To what extent does object knowledge bias the perception of goal-directed actions?

C. Larkman

A thesis submitted for the degree of Doctor of Philosophy

Department of Psychology

University of Essex

October 2022

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Gethin Hughes, for his continuous support and patience throughout my post-graduate studies. I am also grateful to the Economic and Social Research Council, who provided the funding to undertake this research.

I would like to thank my mum, Kathleen Larkman, and my siblings, Rebecca Márta, Maria Fernando, and Robert Larkman, for their love and encouragement throughout my academic journey.

I would also like to express my appreciation for the people who have given so generously of their time over the past year so that I could complete my thesis. In particular, Kay and David Forbes, and my sister, Rebecca Márta, who provided many hours of free childcare, without which I most certainly would not have been able to complete this thesis.

I must also thank my friend Peter Gooding for being an endless source of humour throughout my PhD, particularly during the more difficult periods.

Finally, I am eternally grateful for my son, Dexter, who gave me the motivation to succeed and has made me feel more fulfilled than I could have ever imagined. I cannot say he has made this experience any easier, but he has certainly made it worthwhile.

Abstract

Predictive processing accounts of action understanding suggest that inferred goals generate top-down predictions that bias perception towards expected goals. These predictions are thought to be derived, in part, from the affordances of available objects. This thesis had three aims: (1) to test whether high-level action goals based on object knowledge can bias action perception, (2) to investigate the degree to which this perceptual bias can be influenced by high-level person knowledge, or by expertise in particular objects, (3) to explore the low-level mechanisms underlying the anticipatory representation of action goals associated with objects. Experiments used a modified representational momentum paradigm, as well as RT-based measures. In Chapter 2, we found that the presentation of a prime object led to a predictive bias in the perception of a subsequent action towards a functionally related target object. This bias was present for reaching actions, but not withdrawing actions (Experiment 1a) and persisted even when the functionally related target was simultaneously presented with an unrelated distractor (Experiment 1b). Crucially, this effect was specific to intentional actions, but was eliminated when the hand was replaced by a non-biological object following the same trajectory. This finding supports predictive processing views that action perception is guided by goal predictions, based on prior knowledge about the context in which the action occurs. We found no evidence that this perceptual bias could be influenced by prior knowledge about the gender of the actor (Chapter 3) or by participants' expertise in particular objects (Chapter 4). Chapter 5 tested for motor biases resulting from object-based goal predictions. Originally designed as a TMS study (Experiment 4a), this was tested online using RT measures as an index for motor preparation. We found no evidence that object affordances can be reliably measured using online RTs. Taken together these findings highlight the important role of object knowledge in action perception, while showing the limits to which this might be modulated by person knowledge and expertise. The final chapter highlights the challenges of developing robust behavioural measures for online testing of object affordances.

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

Understanding others’ intentions lies at the heart of social interaction. Understanding others’ goals allows us to infer their mental states, predict what they will do next and coordinate our own actions with theirs (Bach et al., 2014). For instance, in order to help a young child get something that is out of reach, one needs to understand what they are trying to achieve. They may be thirsty and reaching for the beaker, or excited and reaching for the toy bear. While much research has explored the role of motor simulation in action understanding (Iacoboni, 2009; Rizzolatti et al., 2001; Rizzolatti & Sinigaglia, 2010), there has been less focus on other contextual factors that might drive this process. In other words, while actions themselves may provide some clue as to the goal of a particular behaviour, this is normally supplemented by knowledge of the actors, the objects in play, and the situation in which the scene is unfolding. The first section of this chapter will focus on previous research that has begun to investigate how these contextual factors feed into action understanding. It then draws focus to object-based views of action understanding, exploring the role object knowledge might play in inferring the goals of others' actions.

1.1 Mirror neurons and action understanding: a bottom-up approach

Since the discovery of mirror neurons in the premotor and parietal cortices of the macaque, it has been argued that the capacity to understand others’ intentions relies on a process whereby the motor system automatically activates when observing the actions of others. Observed actions are matched to an action in the observer's own motor repertoire. These low-level motor activations then propagate, in a bottom-up manner, to higher levels in the observer's motor hierarchy, allowing for inferences regarding the likely goal of an action (Rizzolatti & Sinigaglia, 2010). For example, when observing somebody reaching to pick up a cup, our brain maps the observed action onto our own motor system, activating the associated goals or intentions we had when previously performing that same action ourselves (e.g., drinking) (Halász & Cunnington, 2012).

Bottom-up accounts of action understanding are largely based on the finding that neural circuitry involved in action observation overlaps extensively with regions that are important for action execution. Direct evidence of mirrors neurons in humans is limited due to the ethical implications of performing single cell recordings in humans (Mukamel et al., 2010). However, research using indirect imaging techniques has demonstrated that observing an action automatically activates premotor and parietal areas of the brain (Buccino et al., 2001; Chong et al., 2008; Gazzola & Keysers, 2009; Kilner et al., 2009). More recently, evidence from multivoxel pattern analysis has shown that common patterns of neural activation emerge from visual and motor brain regions (Oosterhof et al., 2010, 2012).

Although there is little doubt that observing an action activates similar processes and neuronal regions as those required for action execution (for a review, see Rizzolatti et al., 2014), more recent studies have linked action understanding to regions outside of the classical parietal-premotor mirror system, including the superior temporal sulcus (Vander Wyk et al., 2009), the temporo-parietal junction (C. D. Frith & Frith, 2006) and the rostral anterior cingulate cortex (Grézes et al., 2004).

For example, Mukamel and colleagues reported that neurons in the human medial temporal lobe, including the hippocampus, fired both during the execution and observation of similar actions, suggesting the existence of multiple systems in the human brain with neural mechanisms of mirroring (Mukamel et al., 2010).

Others suggest that mirroring may be only one of several action understanding mechanisms (Brass et al., 2007; Rizzolatti & Sinigaglia, 2010). For example, Brass and colleagues found that the mirror network only played a role in situations where no active inferential processing was required to identify the goal of the observed behaviour (Brass et al., 2007). When identifying the goal of a familiar action observed in its typical context, it is easy to map the observed action onto a corresponding motor representation already present in the observer's action repertoire. In contrast, inferring the purpose of an unusual action (e.g. turning on a light switch with one's knee) necessitates a higher level of inferential processing in order to evaluate the efficiency of the action in relation to its situational constraints (e.g., Does the person switching on the light have their hands free?). In this case, Brass et al. found reliable activations along the STS and TPJ; areas previously associated with perception of social stimuli, mentalising and action understanding, but no differential activation in the mirror network. Thus, the explanatory value of simulation theory appears to be limited to familiar actions performed in typical contexts.

1.2 Limits of bottom-up accounts of action understanding

A major problem with bottom-up theories of action understanding is that they fail to consider the ambiguity of human actions. Even in non-social visual perception, where the mapping of meaning to stimulus is much simpler, it is difficult to reliably extract low-level features of natural images, since the same object can generate different images on the retina depending on one's viewpoint, and very different objects can generate identical retinal images (Bach & Schenke, 2017; for a review, see Yuille & Kersten, 2006). Take, for example, an image of a black and white dalmatian dog (see Figure 1.1). Here, low-level cues are not sufficient to activate a high-level *'dog'* model, so naïve participants would take longer to detect the dog. However, participants who have seen this image before and thus have prior knowledge that it contains a dog, will see the dog instantly (Gregory, 2005; Yuille & Kersten, 2006).

Likewise, in social perception, there is no one-to-one mapping between actions and their goals. Most actions can be performed in various circumstances to achieve a variety of goals, and different behaviours can achieve the same goal (Bach & Schenke, 2017; Jacob & Jeannerod, 2005; Kilner et al., 2007). For example, in order to simulate the act of lifting a cup to drink from it, one would need to have prior knowledge that this is the actor's intention, as opposed to moving the cup to clean the table. Indeed, Iacoboni et al. (2005) found increased activation of the mirror neuron system when participants observed reaching actions that were embedded in a context, compared with actions that had no context. Thus, action simulation alone is not sufficient to unambiguously identify the goals of complex human actions.

Figure 1.1.

Dalmatian dog illusion, taken from Gregory. R., (2005).

The Medawar Lecture 2001. Knowledge for vision: vision for knowledge. Phil Trans Biol Sci. 360, 1231 – 1251. doi:10.1098/rstb.2005.1662

1.3 Predictive processing accounts of action understanding

Other theories argue that action understanding is not the result of motor simulation, but rather a precursor to it (Csibra, 2008; Kilner et al., 2009); We first use non-motoric contextual cues, such as language (Hudson, Nicholson, Ellis, et al., 2016; Hudson, Bach, et al., 2018; Todd et al., 2011), gaze and emotional expression (Bayliss et al., 2007), or available objects (Bach et al., 2014), to generate a prediction about the actor's intentions and goals, which is then integrated with incoming sensory information. When the sensory inputs mismatch expectation, prediction errors occur, which propagate back up the system where predictions are revised to better reflect reality (Westra, 2019). These predictive mechanisms help us to anticipate how our environment will change and to adjust our behaviour accordingly.

Studies on action prediction suggest that three factors underlie action predictions: situational context (Iacoboni et al., 2005; Rizzolatti & Craighero, 2004), goal objects (Hunnius & Bekkering, 2010; Nicholson et al., 2017; van Elk et al., 2009) and movement kinematics (Sartori, Betti, et al., 2013). One proposal is that STS, premotor and parietal areas are arranged hierarchically and form an internal generative model that predicts action patterns (at the lowest hierarchical level) as well as understanding action goals (at the higher hierarchical level) (Donnarumma et al., 2017). These hierarchical processes interact continuously through reciprocal top-down and bottom-up exchanges between hierarchical levels, so that action understanding can be variously influenced by action dynamics as well as various forms of prior knowledge; such as the context in which the action occurs (Friston et al., 2011, Kilner et al., 2007). Thus, in predictive processing models, the function of the simulation is not to determine the goal of an action, but to allow for rapid assessment of the degree to which the assumed goal matches the behaviour of the actor. The advantage of these predictive mechanisms is that they allow for minimal processing when the perceptual input matches our predictions, and for ambiguous input to be perceived in light of these expectations. In contrast, cognitive resources are devoted to unexpected actions, which generate a prediction error that is communicated back up the hierarchy, triggering a revision of our prior assumptions until they can better account for the observed behaviour (Hudson, Bach, et al., 2018; Kilner et al., 2009). Taken together, these theories highlight the importance of motor prediction and the covert reuse of our own motor repertoire in understanding others' actions.

1.4 Predictive processing in non-social perception

Predictive processing models of perception theories argue that perception in general, and social perception in particular, is a process of hypothesis testing and revision, whereby predictions about the most likely cause of sensory stimulation are subsequently compared against actual sensory input. A mismatch between the two elicits a predictions error, updating higher-level predictions.

Predictive processing models can explain distortions in perception, like perceptual illusions. Take, for example, the famous ‘Checker Shadow Illusion’ (Adelson, 2005). This visual illusion depicts a checkerboard with light and dark squares, partly shadowed by another object. The optical illusion is that within the object’s shadow, square ‘A’ appears to be a darker shade of grey than square ‘B’, despite being of identical brightness (i.e., they would be printed with identical mixtures of ink, or displayed on a screen with pixels of identical colour). Here, predictive processing models explain how we perceive the “true” colour of a surface, based not only on the bottom-up sensory information, but also our prior expectations of the light source and the way our brain understands contrast. Instead of immediately taking in the greys of A and B and immediately seeing the colour, our visual system takes into account the entire picture in deciding which grey is darker. We perceive that the image is lit from a light source coming from the right. This light source casts a shadow on the checkerboard, supposedly making all of the squares that lie in the shadow's path darker. However, when the cylinder and the rest of the checkerboard is covered (i.e., we remove context), the two squares appear identical.

Similarly, predictive mechanisms can explain how during bistable perception, observers experience fluctuations between two mutually exclusive interpretations of a constant ambiguous input as the brain tests different hypotheses (Weilnhammer et al., 2017).

1.5 Predictive processing in social perception

Top-down theories of action understanding suggest that social perception is predictive and that prior knowledge of others' intentions will shape our perception of their actions. Consequently, when an action is ambiguous, our perception will be biased towards our prediction. This perceptual bias has been demonstrated using a modified representational momentum paradigm (Hudson, Bach, et al., 2018; Hudson, McDonough et al., 2018; Hudson, Nicholson, Ellis et al., 2016; Hudson, Nicholson, Simpson et al., 2016; McDonough, et al., 2019). In a typical representational momentum task, participants see a moving object, which suddenly disappears. They then judge whether a subsequently presented probe photograph is displaced slightly forward or backward in time from the object's last seen position. Forward displacements, which are in line with one's predictions, are perceived as being identical to the object's last position, whereas backward displacements illicit larger prediction errors and are more readily detected (Hubbard, 2005). In the same way that prior knowledge of an object's typical behaviour influences motion perception (Reed & Vinson, 1996), prior knowledge of an actor's intention may influence social perception, biasing perceptual judgements of actions towards their expected goals.

Using this paradigm, Hudson and colleagues (2016) have demonstrated that top-down expectations of others' intentions are directly transformed into perceptual action predictions. Participants observed an actor reach for or withdraw from an object. The action disappeared midway through its trajectory and participants estimated its disappearance point relative to a probe stimulus. Prior to action onset, participants heard the actor state either “I’ll take it” or “I’ll leave it”. Not only was the perceived final position of the hand shifted further along the observed trajectory than it really was (the classic representational momentum effect; Freyd & Finke, 1984), but reaches were perceived as closer to an object than they really were when participants believed that the actor wanted to pick it up, and further away when they assumed a withdrawal (Hudson, Nicholson, Ellis, et al., 2016). This suggests that the perception of others’ behaviour is biased by our prior expectations of what they will do next. Further studies extend these findings, showing similar effects when participants generate expectations on the basis of object type, saying "take it!" or "leave it!" when presented with a 'safe' or 'painful' object prior to action onset (Hudson, Nicholson, Simpson, et al., 2016). Furthermore, the extent to which goal attribution drives action prediction is directly linked to the likelihood that the actor will do as they say (Hudson, Bach, et al., 2018).

Taken together, these findings suggest that perceptual systems might play a pivotal role in understanding others’ actions by not only passively representing the low-level perceptual input (bottom-up), but actively predicting what will be perceived next (top-down), based on inferences about the observed actor's goals.

1.6 Object affordances and action understanding

It has been suggested that object knowledge, including how objects are used and what they are used for, is a primary contributor to how goals are inferred and actions are interpreted (Bach et al., 2014; Nicholson et al., 2017; Schubotz et al., 2014). Gibson's (1979) influential theory of affordances suggests a tight link between perception and action whereby people not only perceive the overt physical properties of objects, but also the possible actions they can afford, based on the motor capabilities of the person.

The term *affordances* was initially used to describe what the environment *affords* the individual (e.g., a hammer affords pounding). Prior to Gibson's theory of affordances, dominant information processing perspectives suggested that humans process information about their environment systematically to build accurate mental representations of the external world. These mental representations could be based on past experience or knowledge (Cutting, 1982). In contrast, Gibson's (1979) ecological approach suggested that perception did not require mental representations. Rather, he suggested a tight link between perception and action whereby people not only perceive the physical properties of objects, but also the possible actions they can afford, based on the context and motor capabilities of the individual. For example, a hammer affords pounding for an adult with the motor capacity to lift it, but not for an infant who lacks the required strength. Such affordances are relationships that exist naturally that do not require pre-existing knowledge and are perceived in a *direct*, immediate way with no sensory processing.

Subsequent interpretations of the affordance concept have described objects as possessing properties that afford actions associated with their use. These affordances lead to the automatic activation of mental representations, regardless of context (Ellis & Tucker, 2000; Kourtis et al., 2018; Phillips & Ward, 2002). Here, *properties* refers to the manipulable aspects of an object that are associated with its typical use. For example, the handle of a mug for drinking (Chong & Proctor, 2020). Ellis and Tucker (2000) proposed the term "micro-affordances" to describe finer-grained possibilities for action than those described by Gibson. While an object may afford grasping, a grasping action can be performed in a number of possible ways, many of which would not be appropriate (Costantini & Stapleton, 2016). Micro-affordances are the specific action components suitable for interacting with specific objects. Here, it is not grasping in general that is facilitated, but a specific grasp appropriate for the viewed object (e.g., a whole-hand grip for grasping a tennis ball). In comparison to Gibson's view, recognising an object is necessary to activate micro-affordances. Furthermore, micro-affordances would be represented in the brain such that the representation of a visual object includes not only a description of its visual properties, but also encodings of all the motor patterns associated with its use (Ellis & Tucker, 2000).

Action goals can be conceptualised in a hierarchical manner such that higher level goals (e.g. turning on a light) depend first on lower level motor goals (e.g. pushing a switch) (Nicholson et al., 2017). As the goal being considered becomes increasingly abstract, the inference becomes less constrained and requires more information than is conveyed by movement kinematics (Chambon et al., 2011). It is here that objects can make a contribution to action understanding, since humans represent each object both in terms of the goal they can achieve with it ('function knowledge') and the actions one has to perform to achieve that goal ('manipulation knowledge') (Bach et al., 2014; Binkofski & Buxbaum, 2013; Buxbaum & Saffran, 2002; Collette et al., 2016; van Elk, van Schie, & Bekkering, 2009). For example, the function of a key is to open or close a door (function knowledge) and this is achieved by holding the key between the thumb and index finger, then inserting into the door lock and turning it (manipulation knowledge). Such object affordances- action knowledge carried by objects, including how they are used and what they are for- are argued to be a primary contributor to how human actions are interpreted, allowing observers to not only infer the goal someone wants to achieve with an object (via function knowledge), but also to predict the actions they would need to carry out to achieve this goal (manipulation knowledge), based on prior knowledge of how the object is used (Bach et al., 2014; Nicholson et al., 2017).

The relationship between object function and manipulation knowledge in the brain has been highlighted by studies of patients with left hemisphere brain damage. Buxbaum, Veramontil and Schwartz (2000) reported two patients with left brain damage who had a preserved capacity to retrieve object function knowledge with an impaired ability to retrieve manipulation knowledge. A subsequent study reported a double dissociation between object function and manipulation knowledge in left-brain damaged patients with and without limb apraxia. Specifically, apraxic patients were found to have preserved function knowledge and impaired manipulation knowledge, whereas nonapraxics were relatively impaired in function knowledge but had preserved manipulation knowledge (Buxbaum & Saffran, 2002).

