects, representational momentum being greater when saying "Backward" compared to saying "Forward". Incongruence between verbal statement and object therefore generally increased representational momentum, irrespective of action direction. There were no further main effects or interactions (all $F < 3.56$, all $p > 0.068$).

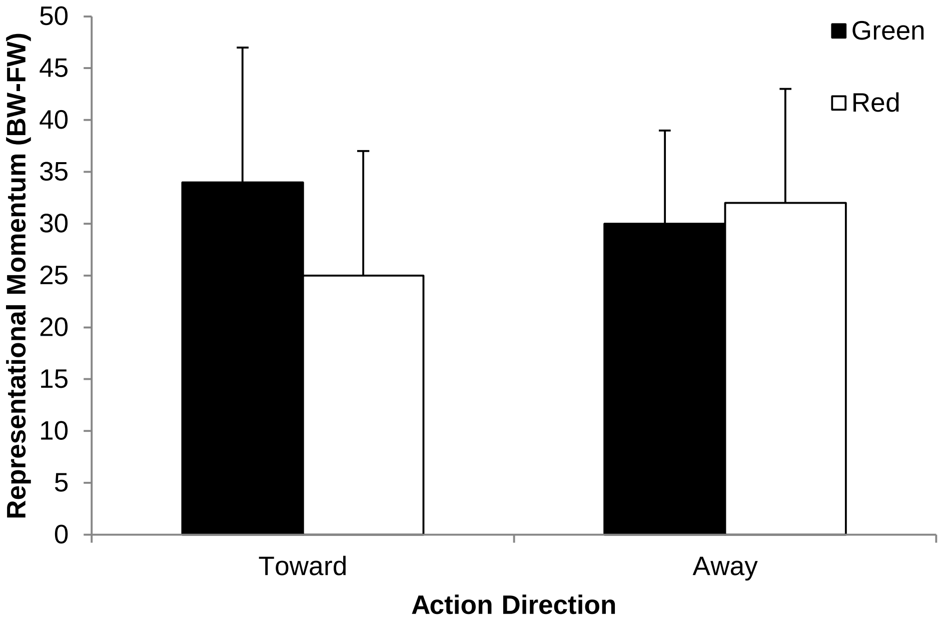


Figure 16. The interaction between prior expectation and action direction on the size of the representational momentum effect ('backward' probe detections – 'forward' probe detections) in Experiment 8. The colour of the object was randomly assigned as green (black bars) or red (white bars) independent of object type and participants said "Forward" if the object was green and "Backward" if the object was red. Error bars represent 95% confidence intervals.

Discussion

The results from Experiment 8 further replicate the representational momentum effect and confirm that the new stimuli and changes in methodology produce a more consistent effect. More importantly, the findings of the current experiment reveal that the effect of expectancies generated by the verbal statements was still present when they were derived from abstract colour cues, independent of object type. Results showed that verbal utterances in response to the colour of the object, regardless of the object type, generated expectations that altered the level of representational momentum in the same way as the previous two experiments. This demonstrates that the effects of the previous two experiments rely on the verbal statements uttered, and not the painfulness of the object.

In contrast, object type had a more general effect on representational momentum, being generally larger whenever verbal statement (forward, backward) and object type (painful, safe) mismatched. This is in line with the view that the representational momentum effect is highly automatic, and that mismatches occupy cognitive resources that cannot be used for accurate (not biased) detection of the displacements.

Taken together, all these findings provide evidence that perception not only relies on bottom-up sensory information, but also involves the influence of prior knowledge. The current results demonstrate that verbal utterances can modify expectations and alter subsequent social perception. When expectations were satisfied representational momentum was larger, suggesting greater perceptual prediction. In contrast, when expectations were not satisfied representational momentum was reduced, suggesting weaker perceptual prediction.

Here though such expectations were always forced on the participant through the requirement to verbally respond to the stimuli. One question about such expectations then is whether they require some form of generation or whether they can also be generated automatically. For example, would the effect of expectation on representational momentum still be visible if the verbal statements were not spoken but only heard, as if from the observed?

Experiment 9 & 10 – Auditory cue

The previous experiments demonstrated that expectancies generated by the spoken statement of a participant could modulate their perception of a subsequently observed action. However, what is less clear is whether these effects rely on the participants producing the verbal statement themselves, or whether the words themselves can generate such perceptual expectancies, for instance when heard by the participant but produced by the observed actor. That is, are expectations created because the verbal statements were self-generated or did they emerge from the linguistic content itself?

In the following two experiments the same stimuli and representational momentum task were implemented but this time rather than performing a verbal response, participants heard verbal statements as if coming from the owner of the observed hand. Participants heard the verbal statements from Experiment 7, “I’ll take it” and “I’ll leave it”. This tested whether the effect of intentionality on representational momentum was present in the absence of any verbal response by the participant. In Experiment 9 the audio stimuli were object specific, so participants always heard “I’ll take it” when a safe object was present and “I’ll leave it” when a painful object was present. This allowed investigation of whether the verbal statements themselves were sufficient to generate expectations and modulate representational momentum, or whether the modulation required a self-generated expectation. In Experiment 10, the verbal statements and object type were not linked, making the audio stimuli unrelated to the object present. This meant that the verbal statements heard were now independent of object type, allowing the influence of movement expectancies alone, regardless of object type, to be investigated when the verbal statements were merely heard.

Method

All aspects of the experiment were identical to previous experiments (4, 5, 6) in methodology and application apart from the following.

Participants. 35 (23 females, mean age = 27.7, SD = 10.3) participants took part in Experiment 9 and 32 (20 females, mean age = 26.8, SD = 9.2) in Experiment 10.

Apparatus and Stimuli. Audio stimuli were recorded using an M-Audio Microtrack 2 Digital Voice Recorder. Two audio stimuli of an actor saying "I'll take it" and "I'll leave it" were created, each of 1000ms duration and played through the headphones at 75% volume. The audio stimuli were biased to the right earphone by 50% to match the position of the actor on the right of the screen. A Logitech PC120 combined microphone and headphone set was used to deliver audio stimuli. All other stimuli and apparatus were the same as previous experiments.

Procedure. The procedure mirrored previous experiments but this time no training session was included. Instead, two probe levels were included in the main experiment. These two probe levels were designed to capture most of the variability in participants' detection threshold to detect the displacements. They should therefore reduce loss of participants in the calibration/practice session prior to the experiment, and should protect the measurement from shifts in a participant's detection threshold over the experiment. The experimental session was identical to the previous experiments but this time when the first frame of the action sequence was presented following a random variable delay between 1000 and 3000ms the audio stimulus – "I'll take it" for safe objects and "I'll leave it" for painful objects – was presented. The action sequence began 200ms after the end of the audio stimulus. The variable SOA between trial onset and auditory stimulus and the highly predictable start of the action sequence after auditory offset again created a causal link between statement of intention and action initiation. The rest of the stimulus presentation followed the previous experiments, with the following exceptions.

While in previous experiments one probe level was used throughout, here two 'different' probe levels (+/-1 and +/-2) were implemented resulting in five possible probe positions (-2, -1, 0, +1, +2). In Experiment 9 there were 160 experimental trials, which were made up of 8 iterations of the Intention ("I'll take it", "I'll leave it") x Action direction (toward or away) x Probe direction (-2, -1, 0, +1, +2). For each trial the object was randomly selected. To ensure that participants processed the audio statements, an extra 16 trials were added which included the actor stating the wrong intention ("I'll take it" for painful objects, "I'll leave it" for safe objects). As soon as participants detected the wrong intention they were required to say "STOP" into the microphone, which ended the trial. In addition to this another 16 catch trials where the probe was +/-4 were also added. In Experiment 10 there were 240 trials which were made up of 6 iterations of the factors Object Type (safe, painful), Intention ("I'll take it", "I'll leave it"), Action Direction (reach, withdrawal), and Probe (-2, -1, 0, +1, +2), and 16 catch trials with +/-4 probes.

Inclusion criteria. As no training blocks were administered, participants were excluded based on their performance in the experimental session. This mirrored the previous three experiments, with the only difference being that the two probe levels were collapsed together when comparing performance in the experimental trials with performance in the catch trials.

Results - Experiment 9

Exclusions. Eleven participants were excluded based on the exclusion criteria.

Main analyses. Participants detected 82% (SD = 15%) of displacements in the catch trials and 58% (SD = 8%) in the experimental trials. Responses were entered into a repeated measures ANOVA with Intention ("I'll take it" vs. "I'll leave it"), Action Direction (reach vs. withdrawal) and Probe Direction ('forward' vs. 'backward') as factors. The 'same' probes were not analysed. There was a significant effect of Probe Direction, $F(1,23) = 68.6$, $p < .001$, $\eta_p^2 = .749$, 95% CI [20, 32], as expected backward displacements were detected more often than forward displacements, replicating the representational momentum effect. Crucially there was a three-way interaction between Intention, Action Direction and Probe Direction, $F(1,23) = 14.9$, $p = .001$, $\eta_p^2 = .395$, 95% CI [10, 30], showing that representational momentum increased when action and intention matched (see Figure 17). There were no further main effects or interactions, all $F < 1.89$, all $p > .182$.

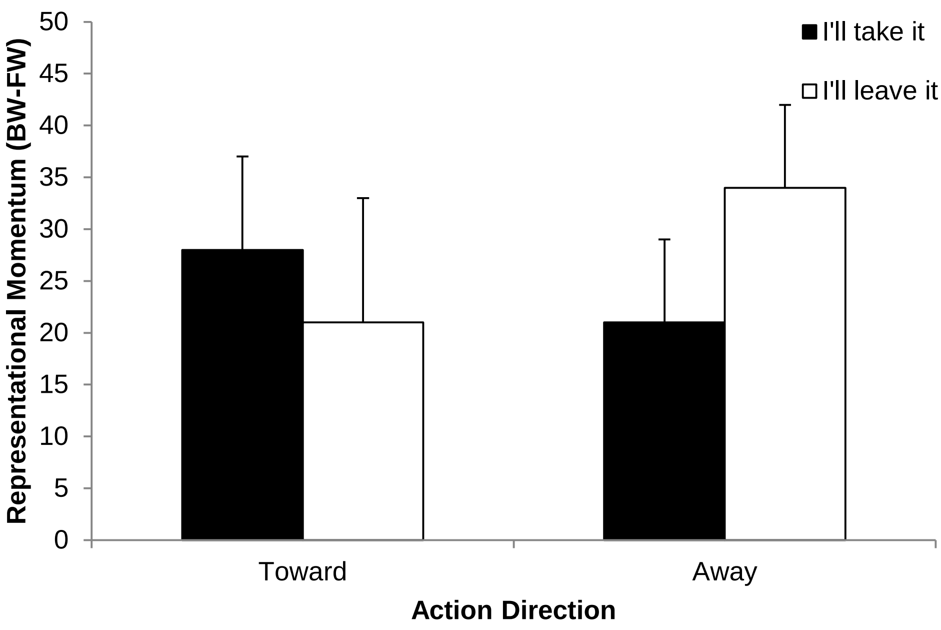


Figure 17. The interaction between prior expectation and action direction on the size of the representational momentum effect ('backward' probe detections – 'forward' probe detections) in Experiment 9. Participants heard "I'll take it" if the object was safe and "I'll leave it"if the object was dangerous. Error bars represent 95% confidence intervals.

Results - Experiment 10

Exclusions. Seven participants were excluded based on the exclusion criteria.

Main analyses. In the catch trials 82% (SD = 15%) of displacements were detected, while in the experimental trials 55% (SD = 17%) were detected. The proportion of displacements were submitted to a repeated measures ANOVA with Intention "I'll take it", "I'll leave it"), Action Direction (reach, withdrawal), Object Type (safe, painful) and Probe Direction ('forward' vs. 'backward') as factors. There was a significant effect of Probe Direction, $F(1,24) = 53.5$, $p < .001$, $\eta_p^2 = .690$, 95% CI [20, 35], once again replicating the representational momentum effect. Most importantly though, there was once again a three-way interaction between Intention, Action Direction and Probe Direction, $F(1,24) = 7.61$, $p = .011$, $\eta_p^2 = .241$, 95% CI [4, 21] (see Figure 18). Like Experiment 9, the representational momentum effect was larger when action and intention matched, compared to when they did not match. There were no further main effects or interactions, all $F < 2$.

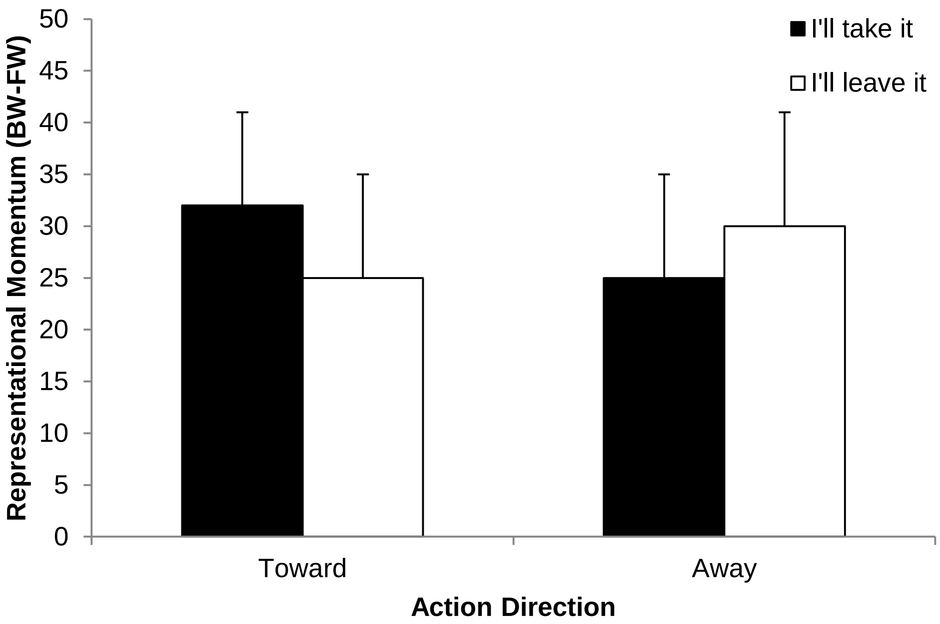


Figure 18. *The interaction between prior expectation and action direction on the size of the representational momentum effect ('backward' probe detections – 'forward' probe detections) in Experiment 10. Participants heard "Take it" or "Leave it" independently of whether the object was safe or painful to grasp. Error bars represent 95% confidence intervals.*

Discussion

The two auditory experiments replicate the robust representational momentum effect found in the first three experiments, with once again 'backward' probes more likely to elicit a response than 'forward' probes. More interestingly, the present two experiments provide evidence that forward bias can be affected by the prior stated intentions of the observed, even when this is task-irrelevant in Experiment 10. Participants were more likely to misperceive the observed action as further along its trajectory when they heard a verbal statement that matched the intention of the subsequently observed reach. Building on the finding that verbally produced statements can modulate the perception of another's action, these results suggest that the modulation of perception can be affected automatically. Hearing verbal intentions in line with the subsequent observed movement increased representational momentum even when there was no overt requirement to process the audio stimuli. This demonstrates the automaticity of these predictions and their sensitivity to intentional language. It shows moreover that predictions relating to observed actions are informed not only by self-generated expectations (Experiments 6-8) but also by expectations generated by the verbal stated intentions of the observed (Experiments 9 and10). Taken together these findings emphasise that social perception can modulate the perception of biological movements in line with higher-level inputs based upon intentions.

General Discussion

The findings of the five studies presented here have demonstrated a robust representational momentum effect and can be seen as strong evidence that the perception of biological motion, like object motion, is subject to a perceptual bias in line with the direction of motion. This supports the idea that we predict the movement of other people's bodies in line with prior knowledge. More importantly, the current findings provide the first evidence that such predictions are sensitive to prior expectations of intention generated by language. In the first three experiments verbal statements performed by the participant prior to any observed arm movement altered their subsequent judgments of where the movement ended. When observed movements were in line with the intention of the verbal statement representational momentum increased, compared to when they contradicted the verbalised intention. This shows that expectations generated by the participant affected their subsequent perception. This is evidence not only that we predict, but that social cues, such as language, can directly affect these predictions. Such an assertion is further supported by the finding in later experiments that the same effect is present when verbal statements were only heard, as if from the observed actor, demonstrating that intentional cues affect prediction automatically. These results support the idea that language production and comprehension rely on an integrated system, which is directly linked to action prediction (Wolpert, Doya & Kawato, 2003; Springer, Huttenlocher & Prinz, 2012;).

The current findings also support recent models of prediction, in which higher-level knowledge interacts with, and guides, perceptual experience (Grush, 2004; Knill & Pouget, 2004; Friston, Daunizeau, Kilner, & Kiebel, 2010). They extend these ideas to include processing and understanding the actions of others. Here, the expected intentions of the observed actor facilitated the generation of predictions about how the action might continue, which are incorporated with incoming sensory information. When the expected intention and observed movement are in line, perceived displacements that contradict expectations ('backward' probes) result in a prediction error, while displacements in line with expectations ('forward' probes) remain undetected. When intention and movement contradict one another, detection of displacements is more equatable, resulting in reduced representational momentum. While previous research studying prediction errors in visual perception have concentrated on low-level aspects such as local movement (Roach et al., 2011) or the probability of presentation (Summerfield et al., 2006), these findings show that similar perceptual effects can result from higher-level expectations, and emphasise the importance of prior expectations during social perception. This supports recent theories of social cognition that emphasise the importance of top-down expectancies when understanding the actions of others (Wilson & Knoblich, 2005; Bach et al., 2014).

The studies in this chapter demonstrate that bottom-up perception can be altered through top-down prior expectations established by verbal cues. Moreover, the results of Experiment 8 showed that the verbal cues generated these top-down expectancies themselves and did not require a connection to the object. However, current theories of object processing suggest that they are important drivers of action predictions (van Elk, van Schie & Bekkering, 2014; Bach et al., 2014). Indeed, a number of studies have provided evidence that objects automatically potentiate motor programs related to the objects' affordance, facilitating perception (Bub et al., 2008; Ellis & Tucker, 2000; Riddoch, Edwards, Humphreys, West, & Heafield, 1998), and that such affordances are also derived for objects near other people (for a review, Creem-Regehr, Gagnon, Geuss, & Stefanucci, 2013).

In the present studies, while objects were always present, the actions related to them were always preceded by a verbal cue. Moreover, the object type was only independent of the action direction in two of the experiments (Experiments 8 and 10). While the object type interacted with the verbal cue and probe direction in Experiment 8 this was independent of the action direction, making any interpretation in relation to top-down predictions meaningless. Moreover, there were no interactions involving object type in Experiment 10. Therefore, it is impossible to know from these studies whether objects do not affect predictive processing or whether the verbal cues in the current experimental set-up were just stronger intentional cues. Indeed, this latter explanation is supported by evidence, which suggests that short-term action intentions can override long-term semantic knowledge (van Elk et al., 2008).

The findings from Chapter 3 failed to show consistently that objects and the actions they afford alter predictive processing, but the cue of grip type may have been too subtle. A more effective way to investigate the role of objects in prediction using the representational momentum paradigm would be to compare how object presence affects expectations of intention. In the experiments outlined above, the verbal cue defined the intention, whereas by comparing representational momentum when observing reaches to objects against reaches to empty space, the object itself defines the intention. Therefore, one would expect that reaches to an object would be viewed as including a clear intention, whereas a reach to empty space does not. If objects are important to predictions, representational momentum should be greater when reaches are object-directed compared to when they are not.