Neuroimaging studies show further support for the segregation of function and manipulation knowledge in the brain, showing stronger activation in the left posterior parietal cortex for the retrieval of object manipulation, compared to object function, knowledge (PET: Boronat et al., 2005; fMRI: Kellenbach et al., 2003). Canessa and colleagues (2008) found a dissociation between function and manipulation knowledge in the brain, with activation of the left frontoparietal region comprising the intraparietal sulcus, the inferior parietal lobule and the dorsal premotor cortex for manipulation knowledge relative to function knowledge, and activation of the anterior inferotemporal cortex for function knowledge relative to manipulation knowledge. Taken together, these results indicate that knowledge about object function and manipulation should be considered separately.

It has been suggested that object function knowledge could make a major contribution to action goal identification over and above motor information (Bach et al., 2014; Nicholson et al., 2017). Using fMRI, Nicholson, et al. (2017) conducted a direct test of this object-based view of action understanding. Participants watched everyday instrumental actions (e.g., posting a letter) while attending to either the movements performed, the objects used, or the actions’ goal, while visual stimulation was kept identical. Their results not only confirmed a unique role of the left inferior frontal gyrus, middle temporal gyrus and medial frontal gyrus in action understanding, but also showed that activation in the goal-identification task overlapped directly with object- but not movement-related activation; namely, activation in left prefrontal and middle temporal regions, regions previously implicated in encoding object semantics (Buxbaum & Saffran, 2002). This suggests that goal understanding draws on the same left prefrontal temporal networks as object identification. Furthermore, they found that the classical mirror regions located in the premotor cortex and parietal lobe were less activated in the goal-identification task compared to the movement task. Thus, while the goal task activated regions involved in object identification, it did not activate the motor regions engaged by the movement task. Taken together, these findings suggest objects, rather than movements, provide the key information about the goals of others' actions.

1.7 Measuring object affordances

In object-based views of action understanding, motor simulation comes about by virtue of knowledge about the way in which objects allow us to achieve particular goals. For instance, an image of a cup might activate the motor chain involved in lifting the object towards the mouth by virtue of our knowledge that cups are used for drinking (Bach et al., 2014; Schubotz et al., 2014). It has already been widely demonstrated behaviourally that there is a direct link between object perception and motor performance. When we see an object, our motor system prepares for the action that object affords and facilitates object-compatible actions (Grèzes & Decety, 2002; McBride et al., 2012; Phillips & Ward, 2002; Tucker & Ellis, 1998, 2004). Early evidence of affordance effects stemmed from Tucker and Ellis's (1998) research using a stimulus-response compatibility (SRC) paradigm. A typical SRC task involves testing whether behavioural responses are faster when the spatial location of the target stimulus is compatible, relative to incompatible, with the required response (e.g., a left-side stimulus is compatible with a left, but not a right, hand response, and is therefore responded to faster; Hommel, 1997). For example, in their seminal study, Tucker & Ellis (1998) found that the handle orientation of a saucepan, though irrelevant to the task, could influence participants to give faster responses when the response hand and the handle orientation were matched. In further experiments, they had participants view objects that would typically be grasped using either a precision grip (PG) or a whole-hand grip (WHG), while being asked to respond to an orthogonal stimulus (a high or low auditory tone) using an apparatus that mimicked a PG or WHG. . Consistent with their previous studies, they found that the task-irrelevant objects influenced participants’ responses such that PG responses were faster on trials displaying objects compatible with a PG than those compatible with a WHG, and WHG responses were faster on trials displaying objects compatible with a WHG than those compatible with a PG(Ellis & Tucker, 2000). These results are commonly interpreted to show that merely perceiving an object leads to the automatic planning of the movements afforded by that object (Bub et al., 2008; Tipper et al., 2006).

Others, however, suggest that SRC tasks do not reflect affordance effects, but are instead due to a simple spatial correspondence effect, whereby the location of the stimulus and response correspond (Cho & Proctor, 2011; Phillips & Ward, 2002). Replicating Tucker and Ellis' (1998) results, Phillips and Ward (2002) found that the left or rightward handle orientation of a saucepan influenced participants to give faster responses when the response hand and handle orientation were matched. However, whether the handle faced toward or away from the participants had no influence on the compatibility effect, as would be expected if compatibility effects were due to action potentiation of the most afforded hand. Furthermore, when participants performed the same task with their hands crossed such that the left key was pressed using the right hand and the right key with the left hand, the spatial compatibility effect was even larger. Compatibility effects have also been obtained with left and right pedal responses using the feet (Phillips & Ward, 2002; Symes et al., 2005). Others suggest that this bias is produced by the asymmetry of the object, which renders the handle more salient than other parts of the object, thus capturing attention (Anderson et al., 2002; Matheson et al., 2014).

Nevertheless, supporters of the affordance account have argued that if attending to a grasp-related property of the stimulus is necessary to evoke the corresponding grasp response, then object-based spatial compatibility effects should be found for other judgements that involve an action-related property, such as the object's shape, but not for judgements that involve properties unrelated to action, such as the object's colour (Tipper et al., 2006; Tucker & Ellis, 1998). Accordingly, Tipper, Paul and Hayes (2006) had participants make key-presses in response to the shape or colour of a door handle, which was oriented to the left or right. They observed spatial compatibility effects for judgements of object shape, but not object colour. Similar effects have also been observed for judgements of texture (Loach et al., 2008). These behavioural results, along with many others (Borghi et al., 2007; I. Chong & Proctor, 2020; Girardi et al., 2010; Riggio et al., 2008; Tucker & Ellis, 2001) suggest that when we see an object, our motor system prepares for the action that object affords, independent of intentions, giving a reaction time advantage to the congruent motor response. However, the activation of affordances is not necessarily automatic, but occurs when the task requires attending to the action-related features of the object.

Neurophysiological evidence shows that observing objects activates possible actions to perform with them (Chao & Martin, 2000; Grafton et al., 1997; Grèzes at al., 2003; Valyear et al., 2007). In both humans and monkeys, neural circuits in the parietal and premotor cortex are devoted to coding the pragmatic features of objects

(Jeannerod et al, 1995). In the monkey intraparietal sulcus (AIP), neurons are endowed with visual and motor properties such as the shape, size and orientation of an observed object (Sakata et al., 1995). The AIP is strongly connected with area F5 in the premotor cortex, where canonical neurons have been discovered that are activated not only when the monkey executes a specific grip (e.g. a precision grip), but also when merely observing an object which requires such a grip (a small object). The interpretation of the discharge of canonical neurons is that these cells encode a potential motor act congruent with the properties of the presented object, independently of whether the act with will be executed or not (Murata et al., 1997). It has since been hypothesised that AIP and F5 interact in affording the most suitable motor program for acting upon an object (Buccino et al., 2009; Jeannerod, 1995).

In humans, an early PET study by Grafton, Fadiga, Arbib & Rizzolatti (1997) demonstrated that the observation of common everyday objects leads to activation of the left premotor cortex, thus suggesting the recruitment of the motor system during object observation in the absence of any motor output. More recently, research has shown that passive observation of objects, with no intention to interact with them, elicits activation in frontal and parietal regions associated with motor processes (Chao & Martin, 2000; Grafton et al., 1997; Grèzes & Decety, 2002; Grèzes et al., 2003; Kourtis et al., 2018; Valyear et al., 2007, 2012). This automaticity is akin to the motor affordances described in the behavioural literature. Indeed, Grèzes and colleagues (2003) found a strong correlation between the size of the affordance effect demonstrated by Tucker and Ellis (2001) and neural activity in anterior parietal, dorsal premotor and inferior frontal cortex. These findings suggest that passive viewing of objects not only involves the processing of its visual properties, but also evokes a representation of actions that are afforded by the specific features of the object.

1.8 Representational momentum as a measure of action prediction

Top-down theories of action understanding suggest that social perception is predictive and that prior knowledge of others' intentions will shape our perception of their actions. Thus, when an action is ambiguous, our perception will be biased towards our prediction. In order to test this, a paradigm is required that can test visual perception during action observation and make such forward predictions measurable. Representational momentum provides such a measure due to its robust demonstration of perceptual modulation based on prior expectations (Hubbard, 2005; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016).

Memory for the final position of a previously viewed moving target is often displaced forward in the direction of target motion. This forward displacement has been referred to as *representational momentum* (Freyd & Finke, 1984). The first demonstration of representational momentum was provided by Jennifer Freyd (Freyd, 1983), who presented observers with single frozen-action photographs drawn from longer motion sequences (e.g., a person mid-jump). Observers judged whether a subsequently presented probe photograph was the same as or different from the original photograph. Observers were slower to reject probes drawn from later in the action sequence than probes drawn from earlier in the sequence. They concluded that this delay resulted from a conflict generated by an internal representation of the implied movement, making it more difficult to detect differences.

In a subsequent study, Freyd and Finke (1984) presented observers with a sequence of three static rectangles, each presented at a different angle to produce the appearance of an ongoing rotation. A fourth probe rectangle was then presented at the same angle, or slightly forward or slightly backwards from the orientation of the last rectangle. Observers judged whether the probe was at the same position as the final inducing stimulus or at a different position. Both errors and reaction times increased when judging the forward probes, suggesting that observers continued to internally represent the movement even when it had disappeared, making forward probes more difficult to distinguish from this forward-displaced mental image.

A wide range of theories and models have been proposed to account for representational momentum (Hubbard, 2010). These range from theories addressing specific low-level mechanisms, such as pursuit eye movements (Kerzel, 2000, 2001), to more general high-level mechanisms based on beliefs (Kozhevnikov & Hegarty, 2001) and mental representations (Finke et al., 1986).

Bottom-up accounts of representational momentum suggest that the mechanisms that produce forward displacement occur primarily at the perceptual or motoric level rather than at the cognitive level, with no internalisation or incorporation of physical principles (Hubbard, 2010). For example, Kerzel (2000) observed that pursuit eye movements overshot the final position of a continuously moving target and that the target remained subjectively visible 50 – 60 milliseconds after the target had objectively vanished, suggesting that the displacement results from eye movements that move the persisting image of the target in the direction of motion.

Kerzel (2002) further claimed that perceptual factors account for a large proportion of a mislocalisation that was previously thought to result from processes operating in memory. Nevertheless, such theories cannot account for displacements occurring in the direction opposite to motion (Brehaut & Tipper, 1996), displacements along the orthogonal axis of horizontally moving targets (Hubbard & Bharucha, 1988) or displacements with implied motion stimuli (Freyd and Finke, 1984), none of which elicit pursuit eye movements. Furthermore, given that anticipatory eye movements are driven by high-level expectations, eye movements would not be the primary cause of representational momentum (Hubbard, 2005; 2010).

More recent accounts of representational momentum have taken a top-down approach, focussing on high-level cognitive processes rather than low-level sensory processes (Hubbard, 2005). Representational momentum is thought to occur due to the encoding of specific contexts and semantic knowledge of the stimuli. These contexts are encoded based on prior knowledge of how objects behave in the real world (Reed & Vinson, 1996; Senior et al., 2002). Reed and Vinson (1996) investigated whether prior knowledge of an object’s typical behaviour in the real world affected representational momentum. In the first experiment, an ambiguous target object was labelled either a rocket or a steeple, and representational momentum was measured when the object moved either up, down, left or right. If prior knowledge about an object’s motion influences representational momentum, the effect should be larger for the rocket than the steeple because rockets move and steeples do not. The results showed that across all directions, the rocket elicited greater representational momentum, despite the object itself being identical in both conditions. Other experiments compared objects of different shapes and found that ‘rockets’ showed greater upwards representational momentum than weights, boxes or churches. These results suggest that conceptual knowledge about objects influences the representational momentum effect.

Subsequent experiments investigated the extent to which representational momentum depended on the object’s label or the object’s visual features (Vinson & Reed, 2002). In this study, the target object was labelled a ‘rocket’ but had either a typical or atypical appearance (e.g. a rounded top rather than a pointed top). This allowed for a test of whether the conceptual effects emerged from the label alone or an interaction of the label and the appearance of the object. They found that the atypical rocket did not produce a similar increase in representational momentum as the typical rocket, suggesting that the conceptual context alone is insufficient to produce object-specific effect in representational momentum. Taken together, these findings suggesting that the representational momentum effect can be attributed to inferred motion perception, which is modulated by high-level semantic knowledge.

This paradigm can also be applied to social perception to test whether top-down knowledge of others’ intentions is directly transformed into perceptual action predictions. Prior knowledge about the actor’s intention should be automatically integrated with the observed kinematics and bias perceptual judgements towards this goal (Hudson, Nicholson, Ellis, et al., 2016). Participants observed an actor either reach for or withdraw from an object. The action disappeared midway through its trajectory and participants estimated its disappearance point. Prior to action onset, participants heard the actor state either “I’ll take it” or “I’ll leave it”. They found that people perceived reaches further towards an object than they really were when they believed that the actor wanted to pick it up, and further away when they assumed a withdrawal. This suggests that the perception of others’ behaviour is biased by our prior expectations of what they will do. Thus, as suggested by recent predictive processing models, perceptual systems might play a key role in understanding others' actions by not only passively representing the perceptual input (bottom-up), but actively predicting what will be perceived next, from knowledge about the people and objects in the situation.

Probe judgement is the most common response method used in representational momentum tasks. However, Hubbard and Bharucha (1988) used a more direct measure of displacement that involved observers using a computer mouse to indicate the display coordinates where a target was judged to have vanished. Displacement was measured by calculating the difference between the judged vanishing point and the actual vanishing point. Results from their study converged with the idea that memory for the final position of a moving target was displaced forward in the direction of target motion. However, a limitation of representational momentum studies using cursor positioning is that the cursor appears at a random location on the screen after the target has disappeared. Thus, participants must shift their attention from the remembered target position to locate the cursor before moving it to the remembered target position (Motes et al., 2008). Touch screen technology allows us to overcome such limitations, since this attention shift would not occur.

1.9 Thesis overview

This thesis aims to provide new insight into the way in which object knowledge might contribute to understanding the actions of others. Object-based views of action understanding suggest that object knowledge directly informs action goal predictions (Bach et al., 2014). Previous research has shown that people integrate action kinematic information with the affordances of available objects to derive the likely goals of observed actions, which in turn biases the perception of the action towards the expected goal (Bach et al., 2011; Mcdonough et al., 2020). Previous research has typically focussed on the low-level affordances triggered by objects, such as grip size (Buccino et al., 2009; Chainay & Humphreys, 2002). However, while previous evidence has shown that objects prime specific actions (Ellis & Tucker, 2000; Grèzes & Decety, 2002; Tucker & Ellis, 2001, 2004) it has not been assessed whether objects also activate higher-level action goals. Prior work by Bach and colleagues suggests that action perception is influenced by higher-level social information provided directly to participants prior to action onset (Hudson, Bach, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016). We aim to build on these findings in three ways: First, testing whether high-level action goals that are based on object knowledge can influence perception in a similar way (Chapter 2). Second, assessing the degree to which this perceptual bias can be influenced by person knowledge (Chapter 3), or by expertise in particular objects (Chapter 4). Third, exploring the low-level mechanisms of these anticipatory representations, by investigating the possible involvement of the motor system in the anticipatory representation of action goals associated with objects (Chapter 5).

The experiments in Chapter 2 use a representational momentum paradigm (see section 1.8) to investigate whether objects prime high-level action goals. Participants watched videos of a hand reaching for or withdrawing from a single target object (Experiment 1a) or a target object and a distractor object (Experiment 1b). Midway through the action, the hand disappeared and participants were asked to judge the index finger’s final position on the touchscreen. Prior to action onset, we implicitly manipulated the observer’s knowledge of the actor’s intention by presenting a static image of a functionally related or functionally unrelated prime object. Predictive processing accounts suggest that action perception incorporates predictions of the action’s future course, which in turn biases action perception toward these expectations (Bach & Schenke, 2017; Kilner et al., 2007). If action perception is guided by such goal predictions, and these predictions are based on the affordances of available objects (Bach et al., 2014), we would expect the hand to be (mis)perceived as disappearing further along the observed trajectory than it really was when (a) the hand is reaching towards the target rather than withdrawing from it, and (b) the target object is functionally related to the prime.

Chapter 3 aims to further investigate whether action goal predictions are guided not only by object knowledge, but also by higher-level person knowledge. In three online experiments, we tested whether prior knowledge about the gender of an individual can bias the perception of a subsequently observed reaching action towards an object that is stereotypically associated with a particular gender, versus an object that is stereotypically associated with another gender. Utilising the same representational momentum paradigm as in Chapter 2, participants were primed with a photograph of a male or female face before viewing a video of a hand reaching towards two objects; one ' highly feminine' object (e.g., a lipstick) and one 'highly masculine' object (e.g., a wrench). Part-way through the action, the hand disappeared and, as before, participants' task was to indicate the index finger's final position on a touchscreen. We hypothesised that identifying a face as male or female would bias the perception of a subsequently observed action towards the object stereotypically associated with that gender, versus an object stereotypically associated with another gender.

In Chapter 4, we test whether action goal predictions are guided by participants' expertise in particular objects. Expert chess players, through repeated exposure to the objects and rules of chess, are able to quickly recognise individual chess pieces and their associated functions. This domain-specific knowledge enables top-down predictions regarding what might occur next within a game of chess (Bilalić et al., 2010). In an online task, using a representational momentum paradigm adapted from our previous studies, we presented skilled and novice chess players with a series of videos showing a hand reaching towards a chessboard, which was set up in a checking configuration. Part-way through the action, the hand disappeared and participants indicated the index finger’s final position on the touchscreen. We predicted that chess players' superior object knowledge would bias their perception of the observed action towards the relevant target chess piece. By contrast, novice players, would not be expected to exibit a perceptual bias as they lack the required object knowledge and would therefore have no expectations as to which object the actor might be reaching for.

Chapter 5 changed focus to explore the possible involvement of the motor system in the anticipatory representation of action goals associated with objects. Experiment 4a aimed to assess this using TMS. However, we were unable to run this study due to COVID-19 restrictions. In Experiment 4b, we adapted our design to see if this question could be addressed using reaction time measures. In this online task, participants watched a series of short videos of an actor reaching to grasp an object (the "prime") with one hand before reaching towards two other objects with their other hand; one object which was functionally related to the prime, and another which was functionally unrelated to the prime. One object required a PG and the other a WHG. At the end of each video, participants heard a high or low tone. Using the keyboard, their task was to respond with the index finger of one hand when hearing a high tone, or *both* the index and little finger of the other hand when hearing a low tone. We predicted that participants would be faster to respond when the functionally related object required a grip that matched the required response. Experiments 4c and 4d used progressively simpler designs to test the extent to which object affordances can be reliably measured using reaction times.

Chapter 2: Do objects prime action goals?[1]

Understanding others' intentions lies at the heart of social interaction. Understanding others' goals allows us to infer their mental states, predict what they will do next and coordinate our own actions with theirs (Bach et al., 2014). For instance, in order to help a young child get something that is out of reach, one needs to understand what they are trying to achieve. They may be thirsty and reaching for the beaker, or excited and reaching for the toy bear. While much research has explored the role of motor simulation in action understanding (Iacoboni, 2009; Rizzolatti et al., 2001; Rizzolatti & Sinigaglia, 2010), there has been less focus on other contextual factors that might drive this process. In other words, while actions themselves may provide some clue as to the goal of a particular behaviour, this is normally supplemented by knowledge of the actors, the objects in play, and the situation in which the scene is unfolding.

The capacity to understand others' intentions has typically been thought to rely on a process whereby the motor system automatically activates when observing the actions of others. Observed actions are matched to an action in the observer's own motor repertoire. These low-level motor activations propagate, in a bottom-up manner, to higher levels in the observer's motor hierarchy, allowing for inferences regarding the likely goal of an action (Rizzolatti & Sinigaglia, 2010). Although there is little doubt that observing an action activates similar processes and neuronal regions

---

[1] Note: Chapter 2 was written as an independent piece of research with the aim of being submitted as a peer-reviewed manuscript for publication. As such, there is an overlap in the literature reviews of Chapters 1 and 2.

as those required for action execution (Gazzola & Keysers, 2009; Kilner, Neal, Weiskopf, Friston, & Frith, 2009; Mukamel et al., 2010; Rizzolatti et al., 2014), in complex visual scenes action simulation alone might be not sufficient to unambiguously identify the goals of human actions. Most actions can be performed in various circumstances to achieve a variety of goals, and different motor behaviours can achieve the same goals (Jacob & Jeannerod, 2005).

Other theories argue that action understanding is not the result of motor simulation, but rather a precursor to it (Csibra, 2008; Kilner et al., 2009); We first use non-motoric contextual cues, such as language (Hudson, Bach et al., 2018; Hudson, Nicholson, Ellis et al., 2016; Todd et al., 2011), gaze and emotional expression (Bayliss et al., 2007), or available objects (Bach et al., 2014), to generate a prediction about the actor's intentions and goals, which is then integrated with incoming sensory information.

Again, other theories highlight the importance of motor prediction and the covert reuse of our own motor repertoire and internal models in this process. For example, one influential proposal is that STS, premotor and parietal areas are arranged hierarchically (in a so-called predictive coding architectural scheme) and form an internal generative model that predicts action patterns (at the lowest hierarchical level) as well as understanding action goals (at the higher hierarchical level) (Donnarumma et al., 2017). These hierarchical processes interact continuously through reciprocal top-down and bottom-up exchanges between hierarchical levels, so that action understanding can be variously influenced by action dynamics as well as various forms of prior knowledge; such as the context in which the action occurs (Friston et al., 2011, Kilner et al., 2007). In such predictive models, the function of motor simulation is not to determine the goal of an action, but to allow for rapid assessment of the degree to which the assumed goal matches the observed behaviour. The advantage of these predictive mechanisms is that they allow for minimal processing when the perceptual input matches our predictions, and for ambiguous input to be perceived in light of these expectations. In contrast, cognitive resources are devoted to unexpected actions, which generate a prediction error that is communicated back up the hierarchy, triggering a revision of our prior assumptions until they can better account for the observed behaviour (Hudson, Bach, et al., 2018; Kilner et al., 2009).

Top-down theories of action understanding suggest that social perception is predictive and that prior knowledge of others' intentions will shape our perception of their actions. Consequently, when an action is ambiguous, our perception will be biased towards our prediction. This perceptual bias has been demonstrated using a modified representational momentum paradigm (Hudson, Bach, et al., 2018; Hudson, McDonough et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016; McDonough et al., 2019). In a typical representational momentum task, participants see a moving object, which suddenly disappears. They then judge whether a subsequently presented probe photograph is displaced slightly forward or backward in time from the object's last seen position. Forward displacements which are in line with one's predictions are perceived as being identical to the object's last position, whereas backward displacements illicit larger prediction errors and are more readily detected (Hubbard, 2005). In the same way that prior knowledge of an object's typical behaviour influences motion perception (Reed & Vinson, 1996), prior knowledge of an actor's intention may influence social perception, biasing perceptual judgements of actions towards their expected goals.

Using this paradigm, Hudson and colleagues (2016) demonstrated that top-down expectations of others' intentions are directly transformed into perceptual action predictions. Participants observed an actor reach for or withdraw from an object. The action disappeared midway through its trajectory and participants estimated its disappearance point relative to a probe stimulus. Prior to action onset, participants heard the actor state either “I’ll take it” or “I’ll leave it”. Not only was the perceived final position of the hand shifted further along the observed trajectory than it really was (the classic representational momentum effect; (Freyd & Finke, 1984), but reaches were perceived as closer to an object than they really were when participants believed that the actor wanted to pick it up, and further away when they assumed a withdrawal (Hudson, Nicholson, Ellis, et al., 2016). This suggests that the perception of others’ behaviour is biased by our prior expectations of what they will do. Further studies extend these findings, showing similar effects when participants generate expectations on the basis of object type, saying "take it!" or "leave it!" when presented with a 'safe' or 'painful' object prior to action onset (Hudson, Nicholson, Simpson, et al., 2016). Furthermore, the extent to which goal attribution drives action prediction is directly linked to the likelihood that the actor will do as they say (Hudson, Bach, et al., 2018).

Taken together, these findings suggest that perceptual systems might play a pivotal role in understanding others’ actions by not only passively representing the low-level perceptual input (bottom-up), but actively predicting what will be perceived next (top-down), based on inferences about the observed actor's goals. However, in each of these studies, the goal of the actor is explicitly stated prior to action onset. This does not reflect real-world interactions in which people typically do not announce their intentions prior to their actions (Mcdonough et al., 2020). Rather, goals are inferred from contextual information, such as objects or social signals, and linked to their likely motor behaviours.

It has been suggested that object knowledge, including how objects are used and what they are used for, is a primary contributor to how goals are inferred and actions are interpreted (Bach et al., 2014; Nicholson et al., 2017; Schubotz et al., 2014). In object-based views of action understanding, motor simulation comes about by virtue of knowledge about the way in which objects allow us to achieve particular goals. For instance, an image of a cup might activate the motor chain involved in lifting the object towards the mouth by virtue of our knowledge that cups are used for drinking (Bach et al., 2014).

It has already been widely demonstrated that there is a direct link between object perception and motor performance. When we see an object, our motor system prepares for the action that object affords and facilitates object-compatible actions (Grèzes & Decety, 2002; McBride et al., 2012; Phillips & Ward, 2002; Tucker & Ellis, 1998, 2004). For example, people are faster to categorise a large object (e.g., a potato) when responding with a whole-hand-grip compared with a precision grip (Tucker & Ellis, 2001). They also make faster action decisions ('pour or twist?') about objects compared with contextual decisions ('found in kitchen?'), suggesting objects provide privileged access to action information (Chainay & Humphreys, 2002; van Elk et al., 2009).

Action goals can be conceptualised in a hierarchical manner such that higher level goals (e.g. turning on a light) depend first on lower level motor goals (e.g. pushing a switch) (Nicholson et al., 2017). As the goal being considered becomes increasingly abstract, the inference becomes less constrained and requires more information than is conveyed by movement kinematics (Chambon et al., 2011). It is here that objects can make a contribution to action understanding, since humans represent objects both in terms of function knowledge; the goal attainable by using an object, and manipulation knowledge; the gestures one has to execute to achieve that goal (Bach et al., 2014; Binkofski & Buxbaum, 2013; Buxbaum & Saffran, 2002; Collette et al., 2016; van Elk et al., 2009). For example, the function of a key is to open or close a door (function knowledge) and this is achieved by holding the key between the thumb and index finger, then inserting into the door lock and turning it (manipulation knowledge). Such object affordances; action knowledge carried by objects, including how they are used and what they are for, are argued to be a primary contributor to how human actions are interpreted, allowing observers to not only infer the goal someone wants to achieve with an object (via function knowledge), but also to predict the actions they would need to carry out to achieve this goal (manipulation knowledge), based on prior knowledge of how the object is used.

Several studies have demonstrated that object affordances directly inform action goal prediction (for a review see Bach et al., 2014). For example, when people see somebody else close to an object, the most effective grip required to use that object is activated, as if they were in the position of the observed actor (Cardellicchio et al., 2013; Costantini et al.,2011). Furthermore, there is evidence that people integrate observed kinematic information with the affordances of available objects to derive the likely goal of the observed action, which in turn biases their perception of the action (Bach et al., 2011; McDonough et al., 2020).

We aim to provide new insight into the way in which objects activate action goals. In other words, if objects prime actions, do they also prime action goals and their perceptual effects?

Previous studies have focussed on low-level affordances triggered by objects (e.g. grip size; Buccino et al., 2009; Chainay & Humphreys, 2002). However, in most real-life situations, grasping is just an initial component of a broader action in which the grasped object is used to achieve a subsequent goal (Majdandžić et al., 2007). Here, we tested for a bias towards a functionally related object versus a functionally unrelated object, while also controlling for low level visual features of the stimuli and in the absence of any explicit cue to intention. Importantly, we tested the robustness of this effect by, first, investigating the predictive biases for intentional actions (a reaching hand) compared to non-intentional actions (a moving ball) and, second, investigating the predictive biases not only in the horizontal plane, but also in the vertical plane between two objects.

Experiment 1a tested whether the visual presentation of an active object (e.g., a hammer) would lead to a predictive bias in the perception of an action towards a functionally associated object (e.g., a nail). We used a modified representational momentum paradigm based on Hudson et al. (Hudson, Nicholson, Ellis, et al.,2016; Hudson, Nicholson, Simpson, et al., 2016). Participants watched videos of a hand reaching for or withdrawing from an object. Part-way through the action, the hand disappeared and participants indicated the index finger’s final position on a touchscreen. Prior to action onset, we manipulated the observer’s knowledge of the actor’s intention, such that they either expected a reach towards the target object (congruent prime and target) or they did not (incongruent prime and target). We predicted that participants would misperceive the hand’s disappearance point further along the observed trajectory when (1) the hand was reaching towards the target object and (2) the target object was congruent with the prime object. This would suggest that action perception incorporates predictions of the action's future course (representational momentum): Hands are perceived to have disappeared closer to the object when the actor has the (inferred) goal to pick it up, even when the visual stimulation is identical.

In Experiment 1b, we looked to replicate our effects from Experiment 1a by investigating predictive biases in the vertical plane between two objects, while controlling for low level visual features. We also tested whether this effect would only occur when observing intentional actions (a reaching hand) compared to non-intentional actions (a moving ball).

Participants watched videos of a hand reaching, or a ball moving, towards two objects, one of which was functionally related to the prime object (e.g., hammer – nail), the other unrelated (e.g., hammer – cigarette). Objects were matched for the type of grip required. Part-way through the action, the hand/ball disappeared and participants indicated the index finger/ball's final position on a touchscreen. We predicted that the visual presentation of an active object would lead to a predictive bias in the perception of a subsequent action towards a functionally related object, rather than a functionally unrelated object. Moreover, if this bias is driven by predictions about the actor's goals, we should only observe an effect in the intentional-action condition and not in the non-intentional ball condition.

2.1 Pre-test of the stimuli

The aim of the pre-test was to assess the correspondence between pairs of objects that are assumed to be functionally related. Since we hypothesised that the visual presentation of an object will activate a representation of the associated object goal, it is paramount that we ensure our stimuli permit such activation.

2.1.1 Method

Participants

Fifty-seven right-handed volunteers (16 male, 41 female), aged 18 to 30 years (mean age: 25.16 years, SD = 3.33), were recruited from the University of Essex.

Stimuli and procedure

A set of visual object stimuli was constructed by selecting 36 colour photographs of manmade objects. The objects consisted of 18 functionally associated pairs (e.g. hammer and nail). Within each pair, one object was always the active object (e.g. screwdriver) and the other the passive object (e.g. screw). The two objects of each pair were presented side-by-side. All objects were positioned at an orientation compatible with a right-hand grasp. Images were sourced online and processed such that each image was approximately 200 x 200 pixels, presented against a white background.

Visual stimuli were presented via an online questionnaire in Qualtrics. Participants were instructed that they would be presented with pairs of common everyday objects and asked to rate how likely they were to use the two objects together, based on their typical use. Responses were given on a 7-point likert scale, ranging from extremely likely to extremely unlikely. The questionnaire comprised 36 questions; 18 pairs of functionally related objects and 18 pairs of functionally unrelated objects. Question order was randomised for each participant. Questions appeared on the screen one at a time and remained there until participants had given their rating and pressed the ‘next’ button to progress onto the next question.

2.1.2 Results

Object pair ratings were coded from 1 to 7, such that higher scores indicated that participants were more likely to use the objects together. Only ratings from completed questionnaires were included in the analysis. Mean values were calculated for each association. Ten object pairs were selected for the main experiment (mean rating = 6.72, SD = 0.25); five with small passive objects and five with large passive objects. The means and standard deviations of the likelihood ratings are presented in Table 2.1.

Table 2.1.

Mean and standard deviation likelihood ratings for each object pair in Experiments 1a and b. Half of the passive objects were small objects requiring a precision grip and half were larger objects requiring a whole-hand power grip.

Precision Grip				**Power Grip**			
Active	**Passive**	***M***	***SD***	**Active**	**Passive**	***M***	***SD***
Pencil	Sharpener	6.88	0.38	Pen	Notepad	6.98	0.13
Key	Lock	6.81	0.85	Teaspoon	Mug	6.84	0.45
Lighter	Cigarette	6.09	1.93	Tennis racket	Tennis ball	6.84	0.49
Spanner	Nut	6.67	0.87	Jug	Glass	6.77	0.54
Hammer	Nail	6.58	1.13	Can opener	Can	6.70	1.15

2.2 Experiment 1a

2.2.1 Method

Participants

Participants were 72 right-handed volunteers from the University of Essex (20 male, 52 female, mean age: 21 years, range: 18 to 29 years), all with normal or corrected-to-normal vision. Participants were given either course credit or £3 for taking part. A priori power analysis (https://jakewestfall.shinyapps.io/pangea) indicated that 60 participants were needed to have 86% power for detecting an interaction with a "medium" effect size when employing the traditional .05 criterion of statistical significance. Participants were excluded from the analysis if their accuracy in the fls was less than 75% suggesting they were not fully processing the prime or target objects. This left us with a sample size of 60.

Apparatus

Action sequences were filmed using a Sony SF7 camera (50 frames/s, 3840 x 2160 pixels) and edited using Adobe Premiere Pro CC. The experiment was administered using Inquisit on an Aser T231H 23-inch touchscreen monitor (resolution: 1920 x 1080, refresh rate: 60Hz, 2 ms response time).

Stimuli and procedure

Experimental stimuli consisted of 10 photographs of manmade objects (the prime objects) and 60 action sequences of an arm reaching for one of 10 objects (the target objects). Prime and target objects were those selected during the pre-test. The prime object was always an *active* object (e.g. screwdriver) and the target object was always a *passive* object (e.g. screw). Half of the target objects were ‘small’ objects requiring a precision grip and half were ‘large’ objects requiring a whole-hand power grip.

Images of prime objects had the same features as in the pre-test, but were resized (~ 300 x 300 pixels) and presented in the centre of the screen against a black background. Real-life target objects were selected for use in the action sequences. These were chosen to represent the target objects from the pre-test as accurately as possible.

Action sequences were derived from videos of a hand starting in a rest position then reaching for either a small (e.g. a nail) or a large (e.g. a mug) target object located on a table at a distance of ~ 58 cm from the actor’s torso. Only the actor’s right arm was visible. In half of the videos the actor performed a reach-to-grasp action using a whole-hand power grip, during which the pre-shaping of the hand was clearly visible as soon as the action started. In the other half, the actor performed a precision grip, depending on the target object. All background details were replaced with a uniform black background. Each video comprised 11 frames (each 20 ms) showing the complete transport phase, but with the first 5 frames and the final grasp omitted. Action sequences were either 5, 7 or 11 frames long. Reaches and withdrawals were created by stepping either forwards of backwards through the sequence.

Procedure

Before beginning the task, participants were instructed that when presented with a picture of an object (i.e. the prime) they should imagine using this object. For example, if presented with a picture of a teapot, they should imagine pouring a cup of tea. This was to encourage them to engage with the task.