In conclusion, while the present studies show that prior expectations established through language can increase predictive processing of an actor's arm, they leave open the role of objects within this process. To test this, the next chapter set out to investigate how object presence affected predictive processing by directly comparing representational momentum when observing object-directed reaches compared to reaches to empty space.

Chapter 5 - The effect of object cues on visual and tactile perception

The experiments in Chapter 4 have confirmed the tendency for participants to predict the actions of others further into the future, and uncovered that this tendency is enhanced when they match prior expectations of intention. They have therefore provided the first evidence that cues of intention can modulate predictive processing. Interestingly, these verbal cues were shown to be independent of the type of object present (Experiments 8 and 10), suggesting that objects are not an important cue to intention. However, this is at odds with prior research, which suggests that seeing an object automatically elicits motor knowledge thought to result from its affordances, as well as the action goals it helps to achieve, and that this also occurs when observing other people near objects (for a review, see Bach et al., 2014; Tucker & Ellis, 1998, 2001, 2004; Vainio, Tucker & Ellis, 2007; Stoffregen, Gorday, Sheng, & Flynn, 1999; Costantini et al, 2011; Cardellicchio et al, 2013)

If objects do activate object affordances when observing somebody else interacting with them then they should act as a cue to the action and generate a top-down prediction of what might happen next. In previous experiments objects were always present but varied with regard to their painfulness (Chapter 4). Based on a previous study, it was expected that reaches to painful objects would not elicit the same affordances for reaches toward them (Anelli et al., 2012). Yet, only the intentional statements, not object type, were found to modulate predictions of further movement. One reason may have been because while the objects were processed semantically for the task, the subsequent action intention induced by the verbal cue overruled any semantic effect (van Elk et al., 2008).

One way to investigate the role of objects as an intentional cue is to measure the effect of object *presence* on action perception. A new experiment was designed to compare visual perceptual prediction when an object is present with when no object is present (pantomime action). Participants were presented with reaches to objects and the same reaches into empty space, based on the assumption that an object-directed reach would be viewed as having a clear intention, whereas a reach to empty space would be seen as having no clear intention. Thus, if objects provide such cues to intention during action observation, object presence should increase perceptual predictions, and thereby representational momentum. In contrast, when the object is absent any perceptual prediction is likely to be reduced due to the lack of clear intention, and result in a decrease in representational momentum

The focus on object presence allowed a secondary question to be investigated as well. According to prediction theories (Friston, 2010; Friston & Stephan, 2007; Panichello et al., 2013), and seen in the last chapter, top-down expectations affect perceptual judgements. These top-down predictions are informed by intentional cues, such as the verbal instructions of the participants or the utterances of the observed actors. However, at least some of these cues – the objects from which the intentions were derived – were present from before action onset till just before appearance of the probes. Yet, if the observed forward predictions truly stem from top-down expectations, then they should be observed both when the relevant cue is currently present during the perception of the action and when it is absent, and only known about. With regard to the present experiment, this would suggest that object presence should affect forward prediction of movement, both when the goal object of the action is present in the scene, and when it is only *known* to be present.

A study investigating mirror neurons in the macaque monkey by Umilta and colleagues (2001) shed some light on this issue. They demonstrated activation of 'mirror neurons' when the monkey observed grasps of an object, but not when observing pantomimed grasps. Interestingly, when object and destination of the pantomimed grasp were occluded prior to the reach, mirror neuron responses showed the same pattern, firing for real but not pantomimed grasps, even though the bottom-up visual input in both conditions was now identical. This suggests that the firing pattern of the mirror neurons does indeed reflect primarily top-down responses, that reflect higher-level understanding of the event, and which fill in the action information that could not be derived from direct visual experience (Csibra, 2007). However, as this is a monkey single cell study caution should be applied to any interpretation. While mirror neuron activity has previously been seen as equivocal with action understanding, there is now much discussion about what mirror neuron firing represents (Csibra, 2007; Heyes, 2010). However, in humans it has indeed been shown that observation of an action towards an occluded object, compared to a pantomime action, activates part of the somatosensory cortex implying some form of prediction of the sensory consequences of the observed action in the absence of any direct vision of it (Turella et al, 2011).

Therefore, a modified version of the experimental paradigm used in the Umilta study was run with the representational momentum methodology. Stimuli consisted of reaches towards an object or the same reach towards an empty space. In addition, half the time, prior to the beginning of the reach, the object or empty space could be hidden by an occluder. This meant that an observer would only know whether an object was present or absent in the occluded trials, but would not have direct bottom-up access to that information when observing the movement. This allowed firstly a more direct test of whether objects can serve as an intentional cue in their own right, and secondly, whether knowledge of an intentional cue in the absence of direct bottom-up sensory information can modulate top-down predictive processing.

Experiment 11 – Object as cue for intentionality

Object presence itself was used to manipulate intentionality, based on the assumption that predictive processing should be evident when observing an intentional reach (object-directed) compared to a non-intentional reach (directed to empty space). Participants observed an actor reach across a table either towards an object or towards empty space. If reaches towards objects are perceived as more intentional, they should elicit more representational momentum than reaches to empty space, demonstrating the role of the object as an intentional cue. In addition, observed reaches were either fully visible or their target was hidden behind an occluder. If prior knowledge of the object's presence is sufficient to initiate predictive processing then representational momentum should be larger when an object is present in the occluded condition compared to when observers only *know* of its presence. If however, predictive processing relies on constant bottom-up sensory information then representational momentum should not be affected by object presence in the occluded conditions.

A concern from the opening studies was the processing of the object itself. As the verbal cue, which was used to achieve this in the previous chapter, could not be used here, an alternative solution was required to ensure object presence was processed. Therefore a catch trial question was added after performing the representational momentum task (in 20% of the trials) that asked the participant whether, in the just seen action, the object was present or absent. This forced participants to remember whether an object was hidden behind the occluder; they could not simply ignore it.

Method

Participants. 29 participants took part in Experiment 11. Participants were made up of students from Plymouth University or members of the public from the wider Plymouth community and they took part in exchange for participation points or payment (£8 p/h). All had normal or corrected-to-normal vision, were native English speakers and right handed. All provided written informed consent before participating and were debriefed at the end of the experiment. The study was approved by the University of Plymouth's ethics committee. One participant failed the training session leaving Experiment 11 with 28 participants (4 males, mean age = 21.1 years, SD = 5.9)

Stimuli and apparatus. Stimuli were filmed with a Casio Exilim EX-ZR100 at a high-speed frame rate of 240 frames per second and edited using Moviedek and Corel PaintShop Photo Pro x6. They consisted of a side on view of a man sat at a table reaching with his right hand towards one of four different objects (orange, stapler, book, coke bottle) positioned in the same location (see Figure 19). In addition, a video was filmed where an occluder (two black lever arch box files attached together) was pushed into the scene from the opposite side of the table to hide the space where the four objects were located while the man remained static. Every 12th frame (representing a 50ms step in real time) was extracted from each video, and 14 images were chosen for each action that spanned the trajectory, from the start of the reach to approximately half way through its trajectory towards the target location.

These images were then duplicated and in one set the object was digitally removed, resulting in two sets of 14 images where the same reaches either directed to an object or empty space. The first image of the each sequence, where the hand was rested still on the table, was used as the basis for the occluder movies. Upon it the occluder was pasted at different stages of its trajectory across the table until it obscured the whole of the object, creating 13 new images for each action sequence. Next, each of 14 images for each action sequence had the occluder inserted before the start of the sequence. This resulted in 27 images for each action sequence, 13 where the actor was still and the occluder emerged into the picture and 14 showing the reach towards the now hidden target location. This resulted in 4 sets of images, one for each condition. 2 sets of 14 images (spanning 700ms) for the two visible conditions (object/no object), where the target location (object or empty space) was visible, and 2 sets of 27 images (spanning 1350ms, 650ms of occlusion) where the object or empty space was hidden prior to the onset of the reach (see Figure 19, panel C). Four different sequence lengths were used ranging from 5-11 frames in length and proceeding in one frame steps.

As in the earlier experiments, the probe image could either be the same as the last image of the action sequence (same probe), or different. When the probe was different, again it could either be further forward along its trajectory in a future position (forward probe) or back along the trajectory in a past position (backward probe). As in the experiments from Chapter 4, different probes were chosen as the future or past frames in the sequence rather than modifying the last image. Different probes varied between 2 or 3 (depending upon calibration, see below) frames in distance from the last image of the action sequence in experimental trials. The experiment was administered using a 22-inch Philips Brilliance 221P3LPY monitor (resolution: 1920 X 1080, refresh rate: 60 Hz) using Presentation software (Neurobehavioural Systems, Inc.).

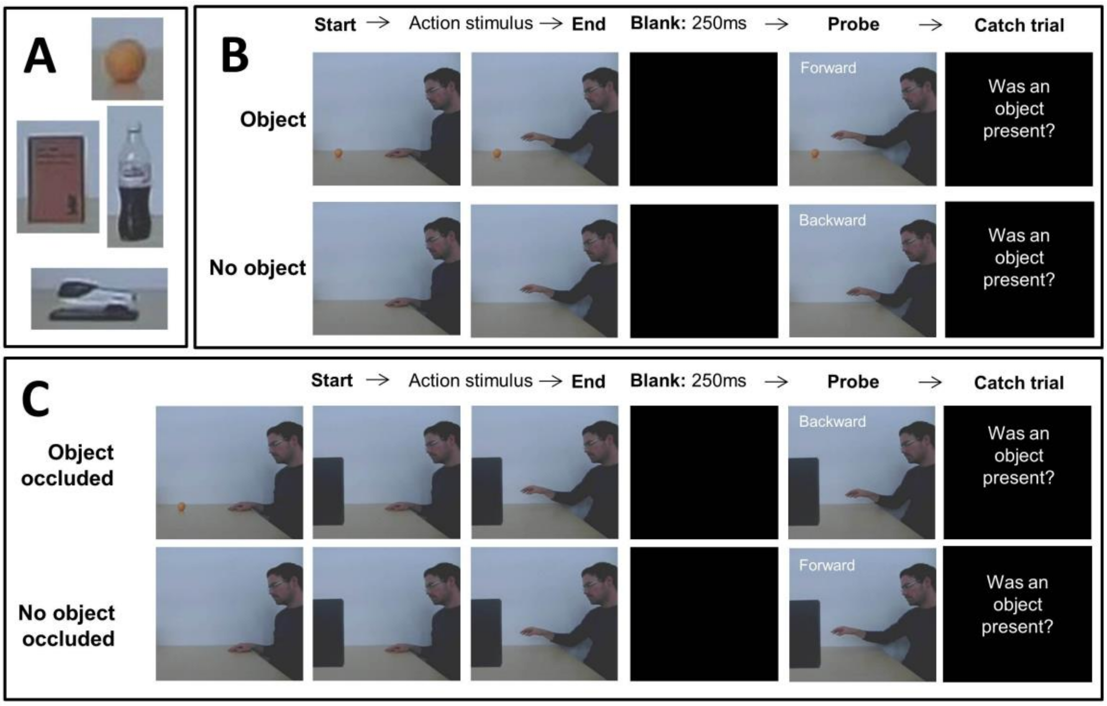


Figure 19. A: The four target objects (orange, book, coke, stapler) used in experiments 11-13. B: Trial sequence for the non-occluded conditions of a reach towards an object (upper row) and empty space (lower row), followed by a blank screen, the probe stimulus and then the catch trial question. In this example, the upper row shows a forward probe (+2) and the lower shows a backward probe (-2).C: Trial sequence for the occluded conditions of a reach towards an occluded object (upper row) and occluded empty space (lower row), followed by a blank screen, the probe stimulus and then the catch trial question. In this example, the upper row shows a backward probe (-2) and the lower shows a forward probe (+2).

Design & procedure. Participants were seated approximately 60 cm from a colour monitor, and were given written and verbal instructions. The experiment began with a training session comprised of 3 training blocks (each

24 trials) that measured participants' ability to discriminate between the different types of probe. The first block contained different probes that were 4 frames from last image, and with each subsequent block they became closer to the same probe (+/-4, +/-3, +/-2). This meant that in the first block different probes were easier to judge, and with each block the task became harder.

Each trial began with the presentation of a fixation cross for 500 ms. Then, a blank screen was presented for 500 ms, before the action sequence was displayed with each image presented for 50ms. After the last image of the sequence was displayed, a blank screen was presented for 250ms, before a probe image was displayed (distance of different probes varying as specified above). All action sequences were presented along the vertical midline of the screen, but, as in the experiments of Chapter 4, the position along the X axis varied across trials and across the full length of the axis, thereby increasing variation and maintaining the participant's attention. Participants' task was to press the spacebar with their left hand if they thought the probe was in a different position to its last position, and not respond if they thought the probe was in the same position. Participants had 3,000 ms in which to make a response.

Participants who did not meet the inclusion criteria did not proceed to the main experiment (see below). The main experimental session was identical to the training session, apart from one important change, the requirement of a post judgement question. In order to direct participants' attention to the presence of an object, a catch trial question was added after the end of every trial. Participants were asked 'Was an object present?' and asked to press the Y key for yes, when an object was present, and the N key for no, when no object was present. Participants had 2,000 ms in which to make this response.

The main experimental session began with computer driven instructions, before a short training phase of 8 training trials that allowed participants to get used to the catch trial question. The experimental session consisted of 192 experimental trials with each stimulus combination shown once (4 objects, 4 conditions, 4 movie lengths, 3 probe types). The experiment was split into 3 blocks of 64 trials with breaks available in between blocks. The whole experiment lasted approximately 30 minutes.

Inclusion criteria. Like the experiments from Chapter 4, performance on the training session was rated on two measures, accuracy and sensitivity. Accuracy was merely the average of correct responses across all probe types. Sensitivity was calculated by subtracting the percentage of "same" responses when the probe was different from the percentage of "same" responses when the probe was the same. If participant performance within a block dropped below chance on either accuracy (50%) or sensitivity (0%) the block was failed. Participants only proceeded to the main experiment if they passed at least block +/-3. If a participant passed block +/-3, but not +/-2, then their experimental session was set at +/-3. However, if block +/-2 was also passed then their experimental session was set at +/-2. In the experimental session if a participant's catch trial accuracy was 1SD below the mean group accuracy they were excluded.

Results

Exclusions. In line with the inclusion criteria, two participants were excluded after the experimental session. This left 26 participants whose data was analysed.

Main analyses. Participants detected 72% (SD = 11%) of displacements in the experimental session. Responses were entered into a repeated measures ANOVA with Intention (intentional/object vs. non-intentional/no object), Visibility (visible vs. occluded) and Probe Direction (forward vs. backward) as factors. The "same" probes were not analysed. There was a significant effect of Probe Direction, $F(1,25) = 65.8$, $p < .001$, $\eta_p^2 = .725$, 95% CI [28, 46]. Backward displacements were detected more often than forward displacements, replicating the representational momentum effect. However, neither the main effect of Intention, $F(1, 25) = 2.29$, p = 0.142, $\eta_p^2 = .084$, 95% CI [-5, 10], nor the two-way interaction between Intention and Probe Direction, $F(1, 25) = 0.44$, $p = 0.514$, $\eta_p^2 = .017$, 95% CI [-3, 7], was significant (see Figure 20). There were no other significant effects, all *F's*< 0.682.

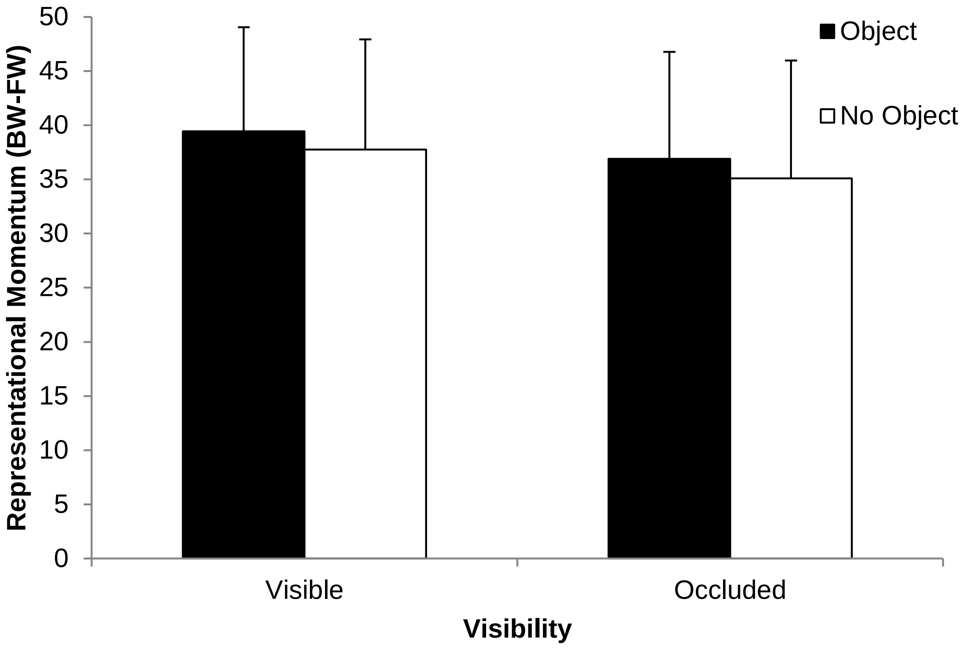


Figure 20. The size of the representational momentum effect ('backward' probe detections – 'forward' probe detections) when observing a reach toward an object (Intentional, black bars) or empty space (Non-Intentional, white bars), when the target location was either visible (left bars) or occluded (right bars) in Experiment 11. Error bars represent 95% confidence intervals.

Discussion

The results of Experiment 11 replicate the previous studies in demonstrating a robust representational momentum effect when observing the object-directed actions of others, but extend these findings to actions where the whole of the actor's body is visible. In addition, it extends them by showing biological representational momentum exists also when the actions are presented in real time. However, crucially no effect of intention was found. While intentional actions – those directed towards an object – again elicited more representational momentum numerically than non-intentional actions directed into empty space, and this held regardless of visibility, there was no statistical difference between them.

These findings suggest that object presence is not a strong enough intentional cue to generate expectations concerning the observed action and increase perceptual prediction. However, previous research has shown that objects can activate affordances even for the potential actions of others (Stoffregen et al, 1999; Costantini et al, 2011; Cardellicchio et al, 2013), and the tactile consequences of observed actions, even when object presence is only known about but not seen (Turella, et al., 2011). These studies suggest that objects do serve as a basis for predictions during action observation. But if this is the case, why was no effect of intention found here for object presence?

One possibility is that the object was not processed fast enough to modulate visual predictive processing. In the visible conditions, participants could wait until the end of the sequence to process the object for the catch trial questions, as it was available throughout the action sequence. However, in the occluded conditions, object presence had to be processed prior to the reach, so here any effect of intention should have been present but was not.

Alternatively, as reaches were kinematically identical, it may have been that action predictions were not distinct enough: in effect, both object-directed and non-object directed body movements were the same, compared to the categorically different reaches and withdrawals in the previous experiments.