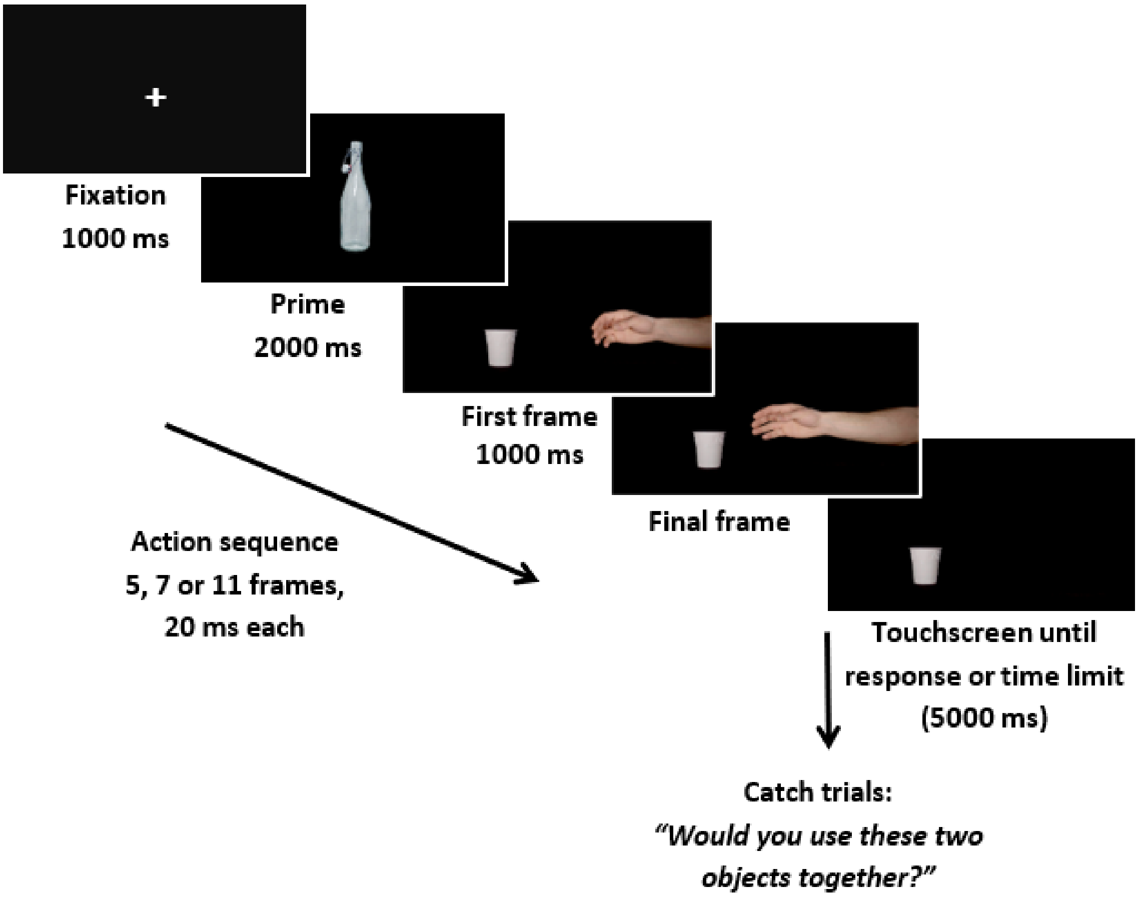
Each trial (see Figure 2.1) began with a fixation cross in the centre of the screen for 1000 ms, followed by the prime stimulus for 2000 ms. The first frame of the video was then presented as a static image, showing the hand in the starting position of the action sequence. After 1000 ms, the action sequence began. The response screen followed immediately after the video, showing the target object with the arm removed. Using their index finger, participants were asked to tap the point on the screen where they last saw the hand’s index finger. Before and after each response, participants placed their finger on a red dot at the bottom of the touchscreen monitor. The response screen remained present until a response was given, or until 5000 ms had elapsed.

Participants completed three blocks of 40 trials consisting of iterating the factors Prime Congruency (congruent, incongruent), Action Direction (reach, withdraw) and Action Sequence (5, 7 or 11 frames). Thus, each target object was presented 12 times. Trials were randomly presented within each block. For incongruent trials, target objects were matched for grip type (e.g. Lighter – cigarette/nail; Tennis racket – ball/mug).

The experiment contained four catch trials per block in which the response stimulus was a question: ‘Would you typically use these two items together?’ Participants responded by pressing either ‘Yes’ or ‘No’ on the touchscreen. The question remained on the screen until a response was given, or 6000 ms had elapsed. Catch trials were used to screen participants who did not comply with task instructions. Participants completely six practice trials to familiarise them with the experimental procedure. Practice trials were selected at random.

Figure 2.1

Trial sequence for Action condition (Experiment 1a).



Data analysis

The experiment employed a 2 x 2 repeated-measures design, with Action Direction (reach, withdraw) and Prime Congruency (congruent, incongruent) as our independent variables. Eleven-frame trials were excluded from the analysis as these limited the capacity for an RM effect to be observed due to the hand's close proximity to the object, despite being a necessary feature of the study design. Participants' selected screen coordinate of the tip of the index finger on each trial was subtracted from the real final screen coordinate. Analysis was conducted on this residual localisation error, which provided a directional measure of how far, in pixels, participant’s responses were displaced along the x-axis. An accurate response would produce a value of zero. Negative values denote a rightward displacement (away from the object) and positive values a leftward displacement (towards the object).

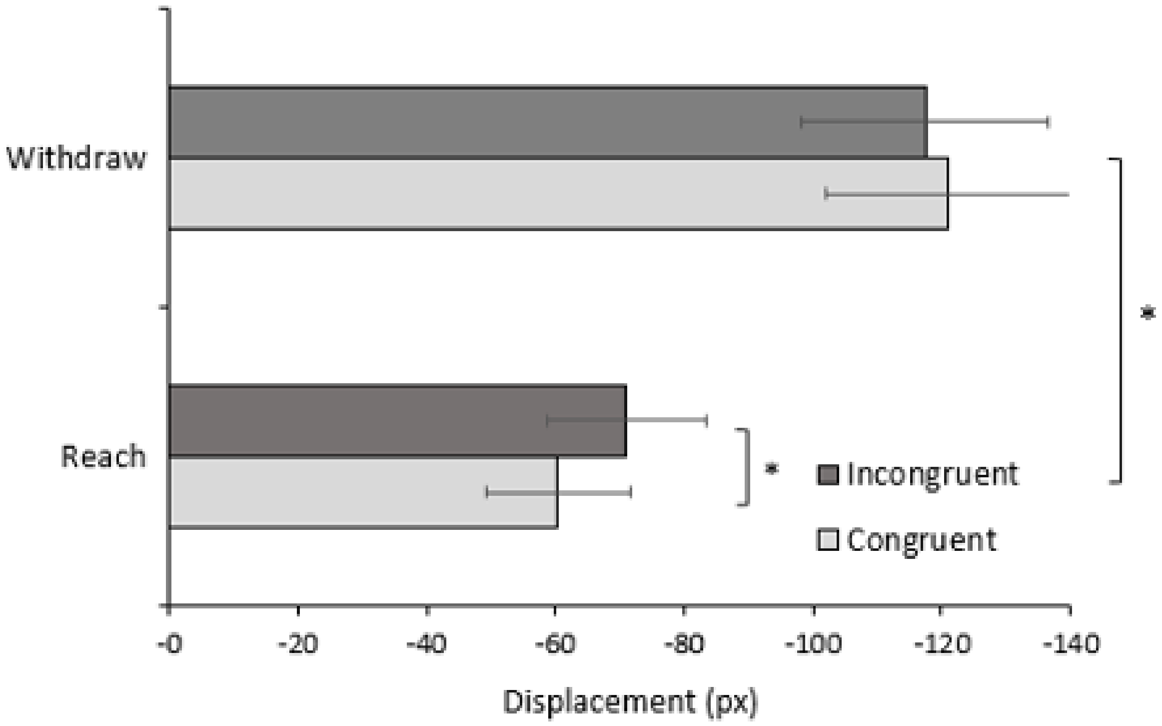
2.2.2 Results

A two-way ANOVA was conducted to examine the effect of Action Direction (reach, withdraw) and Prime Congruency (congruent, incongruent) on Displacement. This revealed a main effect of Action Direction, *F* (1, 59) = 6.36, *p* = .014, hp^2 = .097. Displacements for withdrawing actions were negative and larger than for reaching actions, meaning they were further away from the object (Figure 2.2). As seen in Figure 2.2, both reaching and withdrawing actions were associated with negative displacements, suggesting that participants had a tendency to respond towards the centre of the hand rather than the index finger, as instructed. Importantly, there was a significant interaction between the effects of Action Direction and Congruence on Displacement, *F* (1, 59) = 5.65, *p* = .021, hp^2 = .087, suggesting that the size of the displacement depended on both the direction of the observed action and the congruency of the objects. As predicted, pairwise comparisons, with a Bonferroni adjustment, showed a significant difference in displacement between congruent and incongruent trials for reaching actions (*t* (60) = 2.67, *p* = .010), but not for withdrawing actions (*p* > .05. See Figure 2.2).

Further analyses revealed that the effect was robust to the exclusion criterion and remained significant when no participants were excluded, *F* (1, 71) = 4.96, *p* = .029, hp^2 = .065.

Figure 2.2.

Experiment 1a results.



Mean displacements between the real vanishing point *x*-coordinates and participants' actual touch-response *x*-coordinates are shown for congruent and incongruent prime-target pairs, in reaching- and withdrawing-action trials (Experiment 1a). Error bars show the standard error.

2.3 Experiment 1b

In Experiment 1b, we looked to extend our findings from Experiment 1a to investigate predictive biases in the vertical plane between simultaneously presented two objects, while controlling for low level visual features. We further sought to investigate whether this effect would only occur when observing intentional actions (a reaching hand) compared to non-intentional actions (a moving ball).

In this experiment, participants watched videos of a hand reaching, or a ball moving, towards two objects, one of which was functionally related to the prime object, the other unrelated. Objects were matched for the type of grip required. Part-way through the action, the hand/ball disappeared and participants indicated the index finger/ball's final position on a touchscreen.

2.3.1 Method

Participants

Participants were 101 right-handed volunteers (42 male, 59 female; mean age: 37.56 years, range: 20 to 77 years), recruited from the United States via Amazon Mechanical Turk. Participants were given $2.50 for taking part. None had taken part in the pre-test or Experiment 1. Participants were excluded from the analysis if they did not complete all experiment trials (*N* = 6) or their accuracy in catch trials was less than 60%. This left us with a final sample of 64.

Apparatus

The experiment was administered using Inquisit Web on iPad devices running iOS, with a minimum screen resolution of 1024 x 768 pixels and a minimum refresh rate of 40Hz. This eliminated the possibility of participants using iPhones, and meant that the smallest useable device was an iPad mini with a screen size of 7.9" (a comprehensive list of iOS devices and their respective screen resolutions can be found at: https://www.ios-resolution.com).

Stimuli and procedure

Experimental stimuli consisted of 8 photographs of manmade objects (the prime objects), 24 action sequences of an arm ambiguously reaching towards two of a possible eight objects (the target objects), and 24 sequences depicting a ball moving towards the same target objects. Objects were the same as those of Experiment 1b, but excluded the notebook, which was deemed too large to be presented alongside a second target object, and the nail. New action sequences were filmed showing a hand ambiguously reaching towards two target objects (see Table 2.2); one placed towards the bottom left of the screen, the other placed towards the top left of the screen, forming either a whole-hand power grip or a precision grip. The objects were sufficiently distanced from the hand, allowing all trials to be included in the analysis. In half of the videos, the functionally-related object was orientated towards the bottom of the screen and in the other half towards the top of the screen. Each video comprised 15 frames (each 20 ms) showing the complete reaching action, but with the first 5 frames and the final grasp omitted. Action sequences were either 9, 12 or 15 frames long.

Trial timings were the same as in Experiment 1 (see Figure 2.3). For each condition (Action, Control), participants completed a practice block of 6 trials, followed by three experimental blocks each consisting of 18 trials. These 54 trials comprised 8 objects presented at each level of Target Orientation (Top, Bottom) and Action Sequence (9, 12 or 15 frames), and 6 catch tsrials (total = 120 trials). Trials were randomly presented within each block. Target object pairs were matched for grip type (e.g. cigarette/nail, ball/mug, see Table 2.2).

The experiment contained four catch trials per block in which the response stimulus was a question: ‘Which object would you use with the object in the picture?’ Participants responded by pressing the word ‘Top’ or ‘Bottom’ on the touchscreen. The question remained on the screen until a response was given, or 6000 ms had elapsed.

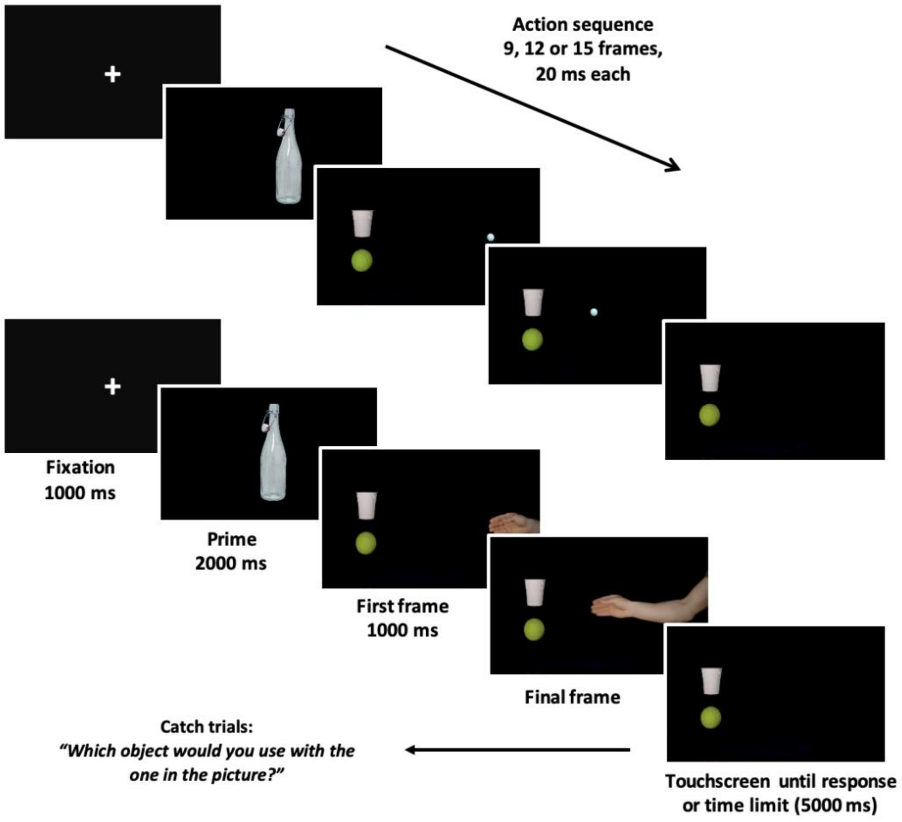
Table 2.2

Object pairs for Experiment 1b.

Precision Grip			**Power Grip**		
Active	**Passive - related**	**Passive – unrelated**	**Active**	**Passive - related**	**Passive - unrelated**
Key	Lock	Sharpener	Bottle	Cup	Ball
Pencil	Sharpener	Lock	Tennis Racket	Ball	Cup
Spanner	Bolt	Cigarette	Teaspoon	Mug	Can
Lighter	Cigarette	Bolt	Can opener	Can	Mug

Figure 2.3

Trial sequence for Action (top) and Control (bottom) conditions (Experiment 1b). Catch trials were present in both Action and Control conditions.



Data analysis

The experiment employed a 2 x 2 repeated-measures design, with Condition (Action, Control) and Target Object Location (Top, Bottom) as our independent variables. Screen coordinates were converted into percentage coordinates, where 0%, 0% represents the top left-hand corner of the screen. Participants' selected screen coordinate (in percentage coordinates) of the tip of the index finger/ball on each trial was subtracted from the real final screen coordinate. Analysis was conducted on this residual localisation error, which provided a directional measure of how far participants' responses were displaced along the X and Y axis. An accurate response would produce a value of 0% on both axes. On the X axis, negative values denote a rightward displacement (against the direction of motion) and positive values a leftward displacement. On the Y axis positive and negative values denote upward and down displacements respectively.

2.3.2 Results

Y-axis

Our main prediction was that perceptual judgments of actions would be displaced towards the expected trajectory, that is, towards the object that is functionally related to the prime object. A two-way ANOVA was conducted to examine the effect of Condition (Action, Control) and Target Object Location (Top, Bottom) on Displacement along the y-axis. The analysis revealed a main effect of Target Location (*F* (1, 63) = 8.27, *p* = .005, hp^2 = .12) and Condition (*F* (1, 63) = 54.28, *p* < .001, hp^2 = .46). Displacements in the Action condition were larger than the Control condition (See Figure 2.4).

Importantly, the interaction was also significant (*F* (1, 63) = 4.06, *p* = .048, hp^2 = .06). The hand was perceived as disappearing lower down on the screen when the target object was presented at the bottom of the screen than when it was presented at the top. As predicted, pairwise comparisons with a Bonferroni adjustment, showed that this effect was only present in the Action condition (*M* = 2.14%, *SD* = 6.12, *t* (63)

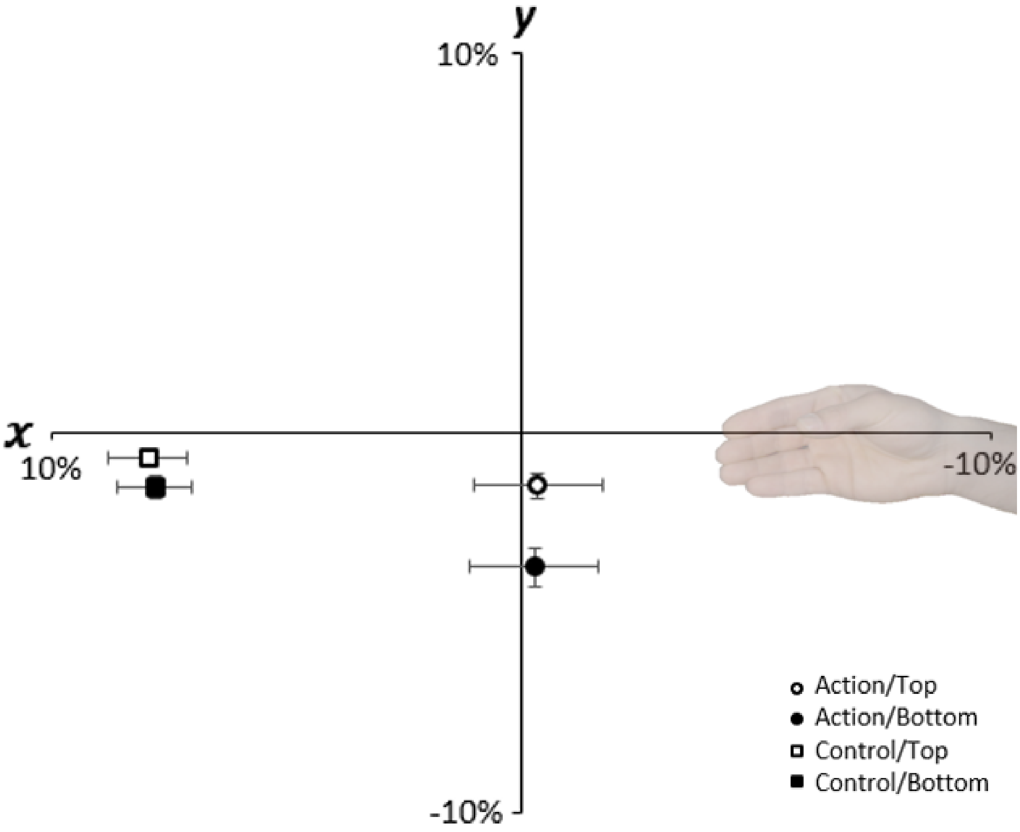
= 2.80, *p* = .007, *d* = .63) and not in the Control condition (*M* = 0.75%, *SD* = 3.17 *t* (63) = 1.89, *p* = .064, *d* = .38).

X-axis

We did not have any specific predictions about perceptual displacements on the X-axis. However, a two-way ANOVA, with Condition (Action, Control) and Target Location (Top, Bottom) entered as repeated-measures factors, found a main effect of Condition (*F* (1, 63) = 45.94, *p* < .001, hp^2 = .42), with no effect of Target Location (*F* (1, 63) = 0.10, *p* = .756, hp^2 = .00) and no interaction (*F* (1, 63) = 0.61, *p* = .439, hp^2 = .01).One-sample t-tests revealed that in the Action condition, displacements were not significantly larger than zero (*M* = -.31%, *SD* = 10.87, *t* (64) = -.23, *p* = .818, *d* = .03.). As in Experiment 1, this is likely explained by participants' tendency to respond towards the centre of the hand rather than the index finger. In the Control condition, the perceptual bias was positive and significant, showing displacement in the direction of motion (*M* = 7.88, *SD* = 6.55, *t* (64) = 9.62, *p* < .001, *d* = 1.20).

Figure 2.4.

Experiment 1b results.



The interaction between task condition (Action/Control) and target-object location in Experiment 1b. The difference between the hand's real final position and participants' selected final position (in percentage coordinates) is plotted for the X and Y axis. The centre of each plot represents the real final position on any given trial (0px difference on each axis). Error bars show the standard error.

2.3.3 Discussion

Previous studies have shown that people integrate action kinematic information with the affordances of available objects to derive the likely goals of observed actions. This in turn biases the perception of the action towards the expected goal (Bach et al., 2011; Mcdonough et al., 2020). Such studies typically focus on low-level affordances triggered by objects, such as grip type (Buccino et al., 2009; Chainay & Humphreys, 2002; Mcdonough at al., 2020), or provide an explicit cue to intention prior to action onset (Hudson, Nicholson, Ellis, et al., 2016). Here, we tested whether objects also activate higher-level action goals by controlling for these low-level visual features and removing any explicit cue to intention. Moreover, we tested the robustness of this effect by, first, comparing perceptual biases for intentional versus non-intentional actions and, second, assessing perceptual biases in both the horizonal and vertical plane between two potential target objects.

Participants watched videos of a hand reaching for or withdrawing from a single target object (Experiment 1a) or a target object and a distractor object (Experiment 1b). Midway through the action, the hand disappeared and participants were asked to judge the index finger's final position on the touchscreen. Prior to action onset, we implicitly manipulated the observer's knowledge of the actor's intention by presenting a static image of a functionally related or functionally unrelated prime.

Results from two experiments confirmed our hypotheses. The overall response bias towards the right of the screen in Experiment 1a suggests that participants exhibited a tendency to respond towards the centre of the hand rather than the index finger, as instructed. Likewise, disappearance judgements in Experiment 1b showed an overall downward bias, consistent with previous touchscreen RM studies reporting shifts in localisation responses toward the object's centre of gravity (Hudson, Bach, et al., 2018; Mcdonough et al., 2019, 2020). Nevertheless, displacements in Experiment 1 were larger for withdrawing actions, reflecting a stronger rightwards expectation for withdrawals than for reaching, as would be expected in a representational momentum (RM) task (Hudson, Bach et al., 2018; Hudson, McDonough et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016). As predicted, the hand was misperceived as being closer to the target object when it was functionally related to the prime (e.g. hammer/nail) than when it was functionally unrelated (e.g. hammer/cigarette). This bias was present only for reaching actions, but not withdrawing actions (Experiment 1a) and persisted even when the functionally-related target object was simultaneously presented with an unrelated distractor object (Experiment 1b). Crucially, the effect was shown to be specific to intentional actions and was eliminated when the hand was replaced with a non-biological object following the same trajectory. Moreover, because our task required participants to accurately judge the hand's disappearance point, irrespective of the target objects, these results reflect automatic and involuntary effects of action expectations on perceptual judgements.

The perceptual displacement towards the expected target object is in accordance with predictive-processing accounts of social perception, which state that action perception incorporates predictions of the action’s future course, which can bias action perception toward these expectations (Bach & Schenke, 2017; Kilner et al., 2007). In other words, the hand is perceived to have disappeared closer to the object when the actor has the inferred goal to pick it up, even when the visual stimulation is identical. This finding builds on previous work in which action goals were overtly conveyed prior to action onset; for example, by hearing the actor verbally state their intention "I'll take it!" or "I'll leave it." (Hudson, Nicholson, Ellis, et al., 2016), or by matching a hand's grip configuration to an available small or large object in the environment (Ambrosini et al, 2011, 2013). Here, no such information was available. Both the target and distractor object were always matched to the grip configuration of the actor's hand. Therefore, the goal could only be inferred from the prime object presented at the start of each trial. Such effects suggest that the perceptual bias toward the anticipated goal was driven by knowledge of the actor's intention, derived from the affordances of the available objects.

A key finding in our experiments was that perceptual displacements towards expected goals were only present when observing reaching actions, but not withdrawing actions. This is in line with predictive-coding models of social perception (Kilner et al., 2007) which argue that observers are constantly testing their inferences about others' goals by predicting how the actor would behave, and integrating this prediction with incoming sensory information. When the goal and sensory input match, the prediction is processed fluently and perception becomes biased towards the expected movement. In contrast, unexpected actions elicit prediction errors and re-evaluations of prior expectations, until they better explain what is being observed (Bach & Schenke, 2017; Kilner et al., 2007). In line with previous findings (Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016), our results showed that the RM effect was stronger when the observed action was congruent with the inferred goal than when it was incongruent. When the prime and target object were functionally related (e.g., a cigarette and lighter), the reach toward the target object was in line with actor's intention, and RM increased. In contrast, a mismatch between the expected intention and the withdrawing action was easily detected and RM reduced.

Predictive processing views of social perception argue that the primary role of predictions is to test one's prior assumptions about the other person against observed behaviour. In such models, any assumptions about the person – '*they are thirsty'* – are automatically translated into concrete expectations of their forthcoming action – *'they will reach for the glass'*. Such predictions not only allow for fluent social interactions, but ensure our beliefs about others remain aligned with reality (Bach & Schenke, 2017). Our results indicate that such goal predictions are driven, at least in part, by information derived from the objects available in the given situation. However, people make a range of inferences about others, from low-level inferences about the goals of single behaviours to higher-level inferences about overarching traits (Van Overwalle et al., 2012). This raises the question to what degree does higher-level knowledge about the other person, such as personality traits or social roles, influence lower-level predictions? For example, if we believe the other person to be health-conscious, would we expect a subsequent reach with a lighter to be towards a cigarette or a candle? And how does our knowledge of the person change when they behave in a way we did not expect?

Taken together, our results support predictive processing accounts of social perception and show that action perception is guided by goal predictions, driven in part by the affordances of available goal objects. Future studies should seek to explore how higher-level predictions feed downwards through the hierarchy to influence lower-level predictions of actions.

Chapter 3: Does person knowledge prime action goals?

The previous chapter demonstrated that action perception is guided by a high-level cognitive mechanism that incorporates predictions about action goals, derived in part from object information. The current chapter seeks to determine whether such goal predictions are guided not only by object knowledge, but by prior expectations about the actors themselves.

Predictive processing models view social perception as a top-down process of hypothesis testing and revision, in which our prior assumptions about others' goals - based on various contextual cues such as objects of social signals - are tested against, and updated by, their observed behaviour (Bach et al., 2014; Gergely & Csibra, 2008; Kilner et al., 2007). In such models, any assumptions one has about the other person – e.g., *'they are on a diet'*– are automatically translated into concrete predictions about their forthcoming actions – *'they will reach for the fruit'*. These predictions not only allow for anticipatory control in social interactions by filling in missing details or biasing perception towards the future (Hudson, Bach et al., 2018; Hudson, McDonough et al., 2018; McDonough, Hudson et al., 2019), they also ensure that our beliefs about others stay aligned with reality. If the person behaves in a manner differently to expected, a prediction error is generated, triggering a revision of one's assumptions until they can better account for the observed behaviour – *'she reached for the cake, she must not be dieting anymore'* (Bach & Schenke, 2017). Direct evidence of top-down goal predictions influencing perception can be found in recent studies using a modified representational momentum paradigm (Hudson, Bach, et al., 2018; Hudson, McDonough, Edwards, & Bach, 2018; Hudson, Nicholson, Ellis et al., 2016; Hudson, Nicholson, Simpson, et al., 2016; McDonough et al., 2019). Participants heard an actor make a verbal statement about his intended goal – "I'll take it" or "I'll leave it" – before seeing him either reach for or withdraw from an object. The action disappeared midmotion and participants indicated the perceived vanishing point. In each of these studies, the hand was consistently misperceived as disappearing further toward the object when the actor said they would take it and further away from the object when the actor said they would leave it (Hudson, Nicholson, Ellis et al., 2016). These results suggest that inferred goals are directly translated into predictions about subsequent actions, which in turn bias perception toward these expectations.

Previous research has focussed on the way in which action predictions emerge from overt social cues, such as verbal statements (Hudson, Bach, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Todd et al., 2011), facial expressions (Adams et al., 2006), gaze (Bayliss et al., 2007), or grip kinematics (Mcdonough et al., 2020). However, it has been argued that prior knowledge about the actors themselves – such as their traits or beliefs – could also shape our perception of their actions (Bach & Schenke, 2017). People make a range of inferences about others, from low-level inferences about the goals of single behaviours - *she is thirsty, she will pick up the glass of water* - to higher-level inferences about overarching traits – *she dropped the glass, she is clumsy* (Van Overwalle et al., 2012). Over time, we learn how other people are most likely to behave in different situations. For example, peoples' attention is automatically biased towards the object someone *usually* looks at rather than where they are actually looking (Joyce et al., 2016). In the absence of overt cues, inferences about others can be drawn from prior knowledge about their past behaviours or the groups to which they belong (e.g., socioeconomic group; Darley & Gross, 1983; race; Bargh et al., 1996; Blair & Banaji, 1996; Sagar & Schofield, 1980; sex; Quadflieg et al., 2011). Consequently, if action predictions reflect higher-level person knowledge, rather than just overt social cues or situational constraints, when an action is ambiguous our perception should be biased towards our prediction. This raises the question to what degree does high-level person knowledge influence low-level predictions about their forthcoming actions?

Evidence from social psychology shows that when observing others' actions, people make implicit inferences about the actor's traits or dispositions (Van Overwalle et al., 2012) and their goals (Todd et al., 2011). Early priming research, in which the perceived traits of others were manipulated to assess their subsequent influence on how the perceiver interpreted their behaviour, showed that primed traits were more readily perceived in others when making subsequent judgements about their personality (Bargh et al., 1996; Higgins et al., 1977) and exerted an unintended passive influence on the interpretation of their behaviour (Bargh & Pietromonaco, 1982; Srull & Wyer, 1979). For example, participants who were unknowingly exposed to words related to hostility were more likely to judge an ambiguous behavioural description of a person negatively on both hostility-related and hostility-unrelated traits (Bargh & Pietromonaco, 1982).

More specifically, research on stereotyping, where traits are primed by merely detecting the group to which a person belongs (Blair & Banaji, 1996; Fazio et al., 1995; Dijksterhuis & Van Knippenberg, 1996; Macrae, Stangor, & Milne, 1994; for a review see Moskowitz & Olcaysoy Okten, 2016), shows that categorising an individual to a group leads people to assume that traits generally associated with that group (i.e. the stereotype) also apply to that particular individual, consequently influencing how one interprets their behaviour (Correll et al., 2002). For example, participants who believed that a child came from a high socioeconomic group reported the child's test performance as indicative of a high-level ability, whereas those who believed the child came from a low socioeconomic group reported the identical performance as indicative of a lower level ability (Darley & Gross, 1983). Moreover, the same behaviour can lead to a different inference if performed by a group member not associated with that stereotype. For example, ambiguously aggressive behaviours were rated as more mean and threatening when the perpetrator was black than when he was white, indicating that participants were more inclined to attribute violence and aggression to black actors than to white actors, in line with the stereotype that black individuals are more prone to acts of violence (Duncan, 1976; Sagar & Schofield, 1980). Although the actions investigated in these studies were more abstract than the single actions investigated in our experiments, taken together, these findings suggest that social categories can be passively primed outside of the participants' awareness and subsequently influence the interpretation of behaviour.

There is evidence that implicit biases and stereotype information can influence not only our interpretation of the world around us, but also our low-level perception (for a review, see Otten et al., 2017). For example, when matched for luminance, faces with African American features appeared darker than faces with European features (Levin & Banaji, 2006). In another study, participants were presented with images of African American or White male targets holding a gun or a nonthreatening object. When asked to shoot only those targets carrying guns, participants were more likely to accidentally shoot black computer avatars carrying tools than white avatars (Correll et al., 2002). Similar results were also obtained for targets wearing Islamic head dress (Unkelbach et al., 2008).

Stereotype information has also been shown to affect emotion recognition. For example, participants from a non-Islamic background more readily recognise fear from women wearing an Islamic headdress (a niqab) than from women wearing a Western headdress (a cap and shawl), with the opposite observed for happy and sad expressions (Kret & de Gelder, 2012). Furthermore, participants more accurately recognise happiness when a face is covered by a Western headdress compared to an Islamic headdress, and have a greater tendency to attribute sadness to a face covered by an Islamic headdress compared to a Western headdress (Kret & Fischer, 2018). Taken together, evidence suggests that high-level stereotype information can affect lower level perception.

There has been some research to address whether person knowledge can be used to predict others' forthcoming actions. For example, there is evidence that people derived motor predictions about other's forthcoming actions from prior knowledge about their typical behaviours (Bach & Tipper, 2006; Tipper & Bach, 2011). Bach and Tipper (2006) had participants identify famous football and tennis players from their faces alone, with hand or foot responses. If seeing a famous person automatically activates some of their well-known characteristics (e.g., Wayne Rooney's advanced motor skills using his feet), and humans represent such motor skills using their own motor system, one might expect the identification of a famous athlete to activate the body part most frequently used in their sport. Interestingly, their results showed contrast rather that facilitation effects: participants were faster to identify non-acting tennis players with a foot response than a hand response, and vice versa for football players (the 'negative compatibility' effect). In other words, perceiving the face of an athlete who is well-known for a specific motor skill, *inhibited* similar motor behaviour in the observer. While these results may seem counterintuitive, they nevertheless suggest that participants' knowledge of the athletes' motor skills influenced their associated motor responses. According to the authors, such negative compatibility effects are in line with predictive processing models which suggest that we make predictions about forthcoming actions based on prior knowledge about the other person (e.g., their action tendencies), which are then compared with the action we actually see them perform. In this view, seeing an athlete not performing their expected action elicits a prediction error resulting from the activation of a predicted action that could not be matched to the actual sensory input in the photograph of the non-acting athlete. Subsequent studies found that when the athletes were seen carrying out their typically associated actions (e.g., footballer players – kicking a ball), the effect was reversed and facilitation effects were observed (Tipper & Bach, 2011). Thus, when an observer sees Wayne Rooney, they predict that he will use his foot to kick a ball. This primes the motor system, activating the foot of the observer, leading to faster foot responses when Wayne Rooney is seen performing that actions (i.e. facilitation) but slower responses when he is not (i.e., negative compatibility).

There is also evidence that people learn which objects others are likely to look at. For example, attention is automatically biased towards the object someone *usually* looks at rather than where they are actually looking (Joyce et al., 2016).

Schenke, Wyer and Bach (2016) investigated the way in which implicit and explicit knowledge about another person's most likely behaviour with an object influenced action prediction. Participants watched the actions of two individuals in two situations – sitting next to a computer or standing next to a football – and reported, with speeded button presses, whether the individual interacted with or turned away from the object. Action expectancies were induced either by manipulating the frequencies of the actor's behaviours across situations (i.e., one actor would be more likely to kick the ball than type on a computer, and vice versa), or by giving participants explicit behavioural descriptions about the actors. They found that actions were identified more rapidly when performed by somebody who typically performed that action, compared to an action that was typically performed by somebody else in the given situation (Schenke et al., 2016). In contrast, participants made more errors when perceived actions conflicted with the prior person description. Taken together, evidence suggests that top-down information about the actor – their action tendencies, goals, beliefs - is integrated with the situational contains, such as the objects available to achieve the current goal, to predict their most likely actions.

The experiments in Chapter 2 harnessed the fact that objects possess functional properties which convey information about how they are used and what they are used for (i.e., *object affordances;* Gibson, 1979). However, objects can also express social meaning, which is relevant to their behavioural use. For example, the features of an object relative to the gender of the actor can determine whether the object's affordances are prohibited or facilitated. This can be seen in the differences between toys that are often aimed at girls (typically associated with domestic behaviours, such as caring for dolls and 'playing house') and toys that are often aimed at boys (typically encouraging activity away from the home, such as vehicles and sports) (Meagher, 2017; Blakemore & Centers, 2005; Jones et al., 2007). This suggests that stereotypic associations regarding men and women are reflected not just in the anticipated characteristics of other people, but also in the inanimate objects that they use. As such, one might expect person information, such as perceived gender, to be integrated with object information when making predictions about others behaviour. The previous chapter demonstrated that action goal prediction is driven, at least in part, by the affordances of available objects, and that such predictions elicit biases in the perception of subsequent actions toward the expected goal. However, predictive social perception models suggest that action predictions reflect not only overt contextual cues, but also higher-level knowledge about the actors themselves (Bach & Schenke, 2017). Thus, our assumptions about others (e.g., their values, beliefs and goals) are integrated with object information and translated into predictions about their forthcoming actions.

Here, we tested whether gender information biases action perception towards objects that are stereotypically associated with a particular gender, versus objects stereotypically associated with another gender. As in Chapter 2, we controlled for the low-level visual features of the stimuli, matching them for grip size. Cues to intention were either implicit (Experiments 2a and 2b) or explicitly implied (Experiment 2c). Moreover, we tested the robustness of the effect by investigating predictive biases for intentional actions (a reaching hand) compared to non-intentional actions (a moving ball).