Especially if one believes that predictions originate from higher-level codes (Clark, 2013; Friston, 2010), a categorical representation might be likely, where actions are distinguished by their gross movement patterns – such as towards or away from an object – rather than by subtle kinematic features, such as, perhaps, a slightly more determined motion towards present compared to absent objects.

The above issues therefore suggest that the representational momentum paradigm may not be the best method to identify differences in top-down and bottom-up perceptual about object presence. While any visual predictions may have been similar regardless of whether an object was present or absent, tactile predictions associated with the sensory consequences the action – whether contact happens or does not – should be far more distinct, and again reflect a categorical difference between predictions in the two conditions. Therefore, one way to address these issues is to shift the paradigm from a purely visual representational momentum task to a cross modal tactile paradigm. This would allow a different test of whether differences in visual perceptual prediction are produced by object presence by measuring how they affect tactile perception.

Tactile responses are a good candidate for investigating the impact of action observation because they are both embodied and perceptual, allowing an insight into how much participants 'feel' the action they see. A number of studies have demonstrated that observing touch activates somatosensory areas of the brain, and can enhance touch perception (Bufalari et al, 2007; Serino, Pizzoferrato & Làdavas, 2008; Cardini, Tajadura-Jiménez, Serino & Tsakiris, 2013; Morrison et al, 2013; Bach, Fenton-Adams & Tipper, 2014). Tactile perception is also enhanced when the affected body part is visible during tactile stimulation (Haggard, 2006; Tipper et al, 2001). This suggests that different sensory inputs sum together for efficient perceptual processing. In this case, vision of touch enhances tactile perception. In fact, Haggard (2006) also found that the sight of another persons' hand compared to the sight of an object improved tactile perception when one's own hand was hidden, showing the extent to which bottom-up visual information can effect tactile perception. These studies therefore provide ample evidence that object-directed actions do elicit predictions about the sensory consequences of the action.

Therefore, using the same stimuli as in the previous experiment, the following two studies investigated the role of object presence on tactile predictions by measuring participants tactile perception when observing an actor either reach to touch an object or reach into empty space. Previous research has shown that the somatosensory cortices are involved in object-directed reaches, but not pantomimed reaches, suggesting a prediction of the resulting contact (Turella et al, 2011). The occluded conditions allowed investigation of whether these tactile predictions continue when touch cannot be seen but only inferred.

In Experiment 12, participants had to detect supraliminal tactile stimulation while watching visible or occluded reaches towards objects or into empty space. Tactile detection times were compared, in order to test if predicted contact speeds up detection of tactile stimulation, both when contact was observed and supplied top-down. In light of the findings of the Umilta study (2001), the hypothesis was that object presence would lead to faster detections to tactile stimulation, both in the visible and the occluded conditions. This would provide evidence that object presence does affect perception, suggesting that objects do provide a cue for predictions.

In Experiment 13 the same experimental set up was employed, but stimulation was administered at detection threshold, to conduct a signal detection analysis to measure the sensitivity and biases of tactile perception when object presence and bottom-up and top-down information varied. Again it was hypothesised that detection profiles would be affected similarly by object presence, across conditions of object visibility.

These experiments therefore allow testing whether prediction of object contact affect the observers' own tactile processes, and whether these effects differ for bottom-up and top-down guided predictions of contact. While the prior mirror neuron work (Umilta et al., 2001) suggests that similar processes occur in the visible and occluded conditions, other findings pointing to altered processing in the top-down cases. For example, Avenanti and colleagues (2013) reported that, when visual processing of an action was disrupted via TMS to the superior temporal sulcus, tactile-motor processing was enhanced, as if tactile processing stood in for the missing visual information (Avenanti et al., 2013). Others have reported that the anticipation – but not perception – of the consequences of one's own actions often lead to changes in sensitivity, rather than the changes in response bias reported above for directly observed contact

(Desantis, Roussel & Waszak, 2014; van Ede, Jensen & Maris, 2010; van Ede, de Lange & Maris, 2014).

Experiment 12 – Object cue – Tactile reaction time

The aim of Experiment 12 was to assess detection times to tactile stimulation on participants' own finger when observing the actions of others, in conditions where object presence and bottom-up information varied. The observed actions were identical to those used in Experiment 12, apart from in two aspects (see Figure 20). Firstly, in order to make sure all conditions were matched temporally to control for sequence length and level of motion in the scene, here the non-occluded conditions were lengthened. The images depicting the introduction of the occluder from the occluded conditions were reversed and added to the non-occluded conditions, so that the object or empty space began occluded and was revealed before the action onset. Secondly, instead of the reach stopping mid-way through, it continued until just prior to contact with the object. In all conditions at this point, the scene disappeared and participants had to respond as fast as possible if they detected tactile stimulation on their finger. If stimulation occurred it always happened immediately after the end of the movie, which in the case of reaches to objects coincided with the moment of touch. This allowed the investigation of how the prediction of touch affected one's own tactile perception.

Previous research has shown that the observation of touch can facilitate one's own perception of touch, providing evidence that we predict the sensory consequences of others' actions (Serino et al., 2008; Serino, Giovagnoli & Làdavas, 2009; Cardini, Bertini, Serino & Làdavas, 2012; Bach, Fenton-Adams & Tipper, 2014). Therefore, it was hypothesised that responses to tactile stimulation would be faster when observing a reach to touch an object, compared to a reach to empty space. This would support the notion that the consequences of touch are predicted when observing the actions of others, using one's own sensory-tactile system, and would provide evidence that, contrary to the results of experiment 11, objects do act as a cue for predictive processing.

In addition, as before, in half the trials the reaches were fully visible while in the other half an occluder obscured the location of the end of the action, and therefore whether contact occurred, from view. This allowed the comparison of top-down predictions of touch when the amount of bottom-up information varied. In the visible condition all of the action up until the moment of contact was available so that touch was all but observed, and could therefore be predicted bottom-up from sensory information. In contrast, in the occluded condition, the end of the action was hidden from view, so that touch could only be predicted based on prior knowledge of object presence. The comparison of observed and predicted touches (relative to observed and predicted reaches into empty space) allowed the investigation of how top-down prior knowledge alters one's own perception.

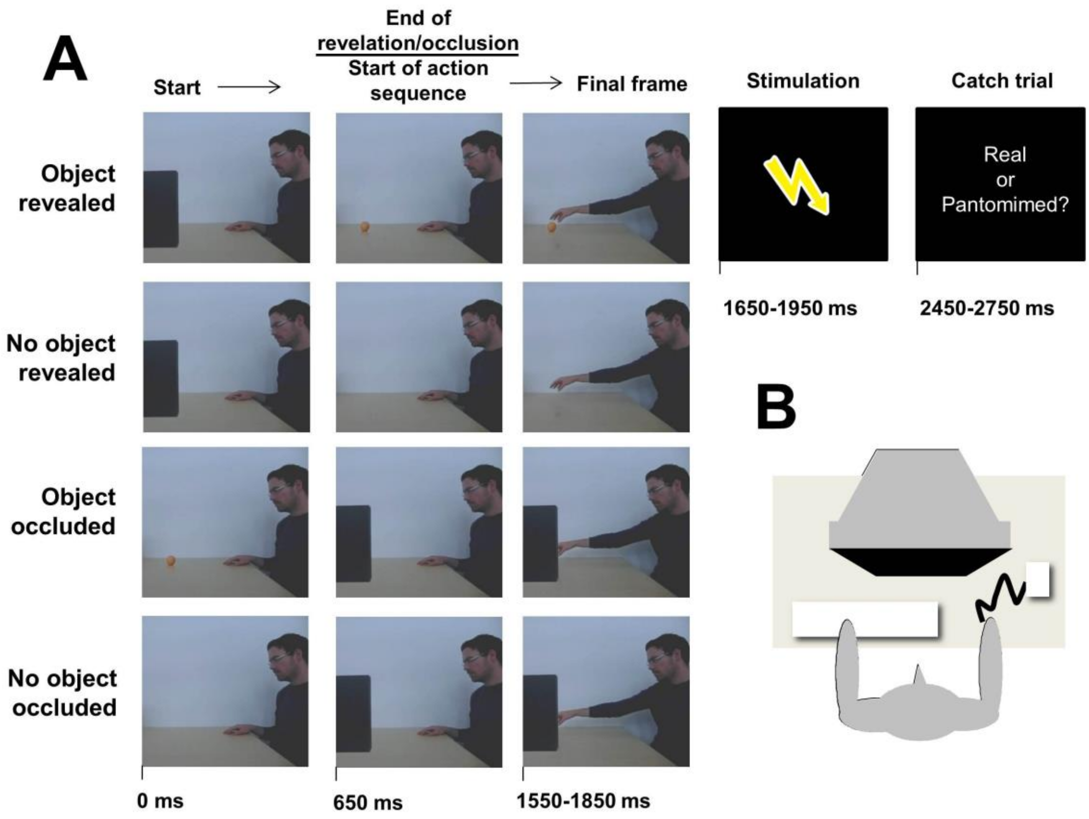


Figure 21. A: Schematic of the design of Experiment 12 and 13. Participants watched reaches towards objects or into empty space while the point of contact was either fully visible or hidden behind an occluder. Just before contact would be made, the scene disappeared and participants had to detect either above threshold (Experiment 12) or at threshold (Experiment 13) tactile stimulation on their own fingers (administered in 50% of trials). In 20% of trials, a catch trial question was presented afterwards, asking participants whether the action they just saw was directed at an object or empty space ("Was the action real or pantomimed?"). B: Schematic illustration of the experimental setup showing stimulator attached to index finger of right hand and left hand over the spacebar to report a tactile detection.

Method

Participants. 36 (26 females, mean age = 20.2 years, SD = 4.3) participants took part in the experiment. Participants were recruited and rewarded in the same manner as Experiment 11, adhering to the same ethical guidelines and participant requirements.

Stimuli and apparatus. The stimuli and apparatus were identical to Experiment 11 apart from the following.

In order to control for the different length of the conditions within Experiment 11 here all conditions were made temporally equivalent. To achieve this the opening frames from the occluded conditions were reversed and added to the start of the non-occluded conditions. This meant that the object or empty space began occluded before the occluder moved out of the frame to reveal the object or empty space before the action started (see Figure 21). This made all conditions temporally identical.

In addition to this for each reach extra frames were extracted, so that the whole of the action was present in the action sequence up to until just prior to object contact (see Figure 21). As before, for the occluded conditions, the extra frames of the sequence were modified using Corel PaintShop Photo Pro x6 to include the occluder in the image. Due to the slightly different lengths of the reaches towards the four different objects this resulted in 3 different sequence lengths ranging from 18 frames (900ms) to 24 frames (1200ms) with identical sequence lengths in the object present and object absent conditions.

Tactile stimulation was delivered via a custom-built amplifier and Oticon BC462 bone conductors (100 *X*), which were attached with a gauze band to the underside of the tip of the participants' right index fingers. The bone conductors convert auditory input from the computer's sound card into vibrations that can be varied in terms of frequency and amplitude. The tactile stimulus was a 200 Hz sine wave overlaid with white noise of 50ms duration. The first and last 10ms were faded in and out to prevent sharp transients.

Design & procedure. Participants were seated in a dimly lit room facing a colour monitor at a distance of 60 cm. After the experiment had been verbally explained to participants, the tactile stimulators were connected to their right index finger and ear defenders were placed over their ears to block out background noise. First, a calibration was performed to find participants approximate detection threshold. The tactile stimuli to be used in the main experiment were administered in a constant stream every 1000ms. Stimulation began at the lowest intensity and was slowly increased until the participant reliably detected the stimulation. This stimulation level was then used for the main experiment.

The main experimental session began with computer driven instructions, before a short training phase of 8 training trials (4 with stimulation). During the training, the catch trial question was administered in every trial in order to train participants to pay attention to the presence or absence of an object. Each trial began with the presentation of a fixation cross for 500ms. After a 300ms blank screen, the stimulus sequence was presented (1,550-1850ms total), followed by a 750ms blank screen. The tactile stimulation was administered 100ms after the start of this blank screen in 50% of the trials. Participants were asked to press the space bar as quickly as possible if they detected stimulation. Participants had 2,000ms in which to make a response. Like in Experiment 11, in order to direct participants' attention to the presence of an object a catch trial question ("Was the action real or pantomimed?") was asked at the end of the trial. The question asked was changed from the previous experiment in order to try and encourage participants to process the object in terms of its consequences to the action rather than its mere presence. Participants were instructed that a reach towards to an object was a "real" action and a reach towards empty space was a "pantomimed" action. Unlike the previous experiment this time the catch trial question was presented randomly with a 20% chance in each trial.

A total of 256 trials were presented in the main experimental session, in which each of the four conditions was presented at equal rates in a randomized order. Half the trials included stimulation (128 trials) while the other half included no stimulation. Stimulation was administered at the previously calibrated threshold intensity. The whole experiment lasted approximately 25 minutes.

Results

Exclusions. As the stimulation intensity was supraliminal, it should have been obvious for participants to detect, and this was reflected in the overall hit rate (M= 92%, SD= 16%). However, due to experimenter error, for some participants a too low stimulation intensity was chosen and this was reflected in hit rates well below 90%. These participants were excluded. In addition, our catch trial question ("Was the action real or pantomimed?") was designed to both draw attention to the presence or absence of the object, but also to measure task attention. Therefore participants whose catch trial accuracy was below 75% were also excluded, resulting in 2 further exclusions. The data of the remaining 27 participants' data was analysed fully. The percentage of hits to stimulation for these participants was 98%, and their mean question accuracy was 94%.

Reaction times. The data for reaction times, hits and false alarms for these participants were then entered into separate 2 x 2 repeated measures analysis of variance (ANOVA) with the factors Object Presence (present or absent) and Visibility (revealed and occluded). The ANOVA revealed no main effect of Visibility, $F(1,26) = 0.48$, $p = 0.52$, $\eta_p^2 = 0.02$, 95% CI [-7, 15]. There was, however, a main effect of Object Presence, $F(1,26) = 8.57$, $p < 0.01$, $\eta_p^2 = 0.25$, 95% CI [6, 29], with participants detecting tactile stimulation more quickly when viewing object-directed reaches compared to reaches into empty space. In addition, there was a significant interaction of Visibility and Object Presence, $F(1,26) = 5.03$, $p = 0.03$, $\eta_p^2 = 0.16$, 95% CI [2, 29], indicating a larger effect of object presence for occluded relative to fully visible objects. Indeed, paired t-tests showed that object presence had only a numerical effect on tactile detection for fully visible actions, $t(26) = 1.37$, $p = 0.18$. For occluded actions, however, participants detected tactile stimulation more quickly when viewing object-directed actions compared to reaches into empty space, $t(26) = 3.79$, $p = 0.001$ (see Figure 22).

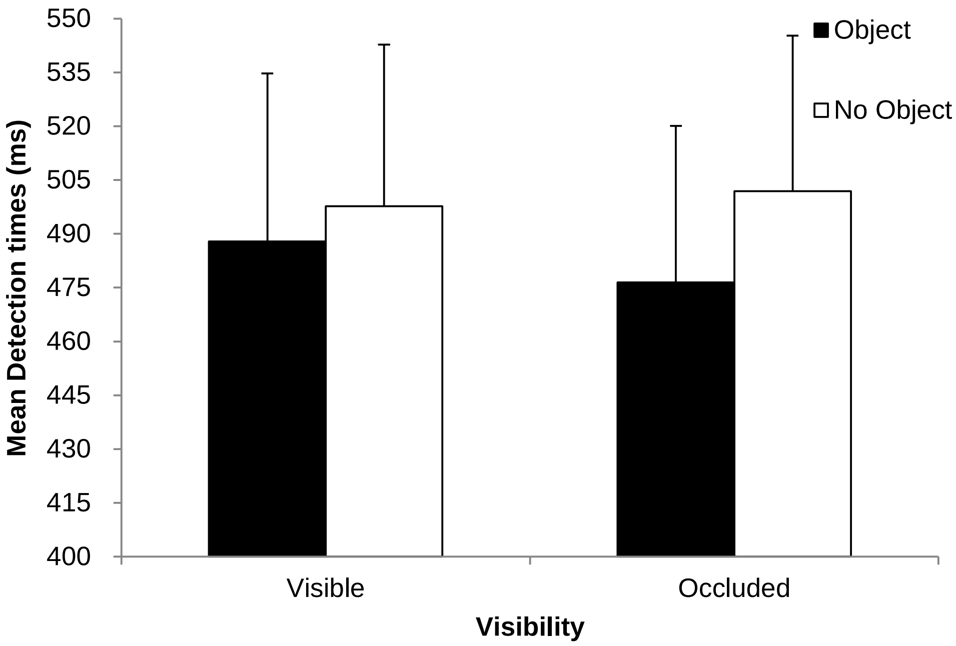


Figure 22. Mean detection times for tactile stimuli depending on whether participants viewed reaches towards objects (black bars) or into empty space (white bars), depending on whether the region of contact was visible (left bars) or occluded (right bars) in Experiment 12. Error bars represent 95% confidence intervals.

Hits. As the stimulation was supraliminal, there was little variation in hits between conditions. There were no significant effects, all *F's* <0.68.

False alarms. As stimulation was supraliminal, false alarms were rare (M = 1%, SD = 1). Nevertheless, there was a trend towards a main effect of visibility, $F(1,26) = 3.98$, $p = 0.06$, $\eta_p^2 = 0.12$, 95% CI [0, 1], with more false alarms when the end point of the action was visible than when it was occluded. There was also a trend towards an interaction, $F(1,26) = 3.46$, $p = 0.07$, $\eta_p^2 = 0.12$, 95% CI [0, 2]. Subsequent paired t-tests revealed that, in the visible conditions, there was a trend for participants to falsely report more stimulation when observing object-directed reaches compared to reaches into empty space, $t(26)= 1.69$, $p = 0.10$. In contrast, for occluded actions, numerically *fewer* false alarms were made for object-directed reaches, compared to reaches into empty space, $t(26)= 1.36$, $p = 0.19$. The main effect of object was not significant, $F = 0.45$, $p = 0.51$.

Discussion

The results of the current experiment showed, as expected, that observing an actor's reach to grasp an object resulted in the faster detection of tactile stimulation on one's own finger, compared to when observing the same reach directed into empty space. This demonstrates the effect of object presence on tactile perception and can be taken as evidence that predictions of contact facilitate tactile perception. In all conditions, the time when stimulation would occur was the same, so participants could anticipate the time of stimulation regardless of the condition. Despite this, when an object was present, reaction times were faster. This complements prior research that has shown that the observation of touch can speed up tactile perception on one's own finger (Bach, Fenton-Adams & Tipper, 2014). It also fits with prior research showing perceptual resonance when observing the actions of others that is taken as evidence for motor matching (for a review see Avenanti, Candidi & Urgesi, 2013). However, of importance here is that the current experiment provides evidence for perceptual - rather than motor - prediction, as the expectation of touch facilitated tactile detections.

Interestingly, in the occluded conditions, when the end of the action was hidden during action observation, knowledge of object presence led to even faster detection of tactile stimulation compared to reaches to empty space, even though visually the actions were identical in both conditions. There are two possibilities to account for this finding. One possibility is that, in the occluded conditions, predictive processing was more necessary, because crucial parts of the action – the goal object – was missing from view. Predictive coding might therefore have been explicitly recruited in these conditions to fill in the missing information. The larger decrease in response times to object-directed reaches in the occluded conditions would therefore reflect this increase in predictive processing due to the reduced visual information.

Alternatively, however, it might be that top-down and bottom-up predictions of contact rely on different mechanisms, and the false alarm data do provide preliminary evidence for this idea. In the visible conditions, object presence increased false alarms, while in the occluded conditions object presence decreased false alarms numerically. This implies that object presence may produce different effects depending upon visibility, as it suggests that when the object is visible, tactile stimulation is detected faster but also induces false alarm when this is none, in line with prior work that visual information about other's hand-object contact lowers one's own tactile detection threshold (Bach et al., 2014; Morrison et al., 2013), In contrast, in the occluded condition, object presence leads to faster detections but with better accuracy, reflected in a relative *decrease* of false alarms. This is suggestive of an increase in sensitivity, in line with other research that has reported such sensitivity shifts for anticipated effects of one's own actions (Desantis et al., 2014).

Of course, as stimulation was supraliminal, false alarms were very rare. Caution should therefore be applied to any interpretation. In addition, the relevant interaction, while close, did not reach full statistical significance (p = .07). Despite this, when taken together the reaction time and false alarm data suggest that the differences observed reflect different processes in tactile perception depending upon the amount of visual information available. To investigate this more directly, a second experiment was conducted in which tactile perception was measured at detection threshold, which allowed a signal detection analysis to be run, which can dissociate effects on detection threshold and sensitivity.

Experiment 13 – Object cue – Tactile Signal detection

Based on the findings of Experiment 12 the detection time effect of object presence on tactile perception is enhanced when the end of the action is occluded. This was an interesting finding and one that suggested either that (1) top-down predictive processing might be stronger when bottom-up information is reduced, or (2) that the effects in both conditions might emerge from different processes, one affecting detection thresholds and the other affecting tactile sensitivity.

In order to better understand these differences a second experiment was run to investigate how the differences in visibility affected participants' sensitivity to tactile stimulation and detection thresholds. To do this, the strength and variety of stimulation was varied. Experiment 13 was identical to experiment 12 apart from the fact that instead of administering supraliminal stimulation to participants' fingers, stimulation was now at threshold (in fact, ranging in intensities from slightly above threshold to slightly below threshold). Applying stimulation at threshold allowed the running of a signal detection analysis that distinguishes two distinct factors determining responses to tactile stimulation: bias and sensitivity. Bias (*c*) measures the overall detection threshold: the amount of tactile evidence required for participants to report tactile stimulation. Within the current paradigm it allows the investigation of how far object presence alone increases the likelihood of a tactile stimulus being detected, while also providing potential evidence for illusory perception, in cases where visually perceived contact is enough to cause participants to report stimulation even though there was none (false alarms).

Sensitivity (*d-prime*) measures the accuracy of detection, that is, correct responses to stimulation combined with correct no responses. This provides information about how accurate participants are at distinguishing stimulation from the neuronal background noise when stimulation is absent. These measures should address the hypothesis motivated by experiment 12 that hand-object contact either produces different effects on tactile detection depending on whether it was directly observed or occluded, or whether it merely leads to stronger effects in the occluded conditions.

Method

Participants. 56 participants (11 males, mean age = 22.6 years, SD = 5.1) were recruited from the Plymouth University student participant pool and the wider Plymouth community. They received either course credit or payment (£8 per hour) for participation. All were right handed, had normal or corrected to normal vision. All provided written informed consent prior to participation and were debriefed following completion of the experiment.

Stimuli and apparatus. All stimuli and apparatus were identical to experiment 1.

Design & procedure. The design of Experiment 13 was identical to Experiment 12 apart from one key difference. Rather than the stimulation being only supraliminal as in Experiment 12, five different stimulation levels were used (90%, 88%, 86%, 84%, 82%), representing a gradient of strength ranging from detectable, 90%, to undetectable (or barely detectable), 82%. Due to the increase in stimulation levels, trial numbers in the main experimental session were increased to 320, 160 with stimulation equally distributed across stimulation levels and conditions, and 160 without stimulation. Participants were instructed to emphasise accuracy over response speed. The whole experiment lasted approximately 45 minutes.

In order to validate the detection gradient, after the calibration session participants completed a simple tactile detection task lasting about 3 minutes. Participants were asked to press the space bar whenever they detected stimulation. To match the visual input to the main experiment, participants were instructed to look at their own hand during this procedure. Sixty tactile stimuli were delivered randomly in a constant train, every 1,500ms, with 36 trials without stimulation randomly interspersed and participants pressed a space bar whenever they felt stimulation. After that, the experimenter analyzed the detection probabilities across these intensities. If the data showed a decrease from accurate detection at 90% stimulus intensity to chance performance at 82% stimulus intensity, the main experiment began. If no such decrease was detectable, a new calibration session was performed.

Results

Exclusions. To be considered for analysis, stimulation needed to be roughly at threshold. 5 participants with calibration errors, who detected stimulation almost never (< 5% of the trials), or in almost every trial (> 95%) were therefore excluded. Such data are inappropriate for signal detection analysis, for which cells with no misses or no hits need to be manually interpolated. Secondly, we excluded 4 participants that did not show at least a minimum improvement (< 10%) of responses in trials with stimulation compared to trials without stimulation and which therefore showed a random response profile without staircase. Finally, as in Experiment 12, participants were excluded if they had catch trial accuracies below 75%. Unfortunately, in the current experiment, probably due to the more demanding at threshold detection task, participants found it harder to pay attention to object presence. A relative high number (8 participants) was excluded due to insufficient accuracy in the catch trials. The data of the remaining 39 participants for reaction times, hits and false alarms for these participants were then entered into separate 2 x 2 repeated measures analysis of variance (ANOVA) with the factors object (object and no object) and visibility (revealed and occluded).

Hits and false alarms. The analysis of Hits (correct detections) did neither reveal a main effect of object, $F(1,38) = 2.27$, $p = 0.14$, $\eta_p^2 = 0.06$, 95% CI [-4, 1], nor of visibility, $F(1,38) = 3$, $p = 0.09$, $\eta_p^2 = 0.07$, 95% CI [0, 4], nor an interaction of these factors, $F(1,38) = 1.34$, $p = 0.25$, $\eta_p^2 = 0.03$, 95% CI [-6, 2]. The analysis of false alarms did not reveal a main effect of object ($p = 0.85$) or of visibility ($p = 0.51$), but a highly significant interaction, $F(1,38) = 10.46$, $p < 0.005$, $\eta_p^2 = 0.22$, 95% CI [1, 4], replicating the previous experiment. In the visible conditions participants were significantly more likely to falsely detect stimulation when viewing object-directed actions (M = 0.04, SD = 0.05) compared to non-object directed actions (M = 0.03, SD = 0.04), $t(38) = 2.13$, $p < 0.05$. Conversely, in the occluded conditions there was a trend for participants to falsely detect stimulation more when viewing non-object directed actions (M

= 0.04, SD = 0.05) compared to object-directed actions (M = 0.03, SD = 0.04), $t(38) = 1.95$, $p < 0.06$ (see Figure 23).

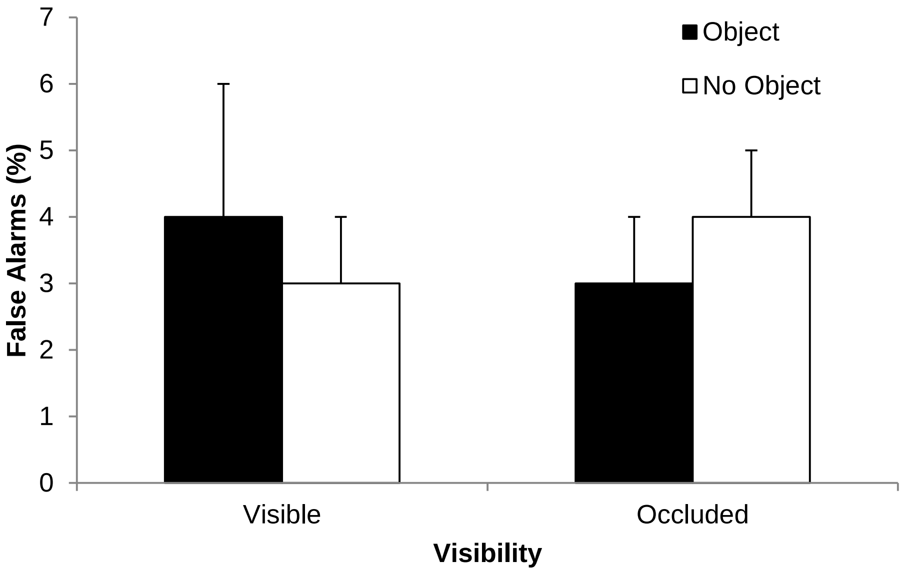


Figure 23. Mean percentage of false alarms (erroneous detections of tactile stimulation) for tactile stimuli depending on whether participants viewed reaches towards objects (black bars) or into empty space (white bars), depending on whether the region of contact was visible (left bars) or occluded (right bars) in Experiment 13. Error bars represent 95% confidence intervals.

Signal detection analysis. A main focus of this experiment was to test whether differences in hits and false alarm reflect differences in sensitivity and bias measures. For each participant both *d-prime* (sensitivity) and *c* (bias) was calculated and these were entered into separate 2 x 2 repeated measures analysis of variance (ANOVA) with the factors object (object and no object) and visibility (revealed and occluded).

The analysis of sensitivity revealed neither a main effect of object, $F(1,38) = 1.03$, $p = 0.32$, $\eta_p^2 = 0.03$, 95% CI [-5, 17], nor or visibility, $F(1,38) = 0.27$, $p = 0.61$, $\eta_p^2 < 0.01$, 95% CI [-7, 13]. However, the interaction was significant, $F(1,38) = 7.33$, $p = 0.01$, $\eta_p^2 = 0.16$, 95% CI [7, 46], (see Figure 24). Post hoc t-tests revealed that there was no significant difference between object-directed (M= 2.31, SD= 0.82) and non-object directed actions (M= 2.39, SD= 0.80) in the visible conditions, $t(38)= 1.08$, $p = 0.29$. However, there was a significant difference between object-directed (M= 2.42, SD= 0.89) and non-object directed actions (M= 2.23, SD= 0.87) in the occluded conditions, $t(38)= 2.46$, $p=0.02$. This demonstrated that participants' ability to detect tactile stimulation was significantly better when viewing object-directed actions in the occluded condition compared to occluded non-object directed actions.

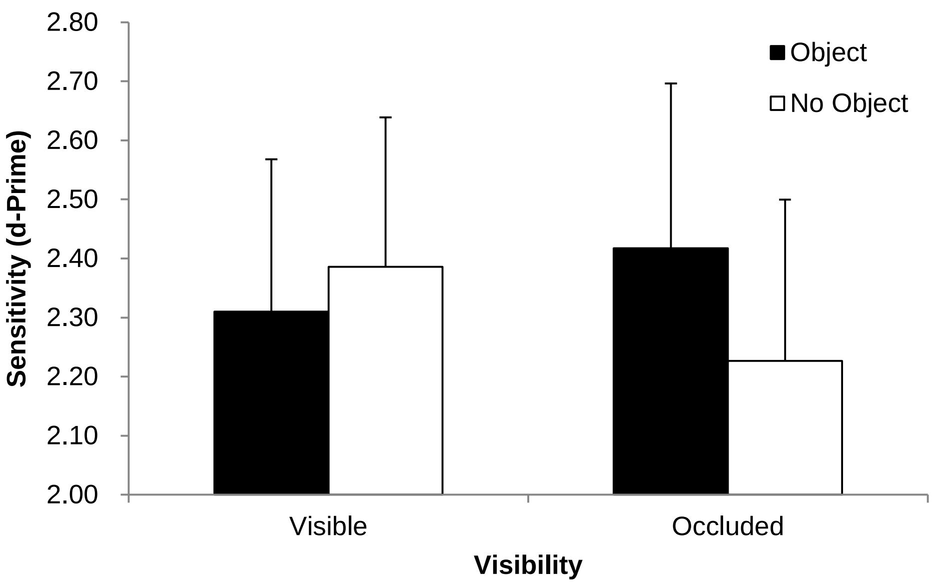


Figure 24. Mean sensitivity (d-prime) scores for the detection of tactile stimuli depending on whether participants viewed reaches towards objects (black bars) or into empty space (white bars), depending on whether the region of

contact was visible (left bars) or occluded (right bars) in Experiment 13. Error bars represent 95% confidence intervals.

The analysis of the response bias again did neither reveal a main effect of object, $F(1,38) = 0.348$, $p = 0.559$, $\eta_p^2 < 0.01$, 95% CI [-11, 21], nor of visibility, $F(1,38) = 0.345$, $p = 0.56$, $\eta_p^2 < 0.01$, 95% CI [-20, 11], but the interaction was significant, $F(1,38) = 9.82$, $p < 0.005$, $\eta_p^2 = 0.21$, 95% CI [16, 67], (see Figure 25). Post hoc t-tests revealed that, for fully visible actions, participants had a stronger bias to respond when observing object-directed actions compared to reaches into empty space, t(38)= 2.23, $p = 0.03$. For occluded actions, there was no such difference, $t(38) = 1.68$, $p = 0.1$, and, if anything, the effect was in the opposite direction.

In summary, therefore, the signal detection analysis revealed two different effects of object presence in the occluded and fully visible conditions. For fully visible actions, object presence increases response bias. In contrast, for occluded actions object presence increases sensitivity, but not response bias.

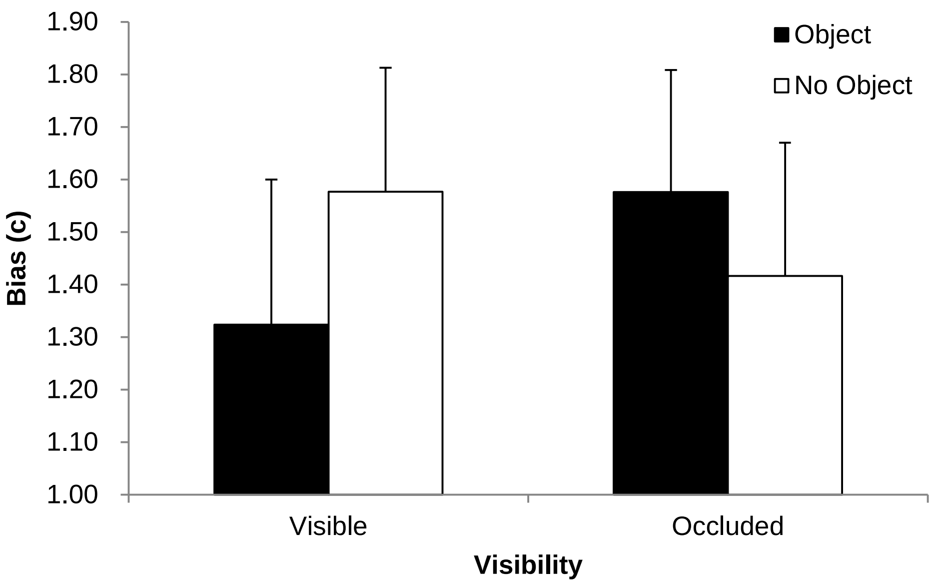


Figure 25. Mean response bias scores for the detection of tactile stimuli depending on whether participants viewed reaches towards objects (black bars) or into empty space (white bars), depending on whether the region of contact was visible (left bars) or occluded (right bars) in Experiment 13. Error bars represent 95% confidence intervals.

Discussion

By measuring tactile perception at threshold, Experiment 13 allowed testing whether the different levels of visual information affect tactile perception differently. When the whole of the action was visible, participants were more likely to both correctly (hits) and falsely report stimulation (false alarms) when viewing reaches to objects compared to viewing reaches into empty space. In contrast, when the end of the action was occluded, and hand-object contact could only be inferred, if anything the converse was true, with participants more likely to falsely report stimulation when viewing reaches to empty space.

This shows that the direct observation of touch elicits a bias to report sensations on one's own fingers, which could be due to the lowering of ones tactile threshold for perceiving touch, akin to some form of illusory perception. This fits with others studies that show observing touch can enhance tactile perception, leading participants to sometimes report stimulation even when there was none (Blakemore et al., 2005; Schaefer, Heinze & Rotte, 2005; Ro, Wallace, Hagedorn, Farne, & Pienkos, 2004; Serino et al., 2008; Bach et al., 2014). More interestingly, in the occluded conditions, the ability to *accurately* detect tactile stimulation was improved by object presence. Here then the prediction of touch did not produce a response bias, but instead improved the sensitivity of participants' tactile perception, allowing participants to more clearly distinguish tactile stimulation from background noise.

These differing effects of object presence on tactile perception when visual access to hand-object contact is varied provides evidence for dissociable processes depending on whether touch is seen or just inferred. It seems that while visual information biases perception in line with expectations, this bias is eradicated if the end of the action is occluded and instead tactile perception is enhanced.

General Discussion

Across three different experiments visual and tactile perception were measured while participants observed reaches towards an object or empty space, when the object or empty space were either visible or occluded prior to the start of the action. The aim of the experiments was to investigate the influence of object presence on perceptual prediction during action observation. In addition, the relationship between bottom-up and top-down prediction processes was tested, by varying the visibility of the object during action observation, to measure how perceptual prediction altered when the level of visual sensory input was changed. While fully visible reaches provide direct bottom-up information about the intention of the observed action, this information can only be supplied by top-down information when the target location is hidden prior to action onset. Taken together, the findings provide evidence that objects do generate predictions based on their cue to intention, which modulate perception, but that this affects predictive processing differently depending upon the sensory domain measured (visual or tactile) and the amount of sensory information available.

When measuring visual perception, once again a reliable representational momentum effect was found demonstrating a robust tendency to predict the future course of an observed action, and report it to be displaced further into the future than it actually was. This supports the findings of the previous chapters and prior research (Hubbard, 2005; 2014), and extends them by demonstrating that the effect remains when observing biological actions in real time. However, the presence of an object as a cue to the intention of the action did not increase visual perceptual prediction. Observing reaches to an object, compared to those to empty space, did not increase the likelihood of perceiving the action further along its trajectory (i.e. representational momentum). Likewise, the occlusion of the object or empty space prior to action observation also did not significantly alter the level of perceptual prediction. This suggests that objects may not be a salient enough cue of intention to affect visual perception. However, the visual similarity between the actions and the concern that top-down predictions may have not had sufficient time to affect visual perceptual processing may explain the lack of an effect of intention.

In two further experiments participants' tactile detection ability was measured while seeing others' full reach for objects or into empty space, when the point of contact (or non-contact) was either visible or occluded. The two experiments revealed that both visually guided and inferred predictions of contact affect the observer's tactile processing. Yet, they also demonstrated that these tactile changes might emerge from different mechanisms. Observing a fully visible reach towards an object (compared to the same reach into empty space) led to a tendency to report stimulation even when there was none. This increase in false alarms was observed both when participants detected supraliminal tactile stimulation (Experiment 12) and when stimulation was at detection threshold (Experiment 13). It replicates prior reports that observing touch enhances the bias to feel touch and report illusory stimulation (Ro et al., 2004; Bach et al., 2014; Morrison et al., 2013).