We tested whether the visual presentation of a 'male' or 'female' face would lead to a predictive bias in the perception of an action towards an object stereotypically associated with that gender, versus an object stereotypically associated with the other gender. Utilising the same representational momentum paradigm as in Chapter 2 (see also Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016), participants were primed with a photograph of a male or female face before viewing a video of a hand reaching, or a ball moving, towards two objects; one 'highly feminine' object (e.g., a lipstick) and one 'highly masculine' object (e.g., a wrench; see Meagher, 2017). Part-way through the action, the hand or ball disappeared and participants indicated the index finger or ball’s final position on a touchscreen. To measure the extent to which participants were processing the prime stimuli, on some trials participants were asked to identify if the face they saw had been male or female (Experiments 2a and b), or explicitly asked which object they thought the actor was reaching for (Experiment 2c). We predicted that perceptual judgments of actions would be displaced towards the object that is stereotypically associated with the gender of the prime face stimulus. We further predicted that if this bias is driven by predictions about the actor's goals, we should only observe an effect in the intentional-action condition and not in the non-intentional ball condition.

3.1 Experiment 2a

Experiment 2a tested whether action observation automatically draws upon not just object knowledge, but upon higher-level person knowledge. Specifically, we tested whether gender information would bias action perception towards an object that is stereotypically associated with a particular gender, versus an object that is stereotypically associated with another gender. Participants were primed with a photograph of a male or female face before viewing a video of a hand reaching, or a ball moving, towards two objects; one 'highly feminine' object (e.g., a lipstick) and one 'highly masculine' object (e.g., a penknife). Midway through the action, the hand or ball disappeared and participants indicated the index finger or ball’s final position on a touchscreen. On some trials, participants were asked to identify if the face they saw had been male or female. We hypothesised that identifying a face as male or female would bias the perception of a subsequently observed action towards the object stereotypically associated with that gender, versus an object stereotypically associated with another gender.

3.1.1 Method

Details of the experimental design and analysis were pre-registered on AsPredicted.org (available at: https://aspredicted.org/wes_tkv).

Participants

Participants were 68 right-handed volunteers (33 male, 21 female, 14 unreported; mean age: 39.34 years, range: 24 to 73 years), recruited from the United States via Amazon Mechanical Turk. Participants were given $2.50 for taking part. None had taken part in the previous experiments.

Apparatus

The experiment was administered using Inquisit Web on touchscreen devices running iOS with a minimum screen resolution of 1024 x 768 pixels and a minimum refresh rate of 40Hz.

Stimuli and procedure

Experimental stimuli consisted of 6 photographs of young adults (3 male, 3 female) portraying neutral expressions, taken from the well-validated NimStim Set of Facial Expressions (Tottenham et al., 2009. Freely accessible via: https://macbrain.org/resources/ see Figure 3.3a and 3.3b for examples), 24 action sequences of an arm ambiguously reaching towards two of a possible eight objects (the target objects), and 24 sequences depicting a ball moving towards the same target objects. Target objects were selected from a study by Meagher (2017) in which 192 participants were asked to rate the degree to which they associated 80 objects with men or women. Ratings were given using a 7-point scale, with 1 labelled as 'very masculine' and 7 labelled as 'very feminine'. Four 'highly masculine' objects were selected and four 'highly feminine' objects. Mean gender association ratings for each object are displayed in Figure 3.1.

Action sequences showed a hand ambiguously reaching towards two objects; one placed towards the bottom of the screen, the other placed towards the top of the screen, forming either a whole-hand grip (WHG) or a precision grip (PG). One of the objects was stereotypically associated with the gender of the prime stimulus (“the target”), while the other was stereotypically associated with the opposite gender (“the distractor”). In half of the videos, the target object was orientated towards the bottom of the screen and in the other half towards the top of the screen. Object pairs were matched for grip type (i.e. both objects were a whole hand grip or both objects with a precision grip). Each video comprised 15 frames (each 20 ms) showing the complete reaching action, but with the first 5 frames and the final grasp omitted. Action sequences were either 9, 12 or 15 frames long. The response screen remained visible until a response was given or 10 seconds had elapsed. For each condition, participants completed a practice block of 6 trials, followed by three experimental blocks each consisting of 20 trials. These 60 trials comprised two factors: Target Location (Top, Bottom), Action Sequence (9, 12 or 15 frames), and catch trials. Trials were randomly presented within each block.

The experiment contained four catch trials per block in which the response screen was followed by a question relating to either the object or face stimuli. For example: *'Was the face you just saw male or female?*' Participants responded by pressing either 'Male' or 'Female' on the touchscreen. The question remained on the screen until a response was given, or 10 seconds had elapsed.

Figure 3.1

Trial sequence for the Action (top) and Control (bottom) conditions (Experiment 2a).

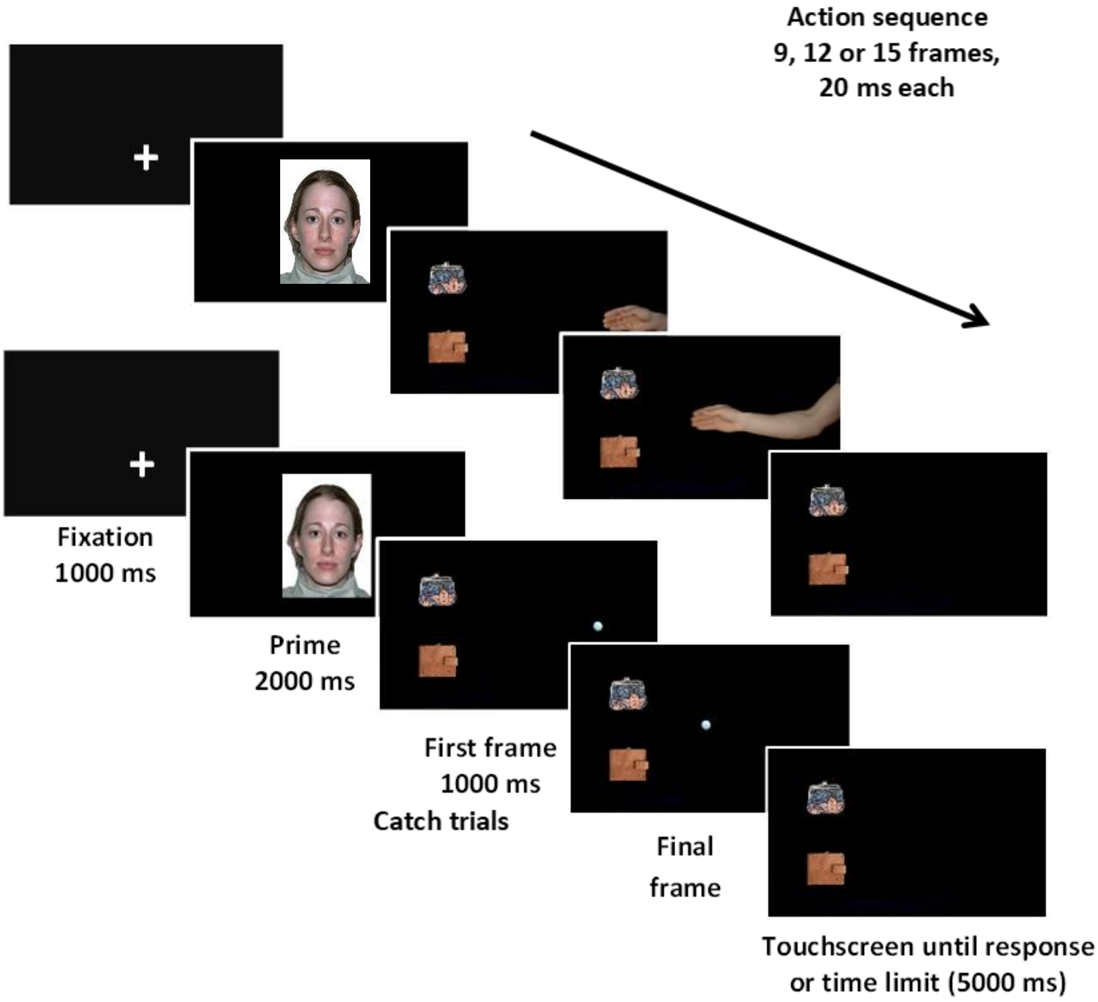


Table 3.1

Mean gender association scores for target objects in Experiment 2a-c.

Precision Grip (PG)				**Whole-Hang Grip (WHG)**			
Male		**Female**		**Male**		**Female**	
Swiss Army Knife	2.46 (1.04)	Lipstick	6.82 (0.50)	Wrench	2.04 (0.94)	Hand mirror	6.29 (0.78)
Razor	2.04 (1.03)	Brush	5.48 (1.14)	Wallet	2.16 (1.04)	Purse	6.73 (0.54)

Taken from Meagher (2017). Objects were rated on a 7-point scale (1 = *Extremely Masculine*, 7 = *Extremely Feminine*). Standard deviations are shown in brackets.

Data analysis

As in Experiment 1b, The experiment employed a 2 x 2 repeated-measures design, with Condition (Action, Control) and Target Object Location (Top, Bottom) as our independent variables. The target object was defined as the object that was congruent with the gender of the prime stimulus. Screen coordinates were converted into percentage coordinates, where 0%, 0% represents the top left-hand corner of the screen. Participants' selected screen coordinate (in percentage coordinates) of the tip of the index finger/ball on each trial was subtracted from the real final screen coordinate. Analyses were conducted on this residual localisation error, providing a directional measure of how far participant's responses were displaced along the X and Y axis. An accurate response would produce a value of 0% on both axes. On the X axis, negative values denote a rightward displacement (against the direction of motion) and positive values a leftward displacement. On the Y axis positive and negative values denote upward and down displacements respectively.

3.1.2 Results

A priori power analysis (https://jakewestfall.shinyapps.io/pangea) indicated that 51 participants were needed to have 80% power for detecting an interaction with a medium sized effect (*d* = 0.52, as in Experiment 2), when employing the traditional .05 criterion of statistical significance. As outlined in the pre-registration, participants with < 60% accuracy on catch trials were excluded from the analyses. This left us with a sample size of 61.

Y-axis

Our main prediction was that perceptual judgments of actions would be displaced towards the expected trajectory; that is, towards the object that is typically associated with the gender of the prime face stimulus. A two-way ANOVA was conducted to examine the effect of Condition (Action, Control) and Target Object Location (Top, Bottom) on Displacement along the y-axis. The analysis revealed a main effect of Condition (*F* (1, 60) = 10.66, *p* = .002, hp^2 = .15). Displacements in the Action condition were larger than the Control condition (See Figure 3.2). However, there was no main effect of Target Location (1, 60) = 2.78, *p* = 1.00, hp^2 = .044) and no interaction (*F* (1, 60) = 3.89, *p* = .053, hp^2 = .06).

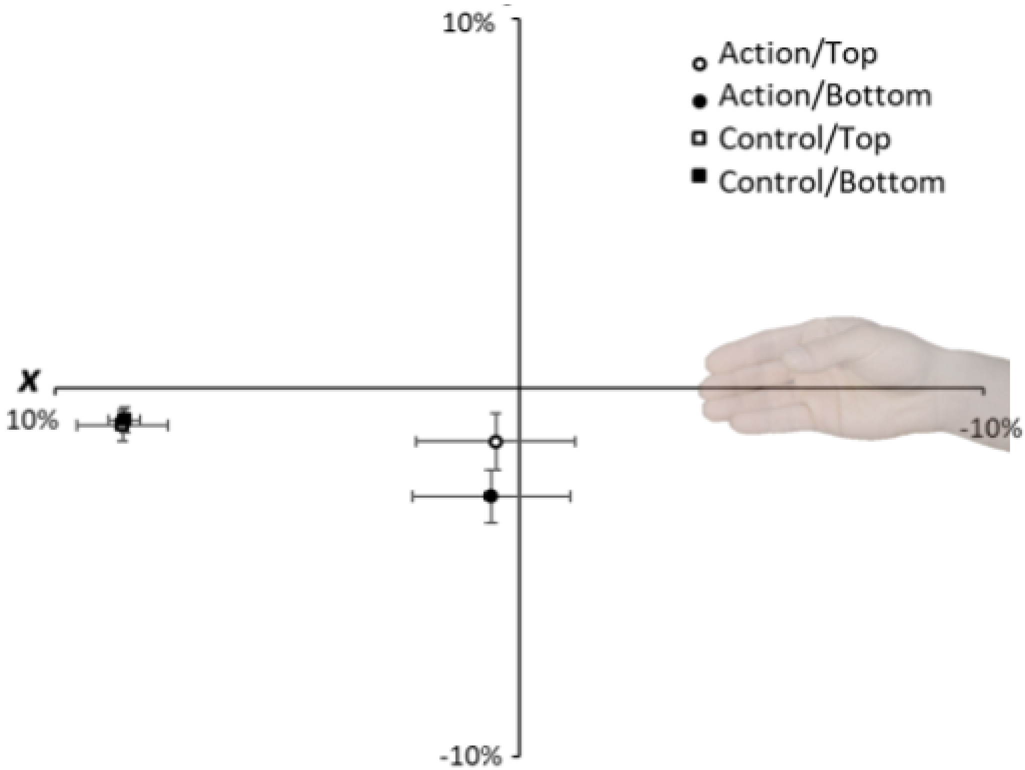
Pairwise comparisons with a Bonferroni adjustment showed no difference in displacements between target locations in the Action condition (*M* = 1.49%, *SD* = 6.31, *t* (60) = 1.84, *p* = .071, *d* = 0.25) or the Control condition (*M* = 0.14%, *SD* = 0.89 *t* (60) = -1.23, *p* = .225, *d* = 0.05).

X-axis

We did not have any specific predictions about perceptual displacements on the X-axis. However, one-sample t-tests revealed that displacements in the Action condition were not significantly larger than zero (*M* = 0.56%, *SD* = 13.31, *t* (60) = .03, *p* = .743, *d* = .04). However, the Control condition demonstrated the classic representational momentum effect, showing displacement in the direction of motion (*M* = 8.52%, *SD* = 7.76, *t* (60) = 8.56, *p* < .001, *d* = 1.10).

Figure 3.2.

Experiment 2a results.



The interaction between task condition (Action/Control) and target-object location in Experiment 2a. The difference between the hand's real final position and participants' selected final position (in percentage coordinates) is plotted for the X and Y axis. The centre of each plot represents the real final position on any given trial (0px difference on each axis). Error bars show the standard error.

3.1.3 Discussion

Experiment 2a tested whether the visual presentation of a 'male' or 'female' face would lead to a predictive bias in the perception of a subsequent action towards an object that is stereotypically associated with that gender (e.g., female – lipstick) versus an object stereotypically associated with the other gender (e.g., male – penknife). We also tested whether this effect was specific to intentional (a reaching hand) versus non-intentional (a moving ball) actions.

Contrary to our predictions, Experiment 2a did not provide evidence that passively viewing a male or female face activates gender stereotypes to such an extent that they bias action perception. One reason for this might be that the effect size is simply too small to detect within the current paradigm. The strength of the association between the male and female faces and the 'masculine' and 'feminine' objects might have been too weak to influence action judgements. We attempted to strengthen this association in Experiment 2b by replacing the prime face stimuli with a new set of faces which were independently rated as being either 'highly masculine' or 'highly feminine' (see Ma et al., 2015). This should allow for greater accessibility to the categories 'male' and 'female' and their associated stereotypes when viewing the prime stimulus.

3.2 Experiment 2b

Experiment 2b sought to enhance the categorisation of the prime stimuli as distinctly 'male' or 'female'. To do this, we used a new set of face stimuli extracted from the Chicago Face Database (Ma et al., 2015). In comparison to the faces used in Experiment 2a, which were drawn from the NimStim Set of Facial Expressions (Tottenham et al., 2009), the standardized neutral faces found in the Chicago Face Database were subjectively rated as being either highly masculine or highly feminine.

3.2.1 Method

Participants

Participants were 52 right-handed volunteers (22 male, 30 female; mean age: 37.39 years, range: 20 to 77 years), recruited from the United States via Amazon Mechanical Turk. Participants were given $2.50 for taking part. None had taken part in the previous experiments.

Apparatus

The experiment was administered using Inquisit Web on touchscreen devices running iOS with a minimum screen resolution of 1024 x 768 pixels and a minimum refresh rate of 40Hz.

Stimuli and procedure

Experimental stimuli and procedure were identical to Experiment 2a (see Figure 2.1), with the exception of the prime stimuli (see Figure 3.3), this time selected from the Chicago Face Database (Ma et al., 2015). The Chicago Face Database provides extensive subjective norming data for 158 standardised photographs of Black and White male and female faces aged between 18 and 40 years. These standardized neutral faces were subjectively rated by 552 males and 308 females with a mean age of 26.75 years, from diverse racial backgrounds. We selected three White female faces with a mean femininity rating of 5.7 and three White male faces with a mean masculinity rating of 5.4 (1 = Not at all, 7 = Extremely). Target objects were the same as in Experiment 2a (see Figure 3.1).

Figure 3.3.

Example face stimuli from Experiments 2a–c.

Experiment 2a used neutral female (a) and male (b) faces taken from the NimStim Set of Facial Expressions (Tottenham et al., 2009). Experiment 2b used neutral female (c) and male (d) faces, with high subjective ratings of femininity or masculinity, from the Chicago Face Database (Ma et al., 2015).

3.2.2 Results

As in Experiment 2a, participants with < 60% accuracy on catch trials were excluded from the analyses, giving a sample size of 50.