In the prior studies, these effects have been interpreted as emerging from a neural summation of contact information from the observed action and the tactile input, such that any response threshold is reached more readily when both are available. They were predicted from the assumption that seeing others reach for objects might induce the same prediction processes that inform observers about the impeding sensory consequences of their own actions, as if they happened on the participants' own body. As such, the current effects are in line with recent views of multisensory integration, which assume that visual and tactile information summate, in a Bayesian manner, to produce an integrated perceptual experience (Ernst & Banks, 2002; Deneve & Pouget, 2004; Wozny, Beierholm & Shams, 2008; Talsma, 2015).

In contrast, touch that could only be inferred – because the point of contact was hidden behind an occluder – did not induce such a bias to report illusory stimulation. Instead, inferred touch improved observers' ability to discriminate between whether stimulation occurred or not. It sped up detection of tactile stimuli on one's own fingers (Experiment 12), while *lowering* false alarms (Experiment 12 and 13). Indeed, the signal detection analysis revealed that this change reflected a change in tactile sensitivity rather than response bias. Thus, while directly observed touch led to a tendency to report stimulation even when there was none, inferred touch made participants better at distinguishing tactile stimulation from no stimulation. This suggests that merely knowing an object is present behind the occluder improves tactile perception, demonstrating that objects are a strong cue for top-down predictive processing.

While the bias shift for visible contact is indicative for neuronal summation of visual and tactile signals, this sensitivity shift for occluded contact reveals an enhancement of tactile processing itself, similar perhaps to the changes present if one anticipates (but not observes) contact on one's own finger (van Ede et al., 2010; van Ede et al., 2014) or foot (Carlsson, Petrovic, Skare, Petersson, & Ingvar, 2000), or anticipates the distal sensory consequences of one's own actions (Desantis et al, 2014). As such, the data from inferred touches are in line with recent predictive coding models of the brain (e.g., Kilner et al., 2007). According to this view prior experience helps to generate predictions related to current perception allowing anticipation and proactive behaviour. Here, therefore, rather than motor simulation providing a basis for action understanding, instead a prediction based on what the observer 'thinks' will happen – based on their own prior experience with tactile interactions with objects – allows observers to test its hypothesis against the incoming sensory input. These predictions are assumed to affect perceptual processing itself, and to lead to changes in coding *precision* of the perceptual input, predicting changes in tactile sensitivity (rather than bias) just as was observed here (Howhy, 2012; Den Ouden, Kok & De Lange, 2012; Clark, 2013; Seth, 2014).

A possible alternative explanation for the tactile effects observed here could be that they result from a general effect of increased arousal – and the associated heightened state of attention – produced by the presence of the object, as opposed to any increase in top-down predictions specifically in somatosensory processing. However, it is not clear how such arousal related explanations could account for both the results on detection threshold and tactile sensitivity. Nevertheless, one way to address such a concern would be to employ a control experiment to see if the same results emerge when participants have to respond to auditory, rather than tactile, stimuli. Indeed, in a previous paper such a control experiment has been used to show that the bias to perceive tactile stimulation when viewing reaches to painful objects, compared to safe objects, did not persist when auditory stimuli were used (Morrison et al., 2013). This suggests that the bias effect found was not due to a general arousal effect, supporting the ideas that the sensory predictions of touch affected tactile perception. This supports the notion that the effects found here are also the result of sensory predictions of touch rather than attention or general arousal. Moreover, the convergence of the findings across both response times and signal detection measures strengthen the interpretation that the effects emerge from top-down sensory predictions.

In conclusion, the experiments here show that predictive processing is influenced by object presence, but that the nature of these effects is dependent on the sensory modality. While object presence did not affect perceptual prediction when measuring visual perception, it was shown to have an effect when measuring tactile perception. Interestingly, this effect of object presence differed depending upon the amount of bottom-up information available. When the object was visible during action observation a bias to report tactile perception was revealed, even when no tactile stimulus was present. While when the object was known about but not visible during action observation the sensitivity of tactile perception increased. This demonstrates that objects do generate predictions and that the nature of these predictions can change depending upon the level of bottom-up information available. While Bayesian theories of multisensory integration can explain the bias effect found, with the summation of different sensory inputs, predictive coding theories can explain the sensitivity effect, with prior knowledge guiding top-down predictions of touch resulting in tactile perception being enhanced.

Chapter 6 - Predictive perception in a social context

Summary of findings

Across a series of experiments, the current thesis has provided robust evidence that people perceptually predict the actions of others (Experiments 1-13). This supports prior research that has demonstrated perceptual prediction for social stimuli (Thornton & Hayes, 2004; Hudson et al., 2009; Wilson et al., 2010; Uono, Sato & Toichi, 2010; Hudson & Jellema, 2011), and extends it to include the observation of object-directed manual actions. In addition, the tendency to perceptually predict increased in line with verbal cues to intention when these cues were both self-produced and produced by the actor being observed (Experiments 6-10). In contrast, intentional cues provided by the grip or gaze of the actor, or by the presence of a target object did not increase perceptual prediction in the visual modality (Experiments 2, 4, 5 and 11). Moreover, perceptual prediction was found to increase in response to subtle differences in kinematics when intentional cues were the same (Experiment 5). Finally, the presence of a target object modulated tactile prediction processes and the nature of the modulation differed depending upon the level of bottom-up information available during the action (Experiments 12 and 13). These studies provide a complex picture of perceptual prediction during action observation, which implies variation depending upon the level of prior knowledge, bottom-up sensory input and the number/types of sensory systems involved.

Predictive visual perception

Visual perception has traditionally been studied as a purely bottom-up process whereby input via the senses is received in low-level regions and propagated through the neural hierarchy in a feed forward manner up to high-level regions where it is interpreted. However, as research in the field has grown, it has become apparent that such a conception is not sufficient to explain how we process the complex visual environment within which we are situated and where the amount of bottom-up information is constantly shifting (Kveraga, Ghuman & Bar, 2007). This has led to theories positing the involvement of top-down feedback projections during perception, which interact with and inform feed-forward sensory inputs (Friston, 2010).

Indeed, research has begun to show that these top-down predictions can sometimes bias visual perception when expectations are met. Evidence for this has been found in relation to how we visually process motion. For example, motion has been shown to induce a very specific prediction of a spatial pattern at the leading, but not trailing, edge of a stimulus, providing evidence for forward predictions during motion perception (Roach et al., 2011). This evidence for forward predictions during visual perception has been supported by other studies investigating low-level vision, which have also demonstrated visual perceptual biases in line with expectations (Denison, Piazza & Silver, 2011; Hisakata, Terao & Murakami, 2013; Schellekens et al., 2014).

Modelling work on the visual system has also emphasised the importance of a top-down predictions for vision (Rao & Ballard, 1999; Lee & Mumford, 2003; Hosoya, Baccus & Meister, 2005). These models suggest that early visual neurons in V1 and V2 are tightly coupled to higher-order visual neurons not just in a feed-forward manner but also in relation to feedback from top-down inferences, assumed to help reduce delays in neural processing (Nijhawan, 2008). Within this conception, the visual system balances bottom-up sensory input against top-down inferences, which can directly affect perception by facilitating timely responses to changes in the environment.

The influence of these top-down predictions is also apparent from research comparing expected and unexpected stimuli. For example, expected stimuli tend to result in a decrease of neuronal activity, whereas unexpected stimuli tend to result in an increase (Kimura et al., 2011; Kimura & Takeda, 2015). Such findings are in line with the expectations of predictive coding in that the increase in neural response when viewing unexpected stimuli can be seen as equivalent to a prediction error (Winkler & Czigler, 2012; Stefanics, Astikainen & Czigler, 2014). In this way, prior knowledge aids perception through implemented top-down predictions that help to reduce the use of neuronal resources and anticipate future events.

These findings and models can explain phenomena such as the representational momentum effect, where the final position of a stimulus is judged as further along its motion trajectory, implying a visual prediction of its most likely future course (Freyd & Finke, 1984; for a review see Hubbard, 2005;

2014). Because the effect emerges when participants are instructed to accurately report the stimulus disappearance point, it reflects an at least partially automatic and involuntary forward prediction that happens even though the task incentivizes participants against it. It helps perceptual judgements of moving stimuli and allows planning of actions towards where it will be in the future rather than where it is in the present (Hubbard, 2006).

The experiments in the current thesis add to this previous research in the non-social domain, and extend research applying the effect to social stimuli (Thornton & Hayes, 2004; Hudson et al., 2009; Wilson et al., 2010; Uono et al., 2010; Hudson & Jellema, 2011), by demonstrating the effect when observing object-directed actions using a number of different social stimuli. This provides support for the notion that visual perception is at least partly predictive (Nijhawan, 2008), and is in line with current theories that suggest prediction is important when considering how the brain operates (Friston, 2010; Panichello et al., 2013). Moreover, it argues against theories of social perception focused on bottom-up mechanisms and emphasises the strong influence provided by predictive mechanisms. The current thesis demonstrates that such top-down predictive processes are directly influenced by social cues that provide information about an actor's intentions.

The effect of intentional cues on social perceptual prediction

A crucial aspect of social perception is the anticipation of people's future actions (Flanagan & Johansson, 2003; Wilson & Knoblich, 2005), and this often relies upon the prediction of the other person's intention (Liepelt, von Cramon & Brass, 2008; Woodward & Cannon, 2013). This ability arises developmentally early (Hunnius & Bekkering, 2010; Bakker, Kochukhova & von Hofsten, 2011), and is key for social interactions (Sebanz et al., 2006; Kunde, Lozo & Neumann, 2011; Ondobaka, de Lange, Newman-Norlund, Wiemers, & Bekkering, 2012) and social competition (Huys et al., 2009; Mann, Abernethy & Farrow, 2010). Prior research has hinted that such predictions are likely to be driven by the social cues generated by the observed, such as their eye gaze (Castiello, 2003; Pierno et al, 2006, 2008), bodily movements (Wilson & Knoblich, 2005; Becchio, Manera, Sartori, Cavallo, & Castiello, 2012; Thioux & Keysers, 2015) and verbal utterances (Baus et al., 2014), as well as the context within which such behaviours occur, stressing an importance for the role of objects (Costantini et al., 2011; Jacquet et al., 2012; Cardellicchio et al., 2013; for a review, see Bach et al., 2014). Current theories of how such predictions emerge in the brain also stress the importance of context for prediction generation (Bar, 2007; Barrett & Bar, 2009; Kilner et al., 2007).

While these studies imply predictive processing during social perception by showing that certain cues direct attention, for example, none of them measure changes in perception itself to assess whether it is directly influenced by intentional cues to the point where what is perceived actually changes. The current set of experiments fill this gap and are therefore novel in showing that intentions generated by language directly increased perceptual prediction of other people's actions, as measured by representational momentum.

We found that people consistently overestimated the vanishing point of a hand reaching towards an object in the direction of motion, such that probe stimuli displaced in the direction of motion were perceived as identical with the hand's last seen position, and probe stimuli displaced against the direction of motion easily detected. Importantly, when spoken verbal statements – such as "Take it!" or "Leave it" – matched the subsequent direction of an actor's arm movement towards or away from objects, this led to a further increase of this overestimation, compared to when statement and direction mismatched. This demonstrates that perceptual predictions of a hand's future course are directly affected by social, verbal cues.

This was true when the cue related to simple kinematic intentions (Forward or Backward, Experiment 6) and to higher level intentions related to the action's goal (Take it or Leave it, Experiment 7). This shows the power of intentional language as a cue that can directly bias perceptual judgements of others' actions. Moreover, the same visual bias was also evident when the verbal statements were not spoken but heard, as if spoken by the actor themselves. This demonstrated that the effect that language can have on subsequent perception does not only apply when the verbal statements are spoken but also when they are passively heard prior to action observation. This shows that both predictions generated internally and those generated externally, based on the language of another person, can bias perceptual judgements in line with the prediction. This perceptual bias can be seen as beneficial in facilitating anticipatory processing during social interactions (Sebanz & Knoblich, 2009). For example, the representational momentum effect is seen as crucial in “bridging the gap” between where a stimulus is now and where it will have to be responded to, considering the neuronal delays associated with perception and action planning (Hubbard, 2006). It also shows the influence of top-down processing on how we perceive bottom-up sensory inputs by demonstrating that a verbal cue can directly bias subsequent visual judgments of an actor’s reach further forward in its path. Here then bottom-up visual perception is not direct and unmediated but influenced by top-down prior knowledge and biased in line with top-down predictions associated with verbal cues specifying an actor’s intentions.

This finding fits in with a body of research that also demonstrates the effect of verbal information on perception. For example, matching verbal labels can bring supraliminally presented objects to visual awareness (Lupyan & Ward, 2013). In addition, self-produced speech has been shown to improve performance on a visual search task, particularly when the association between the visual target and the spoken word is strong (Lupyan & Swingley, 2012). Likewise, listening to task irrelevant directional verbs while performing a motion detection task improved participants’ sensitivity to motion when the heard verbs matched the direction of motion (Meteyard, Bahrami & Vigliocco, 2007). Moreover, it has also been shown that processing language that includes descriptions of motion can induce a motion aftereffect that is in line with the direction specified by the language (Dils & Boroditsky, 2010). All these studies emphasise that language perception (whether spoken or heard) can affect subsequent visual perception in line with the experiments described in Chapter 4.

As in the experiments from Chapter 4, in all these studies the verbal information appears to bias the visual processing, facilitating accurate detection in some situations while in others creating illusory perception. These findings support the idea that language can guide top-down predictions, which can directly influence bottom-up sensory inputs to both aid and distort perception. The new findings presented in this thesis reveal that such effects are not restricted to low-level verbal cues, or abstract point motion at the detection threshold (Meteyard, et al., 2007), but can directly emerge from cues suggesting an action goal, and affect the forward prediction of observed actions. They therefore provide evidence that inferred higher-level goals of other people are translated into the movements in space that would bring them about, such that these actions can be identified more effectively, and one's own actions can be planned in response (e.g., Csibra, 2007). This is line with other recent demonstrations that the anticipation of another person's action can facilitate one's own performance of the same, or a corresponding, action (Pfister, Dignath, Hommel & Kunde, 2013; Genschow & Brass, 2015).

An important question is what the representational momentum effect reflects. Traditionally, the effect has been assumed to take place in the "gap" between the offset of the moving stimulus and the reappearance of the probe, as if the movement continued in the observer's mind after it had disappeared (Hubbard, 1990; Hubbard & Bharucha, 1988). Yet, the effect peaks at gap lengths of 260 ms, much too short to reflect memory processes in the traditional sense, and much closer to lower level perceptual processes in iconic memory. Indeed, more recent research suggests that representational momentum can also reflect processes that happen during the perception of the movement (Jordan, Stork, Knuf, Kerzel, & Müsseler, 2002; Musseler, Stork & Kerzel, 2002). For example, as mentioned above, even during the perception of a moving stimulus, processing at the leading edge of the stimulus is enhanced, reflecting a forward prediction of what the observer will see that happens during movement perception (Roach et al., 2011). On the basis of the current thesis, it is not possible to distinguish between these alternatives. Both, however, reflect perceptual processes that are assumed to serve both cognitive judgments (e.g. where an object will be in the future), to fill in missing perceptual information (where it is while it disappears), and allow the planning of own actions towards it (Hubbard, 2006). Importantly, in at least one prior study (Hudson et al., 2009), gap length was varied while participants performed a representational momentum like task on heads moving in the direction of eye gaze or against it. At least for this study, the length of the gap (0 vs. 1000 ms) did not affect the amount of prediction at all. At both gap lengths, gaze biased the perception of head motion in its direction.

This suggests that the effect could result from processes, which occur throughout the movement and persist into the gap, rather than only occurring during the gap (Jordan et al., 2002; Musseler et al., 2002). One way to investigate whether the representational momentum effect emerges from processing during the movement or the gap would be to utilise another mislocalisation error: the *flash lag effect*. In the *flash lag effect* a flashed stimulus is perceived as lagging behind the position of a moving stimulus, even though when the flash occurs it is spatially aligned with the position of the moving stimulus. While there is some debate as to exactly what causes the effect (Eagleman & Sejnowski, 2000; Brenner & Smeets, 2000; Nijhawan, 2002; Munger & Owens, 2004), it does provide an opportunity to compare whether verbal cues modulate perception during the movement or during the gap, by, for example, comparing the size of the *flash lag effect* when an action either matches, or mismatches, a verbal cue (Chapter 4). If the size of the flash lag effect is the same in both predicted and unpredicted actions then it suggests that the effect of intention on representational momentum occurs during the gap. However, if predicted actions (verbal cue and action are congruent) produce a larger flash lag effect than unpredicted actions (verbal cue and action are incongruent) then it suggests that perceptual prediction occurs during the perception of motion.

Further experiments could also explore if varying the length of the gap affects the influence that verbal cues have on the size of the representational momentum effect. For instance, even if the effect of intention emerges during the movement, this does not rule out the possibility that perceptual prediction also occurs during the gap as well. Therefore, by varying the length of the gap one can test whether the effect of intention decreases or increases in relation to the size of the gap, remains constant and at what point the effect emerges.

The influence of bottom-up and top-down inputs on perceptual prediction

The representational momentum effect found across a range of studies highlights a visual bias to predict the future position of an actor's arm further along its current trajectory. When the direction of the action matched a prior verbal cue this visual bias increased. This reveals an effect of top-down predictions, automatically generated by the meaning of the verbal statement, on perception. The tactile experiments on the other hand, demonstrated that the effect of object presence on tactile perception differed depending upon whether the object was seen during the observed action or whether its presence could only be inferred. This subtle shift in the available bottom-up information resulted in a striking dissociation regarding its effect on tactile perception, with seen goal objects of a reach resulting in a perceptual bias and inferred objects producing an increase in perceptual sensitivity, reflecting an increased ability to distinguish tactile stimuli from the neuronal background noise.

The tactile perceptual bias observed when the object was visible during the observed action is in line with previous research. For example, studies have shown that observing somebody else touch an object can produce a bias to perceive touch on one's own finger even in the absence of a tactile stimulus (Morrison et al., 2013; Bach et al., 2014). This can be seen as an example of the dominant effect vision can have on tactile perception (Tipper et al., 1998, 2001; Press, Taylor-Clarke, Kennett, & Haggard, 2004; Ro et al., 2004; Haggard, 2006). Moreover, a number of studies have shown that observing touch activates somatosensory areas of the brain, suggesting that the social perception of touch is processed similarly to the direct experience of touch (Keysers et al., 2004; Bufalari et al., 2007; Morrison, et al., 2013; Singer et al., 2004; but see Chan & Baker, 2015, for a critical review). These findings are thought to emerge from the multisensory nature of perception, in that observing touch involves the same perceptual code as when perceiving one's own touch.

It appears then that in the visible condition seeing impending touch increased the perceptual evidence in support of the prediction of touch, as if the observed action was processed as one's own, such that it lowered the viewers' threshold for perceiving touch on their own finger. This is line with the idea that such effects reflect a neural summation of signals from the observed action and the tactile input resulting in a reduction in the viewer's response threshold. They are assumed to emerge from the same processes evolved to predict the sensory consequences of one's own actions. Such an interpretation is line with current views of multisensory integration within which different sensory sources (e.g. visual and tactile) summate to combine the information into a unified perceptual experience (Ernst & Banks, 2002; Deneve & Pouget, 2004; Wozny et al., 2008; Talsma, 2015).

In the representational momentum studies, verbal intentional cues biased visual perception. Similarly, here, a visual intentional cue, an object, biased tactile perception. That is object presence specified the intention of the actor to grasp the object, and this predicted intention modulated tactile perception despite object presence being irrelevant to the task. Both these findings demonstrate the affect that predictions can have on one's own perception and the tendency to integrate sensory inputs to unify perception. They are also in line with theoretical models suggesting that predictions are generated through the integration of prior knowledge and intentional cues (Csibra, 2007; Kilner et al., 2007).

In contrast, when the object was known about but visually occluded prior to action observation, any knowledge generated via bottom-up information had to be stored and retained. Therefore, any effect of object here resulted from the memory of the object rather than the perception of it. When the object or empty space was occluded during action observation, object presence resulted in an increase in perceptual sensitivity, instead of producing a response bias. Here it would seem the influence of memory and prior knowledge improved the precision of tactile perception compared to when bottom-up visual cues were permanently available. This effect appears similar to changes observed when contact on one's own finger (van Ede et al., 2010; van Ede et al., 2014) or foot (Carlsson et al., 2000) is anticipated (but not seen), or when the distal sensory consequences of one's actions are anticipated (Desantis et al., 2014). It is also in line with research showing that when visual processing is disrupted via TMS, tactile processing is enhanced, suggesting that tactile processing compensated for the reduction in visual information (Avenanti et al., 2013). These findings support recent predictive coding models of the brain which suggest that prior experience helps to generate predictions about current perception facilitating anticipation and proactive behaviour, and sharpening the representation of the expected input (Kilner et al., 2007).

One explanation for the differential effects may come from the nature of multisensory integration (van Atteveldt, Murray, Thut & Schroeder, 2014). Visual and tactile information, it has been proposed, are integrated in an optimal fashion via a maximum-likelihood estimate, produced by combining the different sensory information streams to increase the power of a given estimate (Ernst & Banks, 2002). However, this integration is fluid, and vision can dominate when variance of visual estimations are lower than variance of tactile estimations. Hence, in the visible condition, object presence may have produced a bias because the visual variance was low due to constant sensory input reinforcing top-down predictions of touch. In contrast, in the occluded condition the removal of visual information would have disrupted visual estimations providing more weight to tactile estimations, which resulted in object presence sharpening the precision of perception. This sharpening can be seen as occurring due to the tendency to integrate signals, explaining why knowledge of object presence did not bias perception but did improve sensitivity. Therefore the memory or awareness of object presence in the absence of any direct visual perception boosted tactile perception without overshadowing it and producing a bias.

Recent models that assume a hierarchical integration of multimodal and top-down information also capture this distinction (Altieri, 2014; Talma, 2015). In such models, information on the same level – such as observed and felt tactile stimulation here – are directly integrated with one another, such that the combined evidence leads to stronger sensations than when only one source of information is available, or one source can compensate for the other. These multisensory integration processes can rely on direct (or thalamus-mediated) connections between primary and secondary sensory cortices (e.g., Falchier, Clavagnier, Barone & Kennedy, 2002; van den Brink et al., 2014; for a review, see Talma, 2015), via which contact information from vision and touch can interact directly and feed from one channel to the other, inducing multisensory summation effects just as were found here when contact was directly observed.

In contrast, when stimulation is not observed visually but inferred from top-down information, such direct interactions cannot take place. Instead, in hierarchical prediction models (cf. Clark, 2013), such top-down expectations are assumed to primarily act on the *precision* of sensory coding in both modalities, reflecting a sharpening of the representation of the expected stimulus (e.g., Kok et al., 2012), or the distribution of attention to the specific time, location and features that characterise the incoming stimulus (Klemen & Chambers, 2011). As found here for inferred contact, these changes in coding precision would go along with changes in tactile sensitivity (rather than threshold), such that the stimulus can be detected more effectively and distinguished from background noise. Indeed, there is now converging evidence that anticipating consequences of one's own actions induces such changes in sensitivity, for both proximal and distal consequences of one's own actions, and across different stimulus modalities (Desantis et al., 2014; van Ede et al., 2010; 2014). Such effects can be likened, perhaps, to the very specific anticipations one experiences when groping, for example, for a door handle in the dark that one knows is there. The studies in this thesis show a similar effect may happen for actions one observes in others.

In sum, then, the different results here can be understood within the context of hierarchical predictive coding models of multisensory integration where different sensory systems and top-down information interact to shape perception. The extra visual information provided by seeing the object during the action strengthened predictions of touch which in turn lowered the threshold for perceiving touch. In contrast, when the object could not be seen but was only known about, predictions of touch did not lower the threshold for perceiving touch but did improve tactile sensitivity. The picture of the collected findings of the current thesis then seem to suggest that both top-down predictions and multisensory integration can interact to influence social perception and produce changes in what is perceived.

Open questions and future directions

In addition to addressing questions about the link between predictive coding and social perception, the current thesis also generated several questions that currently remain unresolved. In the following section some of the more important questions are discussed along with some potential avenues for future research designed to address them.

Multisensory integration as an explanation for the intentional cue effects?

The assumption guiding the representational momentum experiments was that the spoken and heard verbal cues acted as a top-down signal for predictive processing. However, given the results of the tactile detection experiments, an alternative interpretation is that these effects, too, reflect the integration of multimodal signals. Indeed, one important difference between the experiments that showed modulation of perceptual prediction during social perception compared to those that did not was the number of sensory modalities involved. The verbal studies (Experiments 6-10) included visual and auditory (spoken and heard) perception, while the tactile studies (Experiments 12 and 13) included visual and tactile perception, whereas the majority of unsuccessful experiments (Experiments 1, 2, 4, 5 and 11) only included visual perception. Therefore, one reason behind the modulation of perceptual prediction could have been the integration of these multiple sensory systems. While all experiments showed a visual perceptual bias (representational momentum effect), the modulation of this effect by intentional cues was only found when there was the requirement for multisensory integration (Experiments 6-10 & 12-13).

In the experiments from Chapter 4 all verbal statements were causally linked to the subsequent actions. This meant that each action began temporally close (within 200ms) to the offset of each verbal utterance whether spoken or heard. In the tactile experiments, tactile stimulation always occurred at the exact same time that the actor would make contact with the object. This meant that any prediction generated by the auditory stimuli in the verbal studies would be causally linked to the subsequent visual perception. Likewise, any prediction generated by the visual stimuli in the tactile studies would be causally linked to the subsequent tactile perception. In models of multisensory integration, such temporal and causal couplings are crucial for creating crossmodal effects (Ohshiro, Angelaki & DeAngelis, 2011; Zmigrod & Hommel, 2013). Therefore, it could be that the modulation of perceptual prediction in line with the action's intentions relied upon the integration of the different perceptual inputs. This would suggest that the initial sensory information, whether verbal or visual, activated the intention and therefore top-down prediction of the action, which influenced subsequent perception, whether visual or tactile through multisensory integration. The representational momentum effects could therefore emerge from the same multisensory integration mechanisms as the tactile prediction effects in the non-occluded conditions.

A number of different studies have shown that vision can affect tactile perception. For example, studies have shown that the presentation of a visual stimulus at the same time as a tactile stimulus can improve tactile perception (Johnson, Burton & Ro, 2006; Arabzadeh, Clifford & Harris, 2008). Likewise, merely observing a body part can also enhance the perception of touch (Kennett, Taylor-Clarke & Haggard, 2001; Taylor-Clarke, Kennett & Haggard, 2002; Tipper et al., 1998, 2001). In addition, observing oneself being touched enhances tactile perception and this also extends, albeit to a lesser extent, to observing another person being touched (Serino et al., 2008). Other studies have shown how observing touch can bias observers to perceive touch when it is not really there (Ro et al., 2004; Morrison et al., 2013; Bach et al., 2014).

Moreover, phenomena such as the rubber hand illusion show how perceiving touch while observing a rubber hand being touch in synchrony can lead to observers feeling as the though the rubber hand is their own hand, and thereby shift the perceived position of one's own hand towards the rubber hand (Botvinick & Cohen, 1998). All these studies demonstrate that tactile perception can be modulated depending upon the observer's visual perception in line with the studies reported in Chapter 5 (Experiments 12 and 13).

Likewise a number of effects demonstrate how the integration between auditory and visual information can alter perception (Parise & Spence, 2013). While many studies have shown the impact of vision on audition, a number of studies have begun to show how sound can alter visual perception (Shams, Kamitani & Shimojo, 2004). For example, an auditory stimulus has been shown to improve the perception, identification and perceived intensity of a visual stimulus (Stein, London, Wilkinson & Price, 1996; McDonald, Teder-Salejarvi & Hillyard, 2000; Vroomen and de Gelder, 2000). Similarly, the 'illusory flash effect' has demonstrated that when a single visual flash is accompanied by multiple auditory beeps, it is perceived as multiple visual flashes (Shams, Kamitani & Shimojo, 2000, 2002). The effect of auditory stimuli on visual stimuli has also been shown when perceiving motion. For example, visual motion has been shown to be susceptible to biases in the direction of an auditory motion stimulus (Meyer & Wuerger, 2001). Moreover, an auditory motion stimulus has also been shown to modulate perception of a static visual stimulus (Shams et al., 2004). All these studies demonstrate that visual perception can indeed be modulated by auditory perception, once again showing that in some circumstances this improves the precision of perception while in other cases it biases it.

Recent explanations of such crossmodal effects have appealed to Bayesian probability (Pouget, Deneve & Duhamel, 2002; Ernst, 2006; Talsma, 2015). Bayes theorem was originally applied to statistics but has subsequently been applied to a number of different fields including cognition (Knill & Pouget 2004). Essentially it involves a calculation between the current probability of an event occurring and the events prior probability to produce a more accurate measure of the events overall probability of occurring (Efron, 2013). This combination of prior and current information has made it a good candidate for assessing perception. All the information that we receive from the senses contains an element of uncertainty, which results in the reduction of perceptual precision. Therefore, Bayesian probability allows a method to address such uncertainties and improve precision. This makes it a useful tool for understanding the integration between the senses to optimise perception. Optimal cue integration theory suggests that when multiple sources of independent information are available combining them can reduce uncertainty and therefore improve perception (Fetsch, DeAngelis & Andelaki, 2010).

Therefore, the reason that the experiments with more than one sensory system involved resulted in the modulation of perceptual prediction might be because of the extra evidence that the multiple senses provided to such Bayesian priors. Here then the integration of multiple streams of perceptual information provided predictions more weight and led to more perceptual modulation. According to such a view, the dominant sense, visual in Chapter 4 and tactile in Chapter 5, is modulated via the quality of the additional multisensory integration (Ernst & Banks, 2002). This is line with recent models within which both multisensory and top-down signals are integrated hierarchically (Altieri, 2014; Talsma, 2015). According to such views, when different sensory sources are processed, the evidence from both sources can be integrated, so that any detection threshold is surpassed more readily, leading to a bias in perception. This can explain why the experiments involving multiple sensory systems modulated perception based on the combination of predictions specifying the intention of the action.

In such a view, the brain treats the intentional cues – the self-produced or heard verbal statements – not as top-down signals of the actor's intention, but as multisensory cues to motion that were integrated with what was really perceived. One way to test this hypothesis would be tease apart whether the effects on representational momentum – like those in the tactile detection experiments – reflect bias or sensitivity effects. On the basis of the present data in Chapter 4, both interpretations are possible. On the one hand, a verbal cue could have led to a bias that was "added" to the motion that was perceived, thereby biasing it into the expected direction, in the same way as tactile and visual multisensory signals of contact are summed to bias tactile perception. On the other hand, the representational momentum effects could reflect a sensitivity effect. In this view, representational momentum was increased because the verbal cue led to a top-down sharpening of the representation of these motion codes that matched the expected direction, thereby leading to a sharper impression of motion.

If the results from the tactile experiment are an indication, one way to distinguish these effects would be to present the hand's motion close to detection threshold, perhaps in an image with visual noise, and ask participants whether motion had occurred or not. If verbal cues lead to a top-down effect, then they should make participants better at distinguishing seen motion from the background noise, similar to the top-down effects in the tactile detection experiment. In contrast, if verbal cues lead to a bottom-up biasing in line with multisensory integration, they should lead to the tendency to "see" movement in the expected direction, even if there was no movement at all, similar to the bias effects for fully visible actions in the tactile detection experiment.

Of course, there are reasons why one could be sceptical of the idea that the effects on representational momentum reflect multisensory integration. For example, first, the effects did not differ depending on whether the verbal cues specified an intention (Take it, Leave it) or merely a movement path (Forward, Backward). If the effects reflect multisensory integration, the simpler kinematic cues that are more closely related to the incoming stimulation would perhaps be expected to produce larger effects. Similarly, the verbal cues "Forward" and "Backward" as well as "Take it" and "Leave it" are meaningful only from the perspective to the current actor and the situation they are in. From the perspective of the participants, in contrast, the actions go left and right rather than forward and backward. Thus, if the effects indeed reflect multisensory integration, then this integration does not reflect only lower level sensory cues to action. Instead, these integration processes themselves appear to be related to relatively sophisticated processes related to goal attribution and perspective taking.

One way to investigate this would to see if whether similar effects could be achieved with more high-level verbal utterances. While the verbal experiment discussed in Chapter 3 did not find any effect of emotional verbal labels, the methodology and experimental set up was different. Therefore, running similar experiments to those in Chapter 4 but with more emotional or abstract language, which can still be associated with the different objects, would be an interesting extension.

Are objects a cue to intentions?

Both the spatial matching experiments (Experiments 1 and 2) and the occluder representational momentum experiment (Experiment 11) relied on objects and their integration with actions as the cue to the action's intention, but neither produced (replicable) results. One potential explanation for this is that objects are not a strong enough cue of intention to produce predictions themselves. However, there are a number of reasons to discount such an interpretation. One reason is the large body of research that makes a strong case objects do facilitate top-down expectations. For example, a number of studies have provided evidence suggesting that objects prime the retrieval of manipulation knowledge supporting the assumption that, at least on some level, objects facilitate a prediction of possible actions (Tucker & Ellis, 2001; Myung, Blumstein & Sedivy, 2006; Helbig, Graf & Kiefer, 2006; Bub et al., 2008; Ellis et al., 2013). Research has also shown that premotor and parietal areas of the brain are activated when merely viewing graspable objects, implying some form of action information is involved in object recognition (Chao & Martin, 2000; Creem-Regehr & Lee, 2005). The effect of object affordances also appears to be apparent not just for the action possibilities of one own action, but also for others when they are in the vicinity of objects (Costantini et al., 2011; Cardellicchio et al., 2013). Most tellingly, a recent neuroimaging study showed that when viewing a match between an action and an object the brain reinforces the relationship in order to supress other competing actions, providing support that objects facilitate action predictions that specify the most likely goal of the action (Schubotz, Wurm, Wittman & von Cramon, 2014).

A similar reason is complementary research suggesting that appropriate grip types influence subsequent action processing. For example, seeing a matching grip type enhances judgements regarding the action's appropriateness (Bach, 2004). Likewise grip type has been shown to facilitate eye movements towards an object that affords it (Fischer et al, 2008), speed up recognition of matching objects (Helbig et al., 2010), and aid the prediction of an action's intention (van Elk et al., 2008; Ambrosini et al., 2011). Moreover, other studies have shown that kinematic cues are important for the accurate prediction of observed actions (Manera et al., 2011; Stapel et al., 2012). These studies suggest that grip type and kinematics are useful cues during object processing and action observation. Taken together this body of research suggests that seeing graspable objects accesses the appropriate manipulation knowledge associated with using the object, while observing grip types accesses appropriate objects to perform an action. This is good evidence that both objects and grip-types, and the relationship between the two, could guide action predictions during perception. This does not imply that actions have to be processed in order to extract their intention before a prediction can be made. Instead, it suggests that grips imply certain intentions that can generate predictions that influence subsequent perception. Likewise objects imply certain grips which can also generate predictions linked to the intention of the actor.

Indeed, there might be another reason for the lack of modulations in these experiments. In both the spatial matching experiments (Chapter 3) and the occluder representational momentum experiment (Chapter 5), all actions followed the same direction. In contrast, in the experiments from Chapter 4 (Experiments 6-10) the direction varied between action alternatives (reaches, withdrawals) and was directly linked to the linguistic intentional cues, which modulated perception (Take it, Leave it). In effect, both the observed actions and the associated predictions were *categorically* different. The assumption in the spatial matching experiments was that while matching grips would increase perceptual prediction mismatching grips would not. However, as the actions were all aimed in the same direction, only ever directed to one object and all finished well in advance of contact with the object, it could be predictions for matching and mismatching actions overlapped conceptually and were not so distinct. Representational momentum that would be elicited from these predictions would therefore not show strong differences. Likewise, in the occluder representational momentum experiment, the assumption was that reaches to objects would increase perceptual prediction whereas pantomimed reaches would not. But like the spatial experiments all actions were aimed in the same direction and finished well in advance of the object, meaning once again predictions for object-directed and pantomimed reaches may have overlapped and not been distinguished conceptually.

This might suggest that the difference between both object-directed and pantomimed reaches, and between matching and mismatching grips, may have been too subtle to differentiate the predictions based on their intentions. Especially if one assumes that predictions emerge from higher-level action expectations that are categorically different and, perhaps, distinguished by their gross movement patterns such an interpretation might seem likely. If this the case, then adding multiple, categorically different directions (e.g. reach vs withdrawal) to grip-object stimuli may encourage predictions for reaches forward to a matching object. Similarly, instead of a reach or withdrawal, an additional object could be added so that action observation was categorical in regards to which object one predicts the actor to reach for. This could mean that a horizontally oriented object is visible on the left and an upright object is visible on the right. Therefore not only could the direction of the reach encourage perceptual prediction but so to could the type of grasp visible, and bias perception of the reach towards one object or the other.

Do intentions require a top-down signal prior to action onset?

Another reason for the differences between the experiments might be that prior intentions need to be available well before start of the action to have an effect on perception. In the verbal and auditory experiments this is exactly what happened. The intention was uttered or heard, and then the action was observed. However, in the spatial experiments any intention could only be deciphered during the course of the movement, which may have been too late to have a significant effect on the perception of the hand's final position. Therefore, had the intention of the reach been accessed prior to the beginning of movement onset then matching grips may well have resulted in more perceptual prediction than mismatching grips. When thinking about social perception in real life, action observation is a continuous flow, which is not always easily segmented. This means that on most occasions an observed action follows some previous event or action that can provide a cue to an action's intention.

If such an interpretation were true, however, it would be less clear why object presence did not lead to an increase in prediction in the occluder experiment (Experiment 11), as the catch trial question had forced participants to process the object. One potential explanation, at least for the non-occluded conditions, is that as object presence or absence was fully accessible to participants during the reach, they may not have prioritised attention towards the fact before the reach began. That is, they may have focused on the reach and only processed object presence or absence later in the trial, or even when the probe was shown, as this was closer in time to when the question would be asked. If this was the case, the target of the reach may not have been processed in time to produce an effect of intention. However, this explanation cannot explain why no effect was found in the occluded conditions where object presence had to be processed in advance of the action onset, discounting the possibility of any delay in object processing. Therefore, the lack of object effects here would suggest that top-down information concerning object presence does not modulate visual perception, at least not for non-categorically different actions used here, and that the lack of finding is not related to the prior availability of top-down information. However, there may be other reasons may have led to the lack of perceptual modulation found.