Y-axis

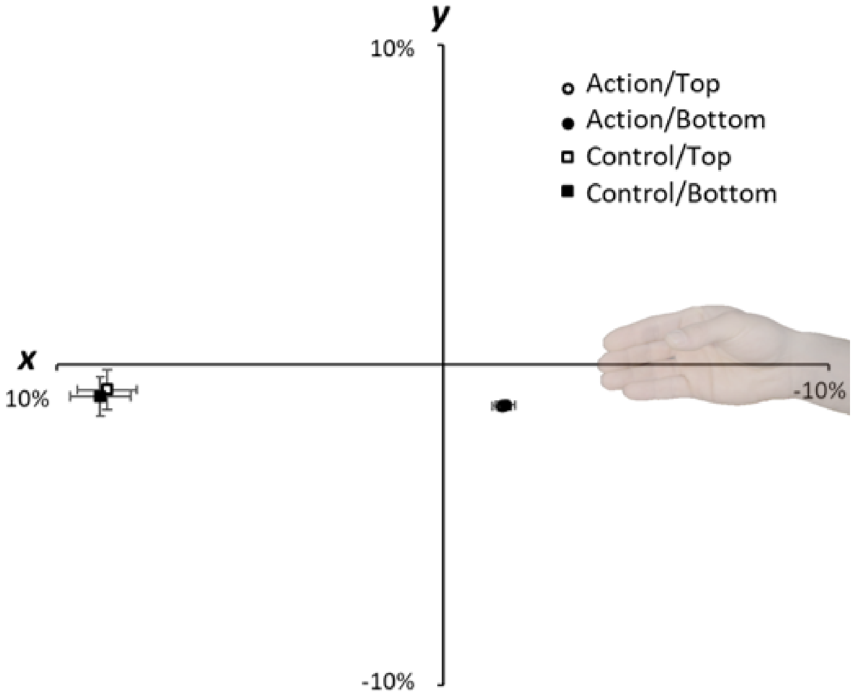
A two-way ANOVA was conducted to examine the effect of Condition (Action, Control) and Target Object Location (Top, Bottom) on Displacement along the y-axis. The analysis revealed no main effect of Condition (F (1, 49) = 0.41, p = .528, hp^{2} = .008) or Target Location (F (1, 49) = 0.89, p = .349, hp^{2} = .02), and no interaction (F (1, 49) = 1.57, p = .217, hp^{2} = .03) (see Figure 3.4).

X-axis

One-sample t-tests revealed that displacements in the Action condition were not significantly larger than zero (M = -0.73%, SD = 5.52, t (51) = -.95, p = .346, d = .13). However, the Control condition demonstrated the classic representational momentum effect, showing displacement in the direction of motion (M = 8.97%, SD = 5.88, t (51) = 11.00, p < .001, d = 1.53).

Figure 3.4.

Experiment 2b results.



The interaction between task condition (Action/Control) and target-object location in Experiment 2b. The difference between the hand's real final position and participants' selected final position (in percentage coordinates) is plotted for the X and Y axis. The centre of each plot represents the real final position on any given trial (0px difference on each axis). Error bars show the standard error.

3.2.3 Discussion

Experiment 2b replicated Experiment 2a, with the exception of the prime face stimuli, which were selected from the Chicago Face Database (Ma et al., 2015). These were chosen because they had been rated as being either 'highly masculine' or 'highly feminine' faces, in contrast to the stimuli used in Experiment 2a which had no such ratings. However, we were unable to provide evidence for such an effect.

Again, the strength of the association between the male and female faces and the 'masculine' and 'feminine' objects appears too weak to exert any influence on action judgements.

So far, our focus has been on drawing participants' attention to the gender of the faces, potentially limiting what attention is drawn to the objects themselves. Indeed, the objects in the videos bear no consequence on participants' ability to correctly answer the catch trial questions, making it entirely possible that they are ignoring the objects altogether. With this in mind, Experiment 2c replicated Experiment 2b, with the exception of catch trials in which participants were explicitly asked which object they thought the actor was reaching for. This was in contrast to Experiments 2a and 2b, in which participants were asked to recall the gender of the face. By asking participants to explicitly think about which object the actor might reach for, we aimed to strengthen the association between the faces and the objects, thus maximising any effect of gender bias on action perception.

3.3 Experiment 2c

Experiment 2c replicated Experiment 2b, differing only in that participants were explicitly asked which object they thought the actor was reaching for. By asking participants to explicitly think about which object the actor might reach for, we aimed to maximise any effect of gender bias on action perception.

3.3.1 Method

Participants

Participants were 57 right-handed volunteers (22 male, 35 female; mean age: 35.63 years, range: 18 to 71 years), recruited from the United States via Amazon Mechanical Turk. Participants were given $2.50 for taking part. None had taken part in the previous experiments.

Apparatus

The experiment was administered using Inquisit Web on touchscreen devices running iOS with a minimum screen resolution of 1024 x 768 pixels and a minimum refresh rate of 40Hz.

Stimuli and procedure

Experimental stimuli and procedure were identical to Experiments 2b, with the exception of the catch trials. On these trials participants were explicitly asked: '*Given the face you have just seen, which object do you think the person was reaching for?'* and responded by selecting either 'Top' or 'Bottom' on the touchscreen.

3.3.2 Results

Participants with < 60% accuracy on catch trials were excluded from the analyses, giving a sample size of 49. Accurate responses were defined as responses that fit with the gender stereotype. Participants' mean catch trial accuracy was 88%.

Y-axis

A two-way ANOVA was conducted to examine the effect of Condition (Action, Control) and Target Object Location (Top, Bottom) on Displacement along the y-axis. The analysis revealed a main effect of Condition (F (1, 48) = 4.94, p = .031, hp^2 = .093) and Target Location (F (1, 48) = 87.34, p < .001, hp^2 = .65), but no interaction F (1, 48) = 3.08, p = .086, hp^2 = .06) (see Figure 3.5).

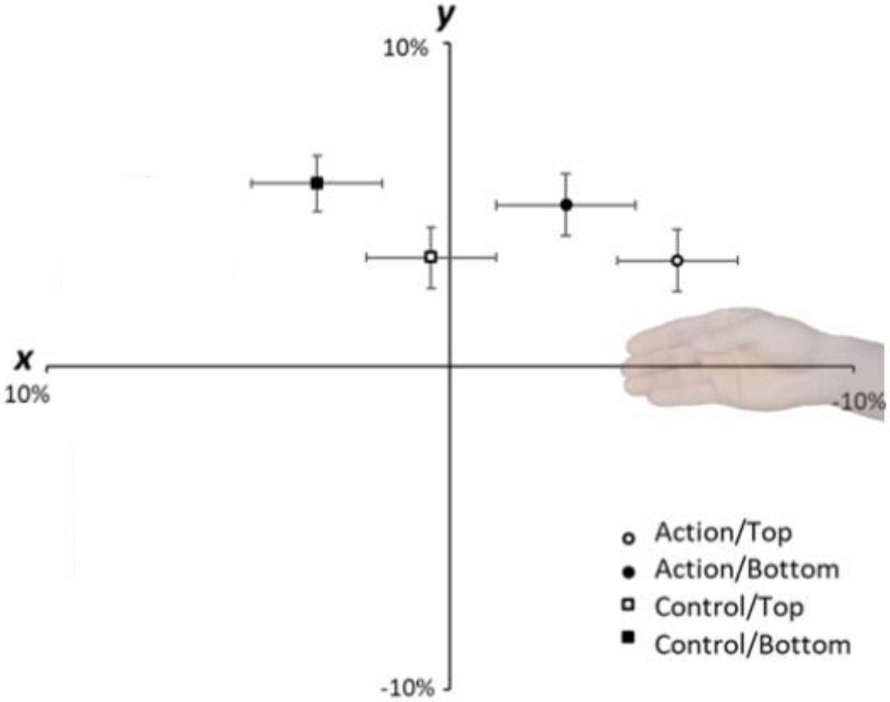
Pairwise comparisons with a Bonferroni adjustment showed that participants consistently misperceived the disappearance point as being closer to the distractor object than it really was in both the Action condition (M = 1.71%, SD = 2.14, t (48) = -5.62, p < .001, d = 0.81) and Control Condition (M = 2.29%, SD = 1.60, t (48) = -10.04, p < .001, d = 1.38).

X-axis

One-sample t-tests revealed that displacements in the Action condition were significantly smaller than zero, indicating an overall rightward bias. As in our previous studies, this suggests that participants had a tendency to respond towards the centre of the hand rather than the index finger (M = -4.24%, SD = 10.87, t (48) = -2.73, p = .009, d = 0.39). However, displacements in the Control condition were not significantly larger than zero (M = 1.88%, SD = 10.91, t (48) = 1.21, p = .233, d = .0.17).

Figure 3.5.

Experiment 2c results.



The interaction between task condition (Action/Control) and target-object location in Experiment 2c. The difference between the hand's real final position and participants' selected final position (in percentage coordinates) is plotted for the X and Y axis. The centre of each plot represents the real final position on any given trial (0px difference on each axis). Error bars show the standard error.

3.3.3 Discussion

The experiments in this chapter aimed to test whether gender information can bias action perception towards an object that is stereotypically associated with a particular gender, versus an object that is stereotypically associated with another gender. Experiments 2a and 2b were unable to detect any such effect of gender stereotyping with regards to objects when participants' attention was drawn to the gender of the person in the prime stimulus. In the present study, we aimed to strengthen the association between the face and the target object by explicitly asking participants which object they thought the actor was reaching for.

Participants' high overall accuracy in catch trials (88%) suggested they were attending to the prime faces and associated the male faces with stereotypically 'masculine objects' and the female faces with stereotypically 'feminine objects'. Yet, surprisingly, when making perceptual judgements, participants had a tendency to misperceive the disappearance point of the actor's hand, and the moving ball, as being closer to the *distractor* object than to the target object. These contrasting results suggest a dissociation between the effects of implicit and explicit information when making action predictions. When asked on catch trials to make a link between the gender of the face and the objects, participants may have been responding in what they considered to be the "correct" way. However, explicitly priming this link between gender and objects could arguably have made participants more aware of *not* portraying sexist behaviours. Consequently, when making perceptual judgements about the hand/ball's final location, responses reflected an implicit reaction *agains*t our portrayal of sexism.

3.3.4 General Discussion

Predictive models of social perception assume that action predictions reflect higher-level person knowledge, not just overt social cues (Bach & Schenke, 2017). Here, we tested whether gender information could bias the perception of others' actions towards an object stereotypically associated with a particular gender, versus an object stereotypically associated with another gender. We also sought to test the robustness of this effect by investigating predictive biases for intentional actions (a reaching hand) versus non-intentional actions (a moving ball). Participants were primed with a photograph of a male or female face before viewing a video of a hand reaching, or a ball moving, towards two objects; one 'highly feminine' object (e.g., a lipstick) and one 'highly masculine' object (e.g., a penknife). Midway through the action, the hand or ball disappeared and participants indicated the index finger or ball's final position on a touchscreen. On some trials, participants were asked to identify whether the face they had seen was male or female (Experiments 2a and b) or explicitly asked which object they thought the actor was reaching for (Experiment 2c).

We hypothesised that identifying a face as male or female would bias the perception of a subsequent action towards the object that is stereotypically associated with that gender, versus an object that is stereotypically associated with another gender. However, none of our experiments provided evidence for such an effect. Indeed, Experiment 2c, in which participants were explicitly asked which object they thought the actor might be reaching for, provided evidence in the *opposite* direction - when presented with a female face, participants more readily misperceived the hand as being closer to the 'masculine object' and when presented with a male face, they more readily perceived the hand as being closer to the 'feminine object'. This bias was present for both reaching actions (a hand) and nonintentional movements (a ball), following the same trajectory.

A consistent finding across our studies was that perceptual judgements in the control condition - where the reaching hand was replaced with a non-biological ball stimulus - were more readily perceived as being further along the horizonal axis than they really were, showing a displacement in the direction of motion. This is perhaps not surprising given the characteristics of the stimulus – a small ball moving in a horizonal direction. Thus, the viewer's knowledge of the ball's trajectory and the way in which a ball typically behaves would have determined the direction of the distortion. Furthermore, a representational momentum effect in the Action condition would be offset by a bias towards the centre of the hand, whereas for the ball this is not the case. A larger unknown object may therefore be more suitable for the control condition in a social representational momentum task like this.

Previous research has shown that when an individual perceives a member of a particular social group, information about that group is automatically activated, including attitudes, traits, and social stereotypes (Blair & Banaji, 1996; Correll et al., 2002; Darley & Gross, 1983; Dijksterhuis & Van Knippenberg, 1996; Fazio et al., 1995; Moskowitz et al., 2016). Since any one person could be judged according to their membership in any number of groups (e.g., race, sex, age), their classification will depend on the relative accessibility of the relevant categories (Higgins, 1996). When a particular social category, such as 'female', becomes activated when perceiving a member of that group, that category remains active (or 'primed') for some time thereafter and can influence the categorization of subsequently presented ambiguous stimuli, including objects. Importantly, any associated knowledge, including stereotypical traits, also remains accessible and is used in the interpretation of others' behaviour. (Ferguson & Bargh, 2004).

Contrary to our predictions, Experiments 2a and 2b could not confirm that passively viewing a male or female face activated any such gender-based information. One possible explanation for this is that the association between the faces and the objects was not strong enough to induce any detectable biases.

Indeed, the magnitude of the priming effect on behaviour has been found to depend upon the strength of the association in memory between the prime (i.e. the stereotype) and the particular behaviour. For example, Dijksterhuis and colleagues (2000) found that people who reported having previous contact with the elderly performed worse on a free-recall memory test after being primed with the 'forgetfulness' stereotype of the elderly, whereas people who reported having little previous contact showed no priming effect. This suggests that that the degree to which abstract knowledge activation impacts on behavioural consequences depends on the strength of that association in memory (Ferguson & Bargh, 2004). Although behaviour is, in part, socially controlled, everyone is exposed to and influenced by gender stereotyping to varying degrees. A low level of exposure to gender stereotypes could have diminished the perceptual bias investigated in our experiments. Future research should include an assessment of the level of gender stereotyping participants are exposed to in their daily lives. If exposure to gender stereotyping modulates the effect of the prime stimulus on action prediction, we would expect to see a stronger perceptual bias when exposure is high than when it is low, if there is one at all.

A similar explanation relates not only to the face stimuli, but to the objects themselves. While the objects used in our experiments were selected on the basis of being rated as either 'highly masculine' or 'highly feminine' objects, the same objects have been found to be differentially associated with men and women on the basis of participant's beliefs about gender and the strength of their identification with their own gender (Meagher, 2017). In particular, benevolent sexism – an ideology that views people who conform to traditional gender roles in a positive manner – was found to predict a greater likelihood of assigning a gender to objects that were rated as neutral by the sample as a whole. In other words, if individuals high in benevolent sexism view men and women differently than individuals low in benevolent sexism, they may also interpret the physical world differently. It is therefore possible that these same attitudes influenced the outcome of our experiment. The inclusion of a measure of sexist attitudes and beliefs, such as the Ambivalent Sexism Inventory (Glick & Fiske, 1996) in future studies could help to determine the extent to which sexist attitudes and beliefs influence object and action perception.

Although not as we predicted, our results are partially compatible with some previous research that demonstrated seemingly counterintuitive priming effects (Bach & Tipper, 2006; Schenke et al., 2021; Tipper & Bach, 2011). For example, Tipper and colleagues asked participants to identify famous football and tennis players from their faces alone, giving foot or hand responses, respectively. They found a 'negative compatibility effect' whereby hand responses were slower when identifying tennis players than football players and foot responses were slower when identifying football players than tennis players. According to the authors, such contrast effects can be explained using predictive processing models: We make predictions about others' forthcoming actions based on prior knowledge about that person (e.g., Wayne Rooney will kick a ball). This prediction is then compared with the action we actually see them perform. Unexpected actions (e.g., a non-performing athlete) elicit a prediction error resulting from the activation of a predicted action that could not be matched to the actual sensory input, thus inhibiting motor performance. Along a similar vein, our results might reflect a prediction error generated by the activation of a predicted action (e.g., a woman reaching for a lipstick) that could not be matched to the ambiguous reaching action observed in the video stimuli.

The achievement of object-directed goals often requires irrelevant distractor objects to be inhibited (Tipper et al., 1994). Tipper (1985) argued that if internal representations of distractor objects are inhibited, when one is subsequently required to select and respond to one of those inhibited objects, that response will be impaired. Since each object is presented on 25% of trials with one other object, and in half of these trials it is considered a distractor and in half of these trials it is the target, the simultaneous activation of both object representations on each trial might lower the overall response tendencies for both of these objects (Tipper & Bach, 2011). A suggestion for future research is that we simplify the paradigm, removing the second object entirely (as in Experiment 1a), and test for a representational momentum effect towards a single target object that is either 'highly masculine' or 'highly feminine'.

A surprising finding in Experiment 2c was that participants had a tendency to misperceive the actor's hand and the moving ball as being closer to the *distractor* object than to the target object, while making explicit judgements about the actor's likely behaviour that indicated an expected reach toward the target object. Although not in the direction we predicted, the dissociation between the effect on participant's action responses and their ability to make explicit judgements about the actor's behaviour is consistent with previous research on the way in which person knowledge shapes action predictions (Schenke et al., 2016). Schenke, Wyer & Bach (2016) had participants perform a simple action identification task in which they judged whether two actors interacted with or withdrew from objects (a computer or a football), while manipulating the actor's action likelihoods across situations such that one actor typically interacted with one object and withdrew from the other, while the other exhibited the opposite behaviour. In a second experiment, participants additionally received explicit behavioural descriptions about the two actors (e.g., *'George typically kicks the ball but rarely types on the computer'*) that either matched or mismatched their observed behaviours and were asked to rate how appropriate this description was after observing their behaviour. They found that implicit likelihoods sped up the identification of typical versus atypical behaviours, irrespective of explicit knowledge about the individual's behaviour, whereas explicit person knowledge led to actions being misidentified altogether for the action that was expected (e.g., identifying a kick as a withdrawal when a withdrawal was expected).

Schenke and others (Shanks & Perruchet, 2002; Shanks & St. John, 1994) have argued that the dissociation between implicit response bias and explicit judgements suggests that participants might not be making explicit judgements during social perception at all. Rather, they form ad-hoc impressions based on their memory of what was previously observed and the re-activation of internal models based on the cues provided. It could therefore be argued that after making the initial action judgement in Experiment 2c, participants attempted to solve the explicit judgement task by reactivating gender-based stereotypes based on the face stimulus presented on the given trial. However, it remains unclear why the response bias observed in Experiment 2c was in the opposite direction to that which we predicted and to what extent explicitly derived person knowledge might influence this effect. One possibility is that by explicitly priming participants to make a link between the gender of the face and the objects, we made them more aware of *not* portraying sexist behaviours. This is particularly likely if participants do indeed score low on measures of sexism, as discussed above. They may have been responding to catch trials by giving what they believed to be the "correct" answer, but their actual behaviour reflected a reactance against sexist stereotypes. Future work could test whether explicit stereotype-reinforcing information about others leads to a similar predictive bias, or whether the effect is reversed to match the explicit behavioural judgements.