For example, one possibility is that the occluder may have been strategically used as a form of landmark or reference in order to perform the task. This may have guided visual perception and dampened any effect of object presence or absence on perceptual prediction. An alternative, but similar, explanation is that the occluder may have been perceived as the object to which the action was directed itself. This may have made any impact of object presence irrelevant to perceptual predictive processing. Both these explanations suggest that the occluded conditions may have led to the use of other mechanisms to drive perceptual prediction other than the presence or absence of the object behind the occluder. This might explain why similar levels of perceptual prediction were found in both occluded conditions, with any effect of intention regarding object presence overridden by the occluder itself. These explanations would suggest that top-down knowledge can be superseded by bottom-up cues as participants were instructed that the actions were directed not to the occluder but to the space behind it.

While such an explanation may explain the failure to observe any effect of intention in the occluder experiment, they cannot explain why object effects were found in the tactile experiments (Experiments 12 and 13). These results demonstrate that object presence can alter perceptual prediction, which could suggest that perceptual prediction is either varies within different modalities or that object effects take longer to emerge, as in the representational momentum experiment all actions ended well in advance of the object.

Summary and conclusions

The current thesis provides evidence that our perception of other people's actions does not rely solely on the passive receipt of bottom-up sensory information, but instead is directly shaped by our prior knowledge. The findings of the present experiments show that social perception is strongly influenced by top-down processes, which can result in the automatic anticipation of an observed action's future course. When saying or hearing a verbal statement, which is in line with the direction of a subsequently observed action, people judged the action as further forward in time than when the verbal statement was not in line with the action's direction. This shows that social cues, such as language, can facilitate perceptual prediction based on the intentions implied by the linguistic content, and directly affect perceptual judgments. The current experiments therefore directly support top-down theories of social perception, which emphasise the predictive nature of perception (Csibra, 2007; Kilner et al., 2007).

Moreover, the findings of the current thesis also uncovered that the influence of top-down predictive processes changes as a result of the level of bottom-up information available. When observing an object-directed action the prediction of contact reduced people's threshold for detecting tactile stimulation on their own finger. Conversely, when observing an object-directed action where contact could only be inferred, due to a reduction in bottom-up input, people's tactile sensitivity increased. These findings demonstrate that top-down predictions of touch can have different effects on tactile perception depending upon the amount of bottom-up available and that these effects are dissociable. Both effects fit with recent models of hierarchical integration of top-down and multisensory signals. While the reduction in tactile threshold when predicting touch would reflect multisensory integration of signals on the same level (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012; van Atteveldt et al., 2014), the increase in tactile sensitivity is in line with a top-down sharpening of expected stimulus representations (Friston, 2010).

Further studies need to test what specifically the effects of representational momentum reflect. Do they reflect processes happening in the gap or during motion perception, and are they better described by top-down predictive processes, or processes of multisensory integration. Do objects serve as a similar intentional cue that leads to predictions of future movement, and do such intentional cues need to be processed before the associated action is observed?

Overall, the current findings provide support for top-down theories of social perception, which suggest that we actively anticipate the actions of others. This can be seen as beneficial for understanding the actions of others and also for planning one's own actions in response. They also suggest that information from multiple sensory inputs is integrated to shape the weight of such top-down predictions. This supports recent models linking multisensory integration to predictive coding (Talsma, 2015). It is also in line with the suggestion that the integration of different cues which is the hallmark of multisensory integration, also provides potential avenues to better understand the mechanisms of social perception (Zaki, 2013). While the precise nature of this interaction between top-down predictions and multisensory integration is not clear, the findings of the current thesis provide useful clues and potential directions for future research to pursue in the search for a more complete understanding of social perception.

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in cognitive sciences*, *4*(7), 267-278.

Altieri, N. (2014). Multisensory integration, learning, and the predictive coding hypothesis. *Frontiers in psychology*, *5*.

Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology*, *106*(3), 1437-1442.

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*(7), e67916.

Amorim, M. A., Lang, W., Lindinger, G., Mayer, D., Deecke, L., & Berthoz, A. (2000). Modulation of spatial orientation processing by mental imagery instructions: A MEG study of representational momentum. *Journal of Cognitive Neuroscience*, *12*(4), 569-582.

Anelli, F., Borghi, A. M., & Nicoletti, R. (2012). Grasping the pain: motor resonance with dangerous affordances. *Consciousness and cognition*, *21*(4), 1627-1639.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and brain sciences*, *28*(02), 105-124.

Arbib, M. A., & Mundhenk, T. N. (2005). Schizophrenia and the mirror system: an essay. *Neuropsychologia*, *43*(2), 268-280.

Arabzadeh, E., Clifford, C. W., & Harris, J. A. (2008). Vision merges with touch in a purely tactile discrimination. *Psychological Science*, *19*(7), 635-641.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological monographs: General and applied*, *70*(9), 1-70.

Avenanti, A., Annella, L., Candidi, M., Urgesi, C., & Aglioti, S. M. (2013). Compensatory plasticity in the action observation network: virtual lesions of STS enhance anticipatory simulation of seen actions. *Cerebral Cortex*, *23*(3), 570-580.

Avenanti, A., Bueti, D., Galati, G., & Aglioti, S. M. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nature neuroscience*, *8*(7), 955-960.

Avenanti, A., Candidi, M., & Urgesi, C. (2013). Vicarious motor activation during action perception: beyond correlational evidence. *Frontiers in human neuroscience*, *7*.

Bach, P. (2004). *Space and function in action comprehension* (Doctoral dissertation, Ludwig-Maximilians Universität München, Germany).

Bach, P., Bayliss, A. P., & Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychonomic bulletin & review*, *18*(1), 171-176.

Bach, P., Fenton-Adams, W., & Tipper, S. P. (2014). Can't touch this: The first-person perspective provides privileged access to predictions of sensory action outcomes. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 457.

Bach, P., Knoblich, G., Gunter, T. C., Friederici, A. D., & Prinz, W. (2005). Action comprehension: deriving spatial and functional relations. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 465.

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in human neuroscience*, *8*.

Bakker, M., Kochukhova, O., & von Hofsten, C. (2011). Development of social perception: A conversation study of 6-, 12-and 36-month-old children. *Infant behavior and development*, *34*(2), 363-370.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of cognitive neuroscience*, *15*(4), 600-609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617-629.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in cognitive sciences*, *11*(7), 280-289.

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1235-1243.

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*(2), 230.

Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1325-1334.

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2391-2404.

Baus, C., Sebanz, N., de la Fuente, V., Branzi, F. M., Martin, C., & Costa, A. (2014). On predicting others' words: Electrophysiological evidence of prediction in speech production. *Cognition*, *133*(2), 395-407.

Bays, P. M., Flanagan, J. R., & Wolpert, D. M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS biology*, *4*(2), e28.

Bays, P. M., Wolpert, D. M., & Flanagan, J. R. (2005). Perception of the consequences of self-action is temporally tuned and event driven. *Current Biology*, *15*(12), 1125-1128.

Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Frontiers in human neuroscience*, *6*.

Bekkering, H., De Bruijn, E. R., Cuijpers, R. H., Newman-Norlund, R., Van Schie, H. T., & Meulenbroek, R. (2009). Joint action: Neurocognitive mechanisms supporting human interaction. *Topics in Cognitive Science*, *1*(2), 340-352.

Bertamini, M. (1993). Memory for position and dynamic representations. *Memory & Cognition*, *21*(4), 449-457.

Blakemore, S. J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. *Brain*, *128*(7), 1571-1583.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551-559.

Blakemore, S. J., Smith, J., Steel, R., Johnstone, E. C., & Frith, C. D. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychological medicine*, *30*(05), 1131-1139.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature neuroscience*, *1*(7), 635-640.

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological bulletin*, *119*(1), 111.

Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Inferring another's expectation from action: the role of peripheral sensation. *Nature neuroscience*, *8*(10), 1295-1297.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel'touch that eyes see. *Nature*, *391*(6669), 756-756.

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*(2), 124-143.

Brenner, E., & Smeets, J. B. (2000). Motion extrapolation is not responsible for the flash–lag effect. *Vision research*, *40*(13), 1645-1648.

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in human neuroscience*, *6*.

Bub, D. N., Masson, M. E., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*(1), 27-58.

Bubic, A., Von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in human neuroscience*, *4*.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European journal of neuroscience*, *13*(2), 400-404.

Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., & Aglioti, S. M. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cerebral Cortex*, *17*(11), 2553-2561.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral cortex*, *15*(8), 1243-1249.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, *16*(19), 1905-1910.

Candidi, M., Vicario, C. M., Abreu, A. M., & Aglioti, S. M. (2010). Competing mechanisms for mapping action-related categorical knowledge and observed actions. *Cerebral Cortex*, bhq033.

Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror neurons: A critical assessment. *Annual review of neuroscience*, *37*, 1-15.

Cardellicchio, P., Sinigaglia, C., & Costantini, M. (2013). Grasping affordances with the other's hand: a TMS study. *Social cognitive and affective neuroscience*, *8*(4), 455-459.

Cardini, F., Bertini, C., Serino, A., & Ladavas, E. (2012). Emotional modulation of visual remapping of touch. *Emotion*, *12*(5), 980.

Cardini, F., Tajadura-Jiménez, A., Serino, A., & Tsakiris, M. (2013). It feels like it's me: interpersonal multisensory stimulation enhances visual remapping of touch from other to self. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 630.

Carlsson, K., Petrovic, P., Skare, S., Petersson, K. M., & Ingvar, M. (2000). Tickling expectations: neural processing in anticipation of a sensory stimulus. *Journal of cognitive neuroscience*, *12*(4), 691-703.

Castiello, U. (2003). Understanding other people's actions: intention and attention. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 416.

Chan, A. W. Y., & Baker, C. I. (2015). Seeing Is Not Feeling: Posterior Parietal But Not Somatosensory Cortex Engagement During Touch Observation. *The Journal of Neuroscience*, *35*(4), 1468-1480.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*(4), 478-484.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893.

Cheung, O. S., & Bar, M. (2012). Visual prediction and perceptual expertise. *International Journal of Psychophysiology*, *83*(2), 156-163.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(03), 181-204.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

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*(2), 302-308.

Creem-Regehr, S. H., Gagnon, K. T., Geuss, M. N., & Stefanucci, J. K. (2013). Relating spatial perspective taking to the perception of other's affordances: providing a foundation for predicting the future behavior of others. *Frontiers in human neuroscience*, *7*.

Creem-Regehr, S. H., & Lee, J. N. (2005). Neural representations of graspable objects: are tools special?. *Cognitive Brain Research*, *22*(3), 457-469.

Csibra, G. (2007). Action mirroring and action understanding: an alternative account. *Sensorimotor foundations of higher cognition: Attention and performance*, *22*, 435-480.

Daum, M. M., Vuori, M. T., Prinz, W., & Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6-and 9-month-old infants. *Developmental science*, *12*(6), 854-862.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., ... & Fazio, F. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain*, *120*(10), 1763-1777.

Deneve, S., & Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *Journal of Physiology-Paris*, *98*(1), 249-258.

Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive context influences perceptual selection during binocular rivalry. *Frontiers in human neuroscience*, *5*.

Dennett, D. C. (1989). *The intentional stance*. MIT press.

Den Ouden, H. E., Kok, P., & De Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in psychology*, *3*.

Desantis, A., Roussel, C., & Waszak, F. (2014). The temporal dynamics of the perceptual consequences of action-effect prediction. *Cognition*, *132*(3), 243-250.

Dils, A. T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences*, *107*(37), 16396-16400.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental brain research*, *91*(1), 176-180.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science*, *287*(5460), 2036-2038.

Efron, B. (2013). Bayes' theorem in the 21st century. *Science*, *340*(6137), 1177-1178.

Ellis, R., Swabey, D., Bridgeman, J., May, B., Tucker, M., & Hyne, A. (2013). Bodies and other visual objects: the dialectics of reaching toward objects. *Psychological research*, *77*(1), 31-39.

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British journal of psychology*, *91*(4), 451-471.

Elsner, C., D'Ausilio, A., Gredebäck, G., Falck-Ytter, T., & Fadiga, L. (2013). The motor cortex is causally related to predictive eye movements during action observation. *Neuropsychologia*, *51*(3), 488-492.

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. *Human body perception from the inside out*, 105-131.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433.

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychological Science*, *20*(4), 438-440.

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of Neuroscience*, *22*(13), 5749-5759.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature neuroscience*, *9*(7), 878-879.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*, *4*.

Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: object-based and context-based contributions. *Progress in brain research*, *155*, 3-21.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, *17*(8), 1703-1714.

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2010). Visual–vestibular cue integration for heading perception: applications of optimal cue integration theory. *European Journal of Neuroscience*, *31*(10), 1721-1729.

Finke, R. A., Freyd, J. J., & Shyi, G. C. (1986). Implied velocity and acceleration induce transformations of visual memory. *Journal of Experimental Psychology: General*, *115*(2), 175.

Fischer, M. H., Prinz, J., & Lotz, K. (2008). Grasp cueing shows obligatory attention to action goals. *The Quarterly Journal of Experimental Psychology*, *61*(6), 860-868.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*(6950), 769-771.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, *33*(6), 575-581.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological review*, *94*(4), 427.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126.

Freyd, J. J., & Finke, R. A. (1985). A velocity effect for representational momentum. *Bulletin of the Psychonomic Society*, *23*(6), 443-446.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, *133*(4), 694.

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature Reviews Neuroscience*, *11*(2), 127-138.

Friston, K. (2011). Embodied inference: Or I think therefore I am, if I am what I think. *The implications of embodiment (Cognition and Communication)*, 89-125.

Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biological cybernetics*, *102*(3), 227-260.

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, *22*(8), 1093-1104.

Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, *159*(3), 417-458.

Gallese, V., Fogassi, L., Fadiga, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. *Common mechanisms in perception and action: attention and performance*, *19*, 247-266.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, *2*(12), 493-501.

Genschow, O., & Brass, M. (2015). The predictive chameleon: Evidence for anticipated social action. *Journal of Experimental Psychology: Human Perception and Performance,* doi:http://dx.doi.org/10.1037/xhp0000035

Gergely, G., Nádasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165-193.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic bulletin & review*, *9*(3), 558-565.

Gordon, R. M. (1992). The simulation theory: Objections and misconceptions. *Mind & Language*, *7*(1-2), 11-34.

Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *Neuroimage*, *36*, T22-T32.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, *74*(6), 1464.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and brain sciences*, *27*(03), 377-396.

Gueugneau, N., McCabe, S.I., Villalta, J.I., Grafton, S.T., & Della-Maggiore, V. (2015) Direct mapping rather than motor prediction subserves modulation of corticospinal excitability during observation of actions in real time. *Journal of Neurophysiology.* Advance online publication. doi: 10.1152/jn.00416.2014.

Haggard, P. (2006). Just seeing you makes me feel better: interpersonal enhancement of touch. *Social Neuroscience,*, *1*(2), 104-110.

Hamilton, A., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology*, *14*(6), 493-498.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in cognitive sciences*, *16*(2), 114-121.

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

Helbig, H. B., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, *174*(2), 221-228.

Helbig, H. B., Steinwender, J., Graf, M., & Kiefer, M. (2010). Action observation can prime visual object recognition. *Experimental Brain Research*, *200*(3-4), 251-258.

Helmholtz, H. (1925). Physiological Optics, Vol. III: The Perceptions of Vision (JP Southall, Trans.). *Optical Society of America, Rochester, NY USA*.

Henrichs, I., Elsner, C., Elsner, B., & Gredebäck, G. (2012). Goal salience affects infants' goal-directed gaze shifts. *Frontiers in psychology*, *3*.

Heyes, C. (2010). Where do mirror neurons come from?. *Neuroscience & Biobehavioral Reviews*, *34*(4), 575-583.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of cognitive neuroscience*, *21*(7), 1229-1243.

Hisakata, R., Terao, M., & Murakami, I. (2013). Illusory position shift induced by motion within a moving envelope during smooth-pursuit eye movements. *Journal of vision*, *13*(12), 21.

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.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Horga, G., Schatz, K. C., Abi-Dargham, A., & Peterson, B. S. (2014). Deficits in predictive coding underlie hallucinations in schizophrenia. *The Journal of Neuroscience*, *34*(24), 8072-8082.

Hosoya, T., Baccus, S. A., & Meister, M. (2005). Dynamic predictive coding by the retina. *Nature*, *436*(7047), 71-77.

Hubbard, T. L. (1990). Cognitive representation of linear motion: Possible direction and gravity effects in judged displacement. *Memory & Cognition*, *18*(3), 299-309.

Hubbard, T. L. (1995). Cognitive representation of motion: evidence for friction and gravity analogues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 241.

Hubbard, T. L. (1996). Representational momentum, centripetal force, and curvilinear impetus. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(4), 1049.

Hubbard, T. L. (1997). Target size and displacement along the axis of implied gravitational attraction: Effects of implied weight and evidence of representational gravity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(6), 1484.

Hubbard, T. L. (1998). Some effects of representational friction, target size, and memory averaging on memory for vertically moving targets. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expÃ© rimentale*, *52*(1), 44.

Hubbard, T. L. (2004). The perception of causality: Insights from Michotte's launching effect, naive impetus theory, and representational momentum. *Fechner Day*, 116-121.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, *12*(5), 822-851.

Hubbard, T. L. (2006). Bridging the gap: Possible roles and contributions of representational momentum. *Psicológica: Revista de metodología y psicología experimental*, *27*(1), 1-34.

Hubbard, T. L. (2010). Approaches to representational momentum: Theories and models. *Space and time in perception and action*, 338-365.

Hubbard, T. L. (2014). Forms of momentum across space: Representational, operational, and attentional. *Psychonomic bulletin & review*, *21*(6), 1371-1403.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, *44*(3), 211-221.

Hudson, M., & Jellema, T. (2011). Resolving ambiguous behavioral intentions by means of involuntary prioritization of gaze processing. *Emotion*, *11*(3), 681.

Hudson, M., Liu, C. H., & Jellema, T. (2009). Anticipating intentional actions: The effect of eye gaze direction on the judgment of head rotation. *Cognition*, *112*(3), 423-434.

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological bulletin*, *139*(1), 133.

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: A study of infants' visual anticipations during action observation. *Developmental psychology*, *46*(2), 446.

Huys, R., Cañal-Bruland, R., Hagemann, N., Beek, P. J., Smeeton, N. J., & Williams, A. M. (2009). Global information pickup underpins anticipation of tennis shot direction. *Journal of Motor Behavior*, *41*(2), 158-171.

Iacoboni, M. (2005). Understanding others: Imitation, language, empathy. *Perspectives on imitation: From cognitive neuroscience to social science*, *1*, 77-99.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, *7*(12), 942-951.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *science*, *286*(5449), 2526-2528.

Jacquet, P. O., Chambon, V., Borghi, A. M., & Tessari, A. (2012). Object affordances tune observers' prior expectations about tool-use behaviors. *PloS one*, *7*(6), e39629.

Jarraya, M., Amorim, M. A., & Bardy, B. G. (2005). Optical flow and viewpoint change modulate the perception and memorization of complex motion. *Perception & psychophysics*, *67*(6), 951-961.

Johnson, R. M., Burton, P. C., & Ro, T. (2006). Visually induced feelings of touch. *Brain research*, *1073*, 398-406.

Jordan, J. S., Stork, S., Knuf, L., Kerzel, D., & Müsseler, J. (2002). Action planning affects spatial localization. *Attention and performance XIX: Common mechanisms in perception and action*, 158-176.

Kalénine, S., Buxbaum, L. J., & Coslett, H. B. (2010). Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. *Brain*, awq210.

Kennett, S., Taylor-Clarke, M., & Haggard, P. (2001). Noninformative vision improves the spatial resolution of touch in humans. *Current Biology*, *11*(15), 1188-1191.

Kerzel, D. (2000). Eye movements and visible persistence explain the mislocalization of the final position of a moving target. *Vision research*, *40*(27), 3703-3715.

Kerzel, D. (2002). The locus of “memory displacement” is at least partially perceptual: Effects of velocity, expectation, friction, memory averaging, and weight. *Perception & psychophysics*, *64*(4), 680-692.

Kerzel, D. (2003). Centripetal force draws the eyes, not memory of the target, toward the center. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 458.

Kerzel, D., Jordan, J. S., & Müsseler, J. (2001). The role of perception in the mislocalization of the final position of a moving target. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 829.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L., & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron*, *42*(2), 335-346.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive processing*, *8*(3), 159-166.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*(6), 522-525.

Kimura, M., Kondo, H., Ohira, H., & Schröger, E. (2011). Unintentional temporal context–based prediction of emotional faces: an electrophysiological study. *Cerebral cortex*, bhr244.

Kimura, M., Schröger, E., & Czigler, I. (2011). Visual mismatch negativity and its importance in visual cognitive sciences. *Neuroreport*, *22*(14), 669-673.

Kimura, M., & Takeda, Y. (2015). Automatic prediction regarding the next state of a Visual object: Electrophysiological indicators of prediction match and mismatch. *Brain research*.

Klemen, J., & Chambers, C. D. (2012). Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neuroscience & Biobehavioral Reviews*, *36*(1), 111-133.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, *27*(12), 712-719.

Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science*, *12*(6), 467-472.

Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(3), 1027-1046.

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child development*, *81*(6), 1729-1738.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *The Journal of Neuroscience*, *33*(41), 16275-16284.

Kok, P., Jehee, J. F., & de Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265-270.

Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S., & Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression?. *Neuron*, *64*(6), 922-930.

Kunde, W., Lozo, L., & Neumann, R. (2011). Effect-based control of facial expressions: Evidence from action–effect compatibility. *Psychonomic bulletin & review*, *18*(4), 820-826.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and cognition*, *65*(2), 145-168.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, *54*(3), 2492-2502.

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception-London*, *28*(11), 1311-1328.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *JOSA A*, *20*(7), 1434-1448.

Liepelt, R., Von Cramon, D. Y., & Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *Neuroimage*, *43*(4), 784-792.

Lupyan, G., & Swingley, D. (2012). Self-directed speech affects visual search performance. *The Quarterly Journal of Experimental Psychology*, *65*(6).

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, *110*(35), 14196-14201.

Macdonald, R. G., & Tatler, B. W. (2013). Do as eye say: Gaze cueing and language in a real-world social interaction. *Journal of vision*, *13*(4), 6.

Macdonald, R. G., & Tatler, B. W. (2015). Referent Expressions and Gaze: Reference Type Influences Real-World Gaze Cue Utilization.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental brain research*, *211*(3-4), 547-556.

Mann, D. L., Abernethy, B., & Farrow, D. (2010). Action specificity increases anticipatory performance and the expert advantage in natural interceptive tasks. *Acta Psychologica*, *135*(1), 17-23.

Maranesi, M., Livi, A., Fogassi, L., Rizzolatti, G., & Bonini, L. (2014). Mirror Neuron Activation Prior to Action Observation in a Predictable Context. *The Journal of Neuroscience*, *34*(45), 14827-14832.

McDonald, J. J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*(6806), 906-908.

Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 491-500.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75-78.

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs language affects low-level visual perception. *Psychological Science*, *18*(11), 1007-1013.

Meyer, G. F., & Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport*, *12*(11), 2557-2560.

Morrison, I., Tipper, S. P., Fenton-Adams, W. L., & Bach, P. (2013). " Feeling" others' painful actions: the sensorimotor integration of pain and action information. *Human brain mapping*, *34*(8), 1982-1998.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current biology*, *20*(8), 750-756.

Munger, M., & Owens, T. R. (2004). Representational momentum and the flash-lag effect. *Visual Cognition*, *11*(1), 81-103.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalizations with moving stimuli: The Fröhlich effect, the flash-lag, and representational momentum. *Visual Cognition*, *9*(1-2), 120-138.

Myung, J. Y., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition*, *98*(3), 223-243.

Nagai, M., & Yagi, A. (2001). The pointedness effect on representational momentum. *Memory & Cognition*, *29*(1), 91-99.

Negri, G. A., Rumiati, R. I., Zadini, A., Ukmar, M., Mahon, B. Z., & Caramazza, A. (2007). What is the role of motor simulation in action and object recognition? Evidence from apraxia. *Cognitive Neuropsychology*, *24*(8), 795-816.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in cognitive sciences*, *6*(9), 387-393.

Nijhawan, R. (2008). Visual prediction: psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences*, *31*(02), 179-198.

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research*, *24*(2), 190-198.

Ondobaka, S., de Lange, F. P., Newman-Norlund, R. D., Wiemers, M., & Bekkering, H. (2012). Interplay between action and movement intentions during social interaction. *Psychological Science*, *23*(1), 30-35.

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nature neuroscience*, *14*(6), 775-782.

Panichello, M. F., Cheung, O. S., & Bar, M. (2012). Predictive feedback and conscious visual experience. *Frontiers in psychology*, *3*.

Parise, C., & Spence, C. (2013). Audiovisual cross-modal correspondences in the general population. *The Oxford handbook of synesthesia*, 790-815.

Parkinson, J., Springer, A., & Prinz, W. (2012). Before, during and after you disappear: Aspects of timing and dynamic updating of the real-time action simulation of human motions. *Psychological research*, *76*(4), 421-433.

Paulus, M., Hunnius, S., & Bekkering, H. (2011). Can 14-to 20-month-old children learn that a tool serves multiple purposes? A developmental study on children's action goal prediction. *Vision research*, *51*(8), 955-960.

Pazzaglia, M., Smania, N., Corato, E., & Aglioti, S. M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *The Journal of Neuroscience*, *28*(12), 3030-3041.

Penn, D. C., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annu. Rev. Psychol.*, *58*, 97-118.

Pfister, R., Dignath, D., Hommel, B., & Kunde, W. (2013). It Takes Two to Imitate Anticipation and Imitation in Social Interaction. *Psychological science*, *24*(10), 2117-2121.

Pierno, A. C., Becchio, C., Tubaldi, F., Turella, L., & Castiello, U. (2008). Motor ontology in representing gaze–object relations. *Neuroscience letters*, *430*(3), 246-251.

Pierno, A. C., Becchio, C., Wall, M. B., Smith, A. T., Turella, L., & Castiello, U. (2006). When gaze turns into grasp. *Journal of Cognitive Neuroscience*, *18*(12), 2130-2137.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, *3*(9), 741-747.

Press, C., Taylor-Clarke, M., Kennett, S., & Haggard, P. (2004). Visual enhancement of touch in spatial body representation. *Experimental Brain Research*, *154*(2), 238-245.

Prinz, W. (1997). Perception and action planning. *European journal of cognitive psychology*, *9*(2), 129-154.

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*(1), 79-87.

Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 839.

Reid, V. M., Csibra, G., Belsky, J., & Johnson, M. H. (2007). Neural correlates of the perception of goal-directed action in infants. *Acta psychologica*, *124*(1), 129-138.

Riddoch, M. J., Edwards, M. G., Humphreys, G. W., West, R., & Heafield, T. (1998). Visual affordances direct action: Neuropsychological evidence from manual interference. *Cognitive Neuropsychology*, *15*(6-8), 645-683.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in neurosciences*, *21*(5), 188-194.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental brain research*, *111*(2), 246-252.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*(9), 661-670.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*(4), 264-274.

Ro, T., Wallace, R., Hagedorn, J., Farne, A., & Pienkos, E. (2004). Visual enhancing of tactile perception in the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *16*(1), 24-30.

Roach, N. W., McGraw, P. V., & Johnston, A. (2011). Visual motion induces a forward prediction of spatial pattern. *Current Biology*, *21*(9), 740-745.

Rossetti, Y., Pisella, L., & Vighetto, A. (2003). Optic ataxia revisited. *Experimental Brain Research*, *153*(2), 171-179.

Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of Neurophysiology*, *96*(3), 1358-1369.

Saxe, R. (2006). Uniquely human social cognition. *Current opinion in neurobiology*, *16*(2), 235-239.

Schaefer, M., Heinze, H. J., & Rotte, M. (2005). Viewing touch improves tactile sensory threshold. *Neuroreport*, *16*(4), 367-370.

Schaefer, M., Xu, B., Flor, H., & Cohen, L. G. (2009). Effects of different viewing perspectives on somatosensory activations during observation of touch. *Human brain mapping*, *30*(9), 2722-2730.

Schellekens, W., van Wezel, R. J., Petridou, N., Ramsey, N. F., & Raemaekers, M. (2014). Predictive coding for motion stimuli in human early visual cortex. *Brain Structure and Function*, 1-12.

Schubotz, R. I., Wurm, M. F., Wittmann, M. K., & von Cramon, D. Y. (2014). Objects tell us what action we can expect: dissociating brain areas for retrieval and exploitation of action knowledge during action observation in fMRI. *Frontiers in psychology*, *5*.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge university press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*, *10*(2), 70-76.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353-367.

Senior, C., Ward, J., & David, A. S. (2002). Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition*, *9*(1-2), 81-92.

Serino, A., Giovagnoli, G., & Làdavas, E. (2009). I feel what you feel if you are similar to me. *PloS one*, *4*(3), e4930.

Serino, A., Pizzoferrato, F., & Làdavas, E. (2008). Viewing a face (especially one's own face) being touched enhances tactile perception on the face. *Psychological Science*, *19*(5), 434-438.

Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive neuroscience*, *5*(2), 97-118.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*(6814), 788-788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147-152.

Shams, L., Kamitani, Y., & Shimojo, S. (2004). Modulations of visual perception by sound. *The handbook of multisensory processes*, 27-33.

Shergill, S. S., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science*, *301*(5630), 187-187.

Shtyrov, Y. (2010). Automaticity and attentional control in spoken language processing: neurophysiological evidence. *The Mental Lexicon*, *5*(2), 255-276.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157-1162.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of experimental psychology*, *81*(1), 174.

Smith–McLallen, A., Johnson, B. T., Dovidio, J. F., & Pearson, A. R. (2006). Black and white: The role of color bias in implicit race bias. *Social Cognition*, *24*(1), 46-73.

Sparenberg, P., Springer, A., & Prinz, W. (2012). Predicting others' actions: evidence for a constant time delay in action simulation. *Psychological research*, *76*(1), 41-49.

Springer, A., Brandstädter, S., & Prinz, W. (2013). Dynamic simulation and static matching for action prediction: evidence from body part priming. *Cognitive science*, *37*(5), 936-952.

Springer, A., Huttenlocher, A., & Prinz, W. (2012). Language-induced modulation during the prediction of others' actions. *Psychological research*, *76*(4), 456-466.

Stadler, W., Ott, D. V., Springer, A., Schubotz, R. I., Schütz-Bosbach, S., & Prinz, W. (2012). Repetitive TMS suggests a role of the human dorsal premotor cortex in action prediction. *Frontiers in human neuroscience*, *6*.

Stadler, W., Schubotz, R. I., Von Cramon, D. Y., Springer, A., Graf, M., & Prinz, W. (2011). Predicting and memorizing observed action: differential premotor cortex involvement. *Human brain mapping*, *32*(5), 677-687.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological science*, *12*(2), 153-156.

Stapel, J. C., Hunnius, S., & Bekkering, H. (2012). Online prediction of others' actions: the contribution of the target object, action context and movement kinematics. *Psychological research*, *76*(4), 434-445.

Stefanics, G., Astikainen, P., & Czigler, I. (2014). Visual mismatch negativity (vMMN): A prediction error signal in the visual modality. *Frontiers in Human Neuroscience*, *8*, 1074.

Stein, B., London, N., Wilkinson, L., & Price, D. (1996). Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *Cognitive Neuroscience, Journal of*, *8*(6), 497-506.

Stevens, J. A., Fonlupt, P., Shiffrar, M., & Decety, J. (2000). New aspects of motion perception: selective neural encoding of apparent human movements. *Neuroreport*, *11*(1), 109-115.

Stoffregen, T. A., Gorday, K. M., Sheng, Y. Y., & Flynn, S. B. (1999). Perceiving affordances for another person's actions. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 120.

Summerfield, C., Egner, T., Mangels, J., & Hirsch, J. (2006). Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cerebral Cortex*, *16*(4), 500-508.

Symes, E., Ellis, R., & Tucker, M. (2007). Visual object affordances: Object orientation. *Acta Psychologica*, *124*(2), 238-255.

Talsma, D. (2015). Predictive Coding and Multisensory Integration: An Attentional Account of the Multisensory Mind. *Name: Frontiers in Integrative Neuroscience*, *9*, 19.

Taylor-Clarke, M., Kennett, S., & Haggard, P. (2002). Vision modulates somatosensory cortical processing. *Current Biology*, *12*(3), 233-236.

Tessari, A., Canessa, N., Ukmar, M., & Rumiati, R. I. (2007). Neuropsychological evidence for a strategic control of multiple routes in imitation. *Brain*, *130*(4), 1111-1126.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: attributed mental states influence perception. *Trends in cognitive sciences*, *14*(8), 376-382.

Thioux, M., & Keysers, C. (2015). Object visibility alters the relative contribution of ventral visual stream and mirror neuron system to goal anticipation during action observation. *NeuroImage*, *105*, 380-394.

Thornton, I., & Hayes, A. (2004). Anticipating action in complex scenes. *Visual Cognition*, *11*(2-3), 341-370.

Tipper, S. P., Lloyd, D., Shorland, B., Dancer, C., Howard, L. A., & McGlone, F. (1998). Vision influences tactile perception without proprioceptive orienting. *Neuroreport*, *9*(8), 1741-1744.

Tipper, S. P., Phillips, N., Dancer, C., Lloyd, D., Howard, L. A., & McGlone, F. (2001). Vision influences tactile perception at body sites that cannot be viewed directly. *Experimental Brain Research*, *139*(2), 160-167.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(05), 675-691.

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 80.

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*(3), 830.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual cognition*, *8*(6), 769-800.

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta psychologica*, *116*(2), 185-203.

Turella, L., Pierno, A. C., Tubaldi, F., & Castiello, U. (2009). Mirror neurons in humans: consisting or confounding evidence?. *Brain and language*, *108*(1), 10-21.

Turella, L., Tubaldi, F., Erb, M., Grodd, W., & Castiello, U. (2011). Object presence modulates activity within the somatosensory component of the action observation network. *Cerebral Cortex*, bhr140.

Umilta, 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*(1), 155-165.

Uono, S., Sato, W., & Toichi, M. (2010). Brief report: Representational momentum for dynamic facial expressions in pervasive developmental disorder. *Journal of autism and developmental disorders*, *40*(3), 371-377.

Urgesi, C., Candidi, M., Fabbro, F., Romani, M., & Aglioti, S. M. (2006). Motor facilitation during action observation: topographic mapping of the target muscle and influence of the onlooker's posture. E*uropean Journal of Neuroscience, 23(9)*, 2522-2530.

Vainio, L., Tucker, M., & Ellis, R. (2007). Precision and power grip priming by observed grasping. *Brain and Cognition*, *65*(2), 195-207.

van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: flexible use of general operations. *Neuron*, *81*(6), 1240-1253.

Van den Brink, R. L., Cohen, M. X., van der Burg, E., Talsma, D., Vissers, M. E., & Slagter, H. A. (2014). Subcortical, modality-specific pathways contribute to multisensory processing in humans. *Cerebral Cortex*, *24*(8), 2169-2177.

van Ede, F., de Lange, F. P., & Maris, E. (2014). Anticipation increases tactile stimulus processing in the ipsilateral primary somatosensory cortex. *Cerebral Cortex*, *24*(10), 2562-2571.

van Ede, F., Jensen, O., & Maris, E. (2010). Tactile expectation modulates pre-stimulus β-band oscillations in human sensorimotor cortex. *Neuroimage*, *51*(2), 867-876.

Van Elk, M., Van Schie, H. T., & Bekkering, H. (2008). Conceptual knowledge for understanding other's actions is organized primarily around action goals. *Experimental brain research*, *189*(1), 99-107.

van Elk, M., van Schie, H., & Bekkering, H. (2014). Action semantics: a unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of life reviews*, *11*(2), 220-250.

Vannuscorps, G., Andres, M., & Pillon, A. (2013). When does action comprehension need motor involvement? Evidence from upper limb aplasia. *Cognitive neuropsychology*, *30*(4), 253-283.

Vinson, N., & Reed, C. L. (2002). Sources of object-specific effects in representational momentum.

Voisin, J. I., Marcoux, L. A., Canizales, D. L., Mercier, C., & Jackson, P. L. (2011). I am touched by your pain: limb-specific modulation of the cortical response to a tactile stimulation during pain observation. *The Journal of Pain*, *12*(11), 1182-1189.

Vroomen, J., & Gelder, B. D. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of experimental psychology: Human perception and performance*, *26*(5), 1583.

Wills, A. J., Lavric, A., Croft, G. S., & Hodgson, T. L. (2007). Predictive learning, prediction errors, and attention: Evidence from event-related potentials and eye tracking. *Journal of Cognitive Neuroscience*, *19*(5), 843-854.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological bulletin*, *131*(3), 460.

Wilson, M., Lancaster, J., & Emmorey, K. (2010). Representational momentum for the human body: Awkwardness matters, experience does not. *Cognition*, *116*(2), 242-250.

Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, *83*(2), 132-143.

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. Series B: Biological Sciences*, *358*(1431), 593-602.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729-R732.

Woodward, A., & Cannon, E. N. (2013). 16 Online Action Analysis: Infants' Anticipation of Others' Intentional Actions.

Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of vision*, *8*(3), 24.

Yoon, E. Y., & Humphreys, G. W. (2005). Direct and indirect effects of action on object classification. *Memory & cognition*, *33*(7), 1131-1146.

Yoshikawa, S., & Sato, W. (2008). Dynamic facial expressions of emotion induce representational momentum. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 25-31.

Zaki, J. (2013). Cue integration a common framework for social cognition and physical perception. *Perspectives on Psychological Science*, *8*(3), 296-312.

Zmigrod, S., & Hommel, B. (2013). Feature integration across multimodal perception and action: a review. *Multisensory research*, *26*(1-2), 143-157.