Finally, it should be acknowledged that the reaching hand observed in these experiments was always of the same male. This could be problematic, particularly for an experiment designed to tap into participants’ gender biases. Although this issue has not been widely addressed in the literature, one study investigating automatic imitation – the finding that movement execution is facilitated by compatible, and impeded by incompatible, observed movements – reported stronger imitative responses for female, compared to male, participants when a female stimulus hand was used (Butler et al., 2015). Although it should be noted that they did not manipulate the gender of the stimulus hand. Nevertheless, one might predict that automatic imitation should be stronger when there is a match between the gender of the stimulus and the gender of the participants.

Across 3 experiments, we aimed to address the degree to which action goal prediction is influenced by prior knowledge about the gender of the actor. In Experiments 2a and 2b, participants' attention was drawn to the gender of the prime face stimulus by asking them to identify whether the face they had seen was male or female. The lack of an observable bias may be due to participants' weak associations between the faces and objects, meaning neither gender schema could be activated to the extent that it led to a predictive bias towards . Experiment 2c sought to strengthen this association by explicitly asking participants which object they thought the actor was reaching for. Although participants associated the male faces with stereotypically 'masculine objects' and the female faces with stereotypically 'feminine objects' when explicitly asked, they had a tendency to misperceive the disappearance point as closer to the *distractor* object than to the target object. We suggest that this likely reflects a dissociation between the effects of implicit and explicit information. When explicitly priming associations between gender and objects, participants became more aware of not portraying sexist behaviours. Thus, when making perceptual judgements about an action, participants implicitly reacted *against* gender stereotypes.

Chapter 4: Do expertise guide action goal prediction?

In the previous chapters, we demonstrated that action perception is guided by top-down predictions about action goals, derived in part from the affordances of available objects (see Chapter 2). We further hypothesised that prior knowledge about the actors themselves – specifically, their gender - would guide action goal prediction. However, our experiments could not detect any evidence for this effect (see Chapter 3). The present chapter aims to test whether action goal predictions are guided by participants' high-level expertise. In this online experiment, participants were either skilled chess players or novices. Participants in each group watched a series of videos showing a hand reaching towards a chessboard, each set up in a checking configuration. Part-way through the action, the hand disappeared and participants indicated the index finger’s final position on the touchscreen. We predicted that prior knowledge of the rules of chess would lead to a predictive bias in the perception of a subsequent action towards the expected target chess piece.

Hierarchal predictive coding models of cognition argue that perception in general (e.g., Spratling, 2016), and social perception in particular (Bach & Schenke, 2017), is not just informed by bottom-up signals from the environment to the brain, but also by top-down predictions based on our expectations about what those incoming signals will be. This is accomplished using contextual cues, such as language (Hudson, Bach, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Todd et al., 2011), gaze and emotional expression (Bayliss et al., 2007), or available objects (Bach et al., 2014), to generate predictions about the actor's intentions and goals, which are checked in turn against the incoming sensory inputs. When the sensory inputs mismatch expectation, prediction errors occur, which propagate back up the hierarchy where predictions are revised to better reflect reality (Westra, 2019).

Previous research has focussed on the way in which action predictions emerge from contextual cues in the environment, such as verbal statements (Hudson, Bach, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Todd et al., 2011), facial expressions (Adams et al., 2006), gaze (Bayliss et al., 2007), or grip kinematics (Mcdonough et al., 2020). However, predictive processing models suggest that action predictions reflect more abstract, high-level knowledge, not just overt social cues (Bach & Schenke, 2017).

In Chapter 2 we demonstrated that action predictions can be derived from object affordances. In our everyday lives we are surrounded by numerous familiar objects that are associated with particular functions. Through repeated experience, we learn what these objects are for ('function knowledge'), how they are used ('manipulation knowledge') and the complex relations between them (Bach et al., 2014; Bilalić et al., 2010). In other words, we become *experts* at mastering our daily routine because of our previous experience with and knowledge about common everyday objects.

The game of chess has been used as a paradigm in cognitive psychology and neuroscience research as it involves multiples cognitive functions, including object recognition, conceptual knowledge, memory and the processing of spatial configurations (Bilalić, 2017; Harel et al., 2013). Similar to the complexities of everyday life, it offers a complex environment that requires a broad range of higher cognitive processes such as problem solving (Bilalić, McLeod, & Gobet, 2008; 2009) and decision making (Bilalić et al., 2014), as well as lower level perceptual processes like object or pattern recognition (Langner, Eickhoff & Bilalić, 2019).

Chess positions consist of individual objects – chess pieces – with characteristic shapes that must be recognised by the player. This object recognition allows the player to ascribe functions to the individual objects, based on prior knowledge of the rules of their movements. In other words, recognising a particular chess piece is tightly connected with activating the representation of potential actions afforded by the object (Langner, Eickhoff & Bilalić, 2019).

Expert chess players possess the remarkable skill of being able to visualise numerous functional relations that form between individual chess pieces, recognising patterns on the board and developing strategies based on these patterns. Evidence suggests that expertise in a task such as chess strongly improves perceptual skills, like the encoding of task-specific stimuli. Experts are said to possess domain-specific knowledge structures, often referred to as 'chunks' (Chase & Simon, 1973), acquired through extensive and focussed exposure to domain-specific stimuli. Knowledge structures include information about individual objects and their connected functions. This information enables chess experts' superior recognition of domain-specific objects, even in the absence of context, compared to novices (Bilalić et al., 2010; Kiesel et al., 2009). Furthermore, knowledge structures enable top-down predictions that automatically direct experts' attention towards the most important stimulus features, thus facilitating pattern recognition. Chess experts can recall the locations of chess pieces from a game position almost perfectly, even when the chessboard is only presented for a very brief period of time (2 – 15s)(de Groot, 1965). In contrast, novices do not have inherently weaker cognitive abilities than experts, but they lack the knowledge structures that guide perception (Chase & Simon, 1973; Bilalić et al., 2010).

The present study tested whether action goal predictions are guided by high-level chess expertise. In this online experiment, participants were either experienced chess players or novices. Participants in each group watched a series of videos showing a hand reaching towards a chessboard, each set up in a checking configuration. Part-way through the action, the hand disappeared and participants indicated the index finger’s final position on the touchscreen. We predicted that the chess experts' domain-specific object knowledge would lead to a predictive bias in the perception of a subsequent action towards the expected target chess piece. In contrast, novice players, who lack the the required knowledge, would have no expectations as to the object the actor might be reaching for, and would therefore exibit no perceptual bias.

4.1 Experiment 3

4.1.1 Method

Participants

Fifty right-handed participants were recruited from the United States via Amazon Mturk (30 male, 20 female; mean age: 37.03 years, range: 18 to 16 years). We initially planned to recruit expert chess players by contacting a selection of UK-based chess clubs. However, this proved more difficult than anticipated and after a several weeks were still unable to recruit the necessary number of participants. Instead, skilled players were recruited via an online pre-test comprising 10 checkmate puzzles. The 40 top-scorers were invited to participate in the main experiment. Twenty-five skilled players responded and completed the main task (this was our ‘Players’ group). Twenty-five unskilled players were recruited via an advert for a 'simple touchscreen task' and given the same checkmate puzzles prior to participation in order assess their chess knowledge (this was our ‘Novice’ group). The distribution of scores on the pre-test can be seen in Figure 4.4. All participants received $3 for taking part. None had taken part in the previous experiments.

Apparatus

The experiment was administered online using Inquisit Web on touchscreen devices running iOS with a minimum screen resolution of 1024 x 768 pixels and a minimum refresh rate of 40Hz.

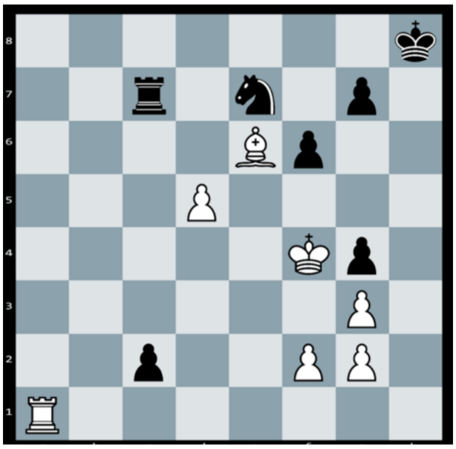
Stimuli and procedure

Pre-Test: Checkmate Puzzles Task

Stimuli consisted of 10 'mate-in-one' checkmate puzzles. Each puzzle was presented as a 2-dimensional chessboard (see Figure 4.1) against a black background. In each of the puzzles, a checkmate could be achieved in a single move of a white chess piece. Participants responded by first tapping on the piece they wished to move, then tapping on the square they wished to move it to. Each tap generated a short tone to indicate a selection had been made. Immediate feedback was given at the end of each trial by presenting either ‘correct’ or ‘error’ in green or red text, respectively, at the bottom of the screen for 1000 ms. Participants were allocated 15 minutes to complete the task. However, prior to beginning the task they were advised to spend no longer that one minute on any given trial. They were also instructed that if they did not know an answer they should take a guess and move on to the next puzzle.

Figure 4.1.

Example checkmate puzzle from the pre-test (Experiment 3).



Chess Task

Experimental stimuli consisted of 24 action sequences showing a hand reaching from the right side of the screen towards a chessboard situated to the left, forming a precision grip. The chessboard was positioned at a sideways angle, as if the participant were observing a chess match (see Figure 4.2) and set up such that a checkmate could be achieved in a single move with a white chess piece. Eight different board configurations were used, consisting of the 4 easiest puzzles from the Checkmate Puzzles Task, each of which was mirrored. In half of the trials the target chess piece was positioned at the top the board (to the right of the actor) and in half of the trials the target chess piece was positioned at the bottom of the board (to the left of the actor). Each video showed the complete transport phase, but with the first

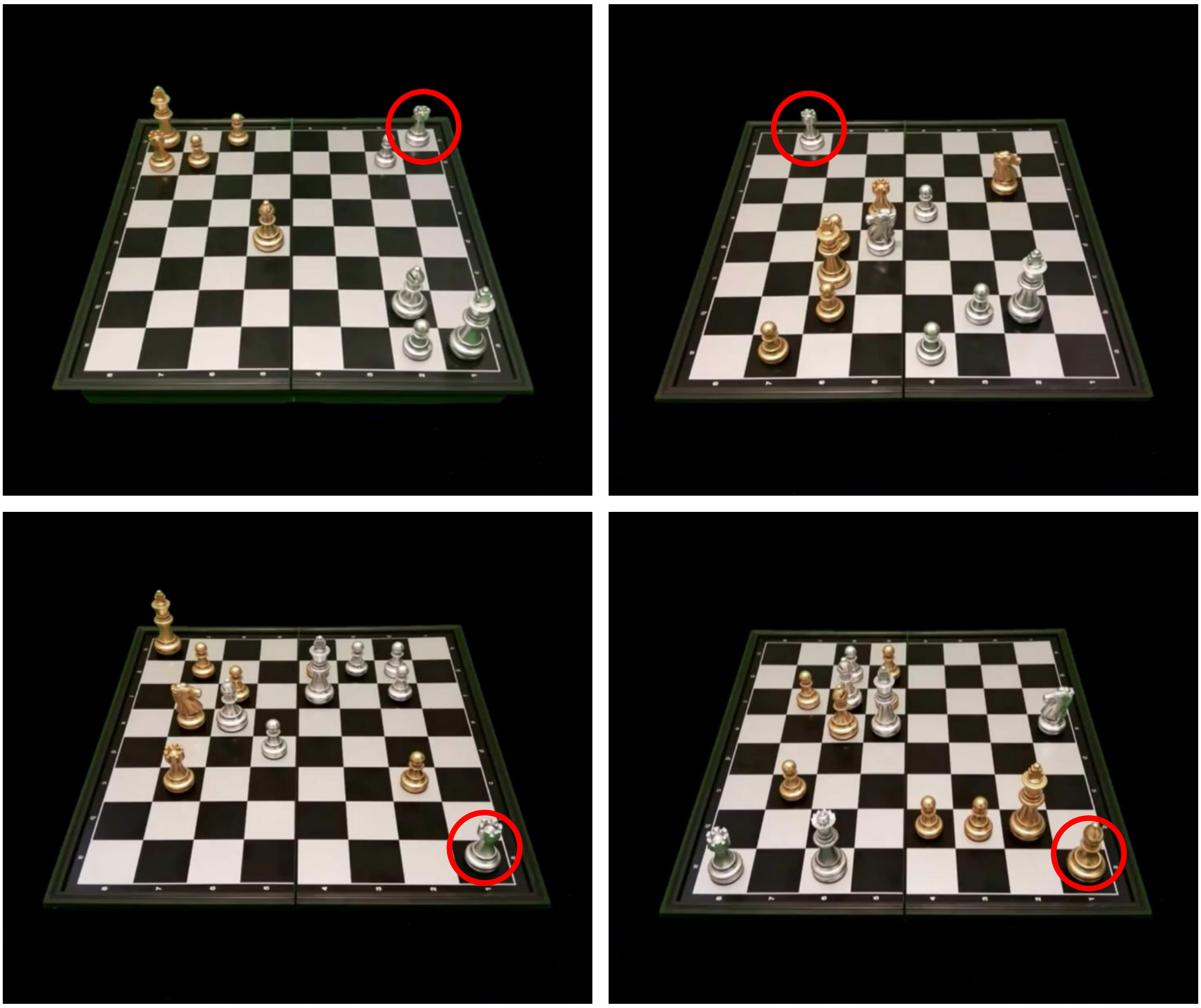
5 frames and the final grasp omitted. Action sequences were either 9, 12 or 15 frames long (frame rate 20 fps).

Each trial (see Figure 4.3) began with a fixation cross in the centre of the screen for 2000 ms, followed by a freeze frame of the chessboard for 2000 ms. The first frame of the video was then presented as a static image, showing the hand in the starting position of the action sequence. After 1000 ms, the action sequence began. The response screen followed immediately after the video, showing the chessboard with the arm omitted. Using their index finger, participants were asked to tap on the screen where they last saw the hand's index finger. The response screen remained present until a response was given, or until 10 seconds had elapsed.

Participants first completed a practice block of 8 trials. These were identical to experimental trials with the exception of the freeze frame, which was 9000 ms, allowing participants sufficient time to familiarise themselves with each chessboard. They then completed 3 blocks of 32 trials. Each block consisted of 8 chessboard setups, each presented three times; once for each action sequence length (9, 12 or 15 frames), and 8 catch trials. On each catch trial, the response screen was followed by a question: *'Which piece should the player move next?'* Participants responded by selecting one of six possible chess pieces (Queen, Rook, Bishop, Knight, Pawn or King; see Figure 4.2). The question remained on the screen until a response was given or 20 seconds had elapsed. Trials were randomly presented within each block.

Figure 4.2.

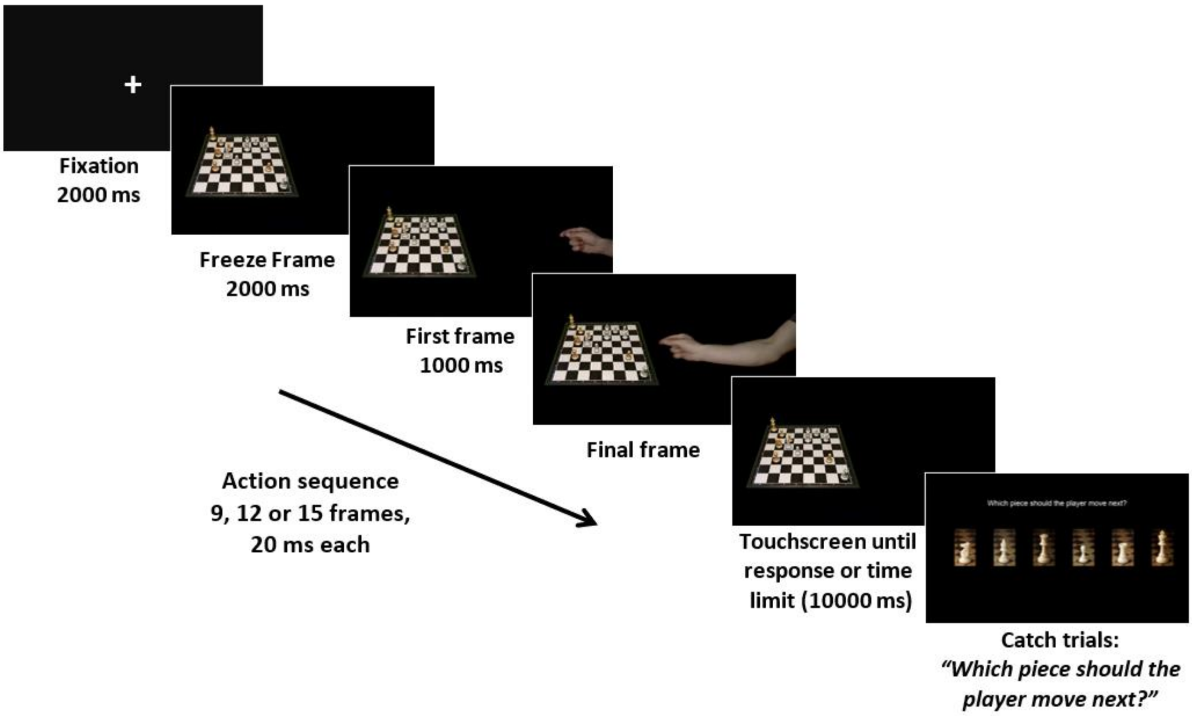
Example chessboard configurations (Experiment 3).



Target chess piece appeared at either the top (panels *a* and *b*) or bottom (*c* and *d*) of the board.

Figure 4.3.

Trial sequence (Experiment 3).



Data analysis

The experiment employed a 2 x 2 mixed design, with Chess Expertise (Player, Novice) and Target Location (Top, Bottom) as our independent variables. As in the previous studies, screen coordinates were converted into percentage coordinates, where 0%, 0% represents the top left-hand corner of the screen. Participants' selected screen coordinate (in percentage coordinates) of the tip of the index finger on each trial was subtracted from the real final screen coordinate. Analyses were conducted on this residual localisation error, providing a directional measure of how far participant's responses were displaced along the Y axis. An accurate response would produce a value of 0%. Positive and negative values denote upward and down displacements respectively.

4.1.2 Results

Participants' scores on the Checkmate Puzzles pre-test are displayed in Figure 4.4. The mean scores for participants recruited as Chess Players and Novices were 8 and 1, respectively. This difference was statistically significant ($t(48) = 15.82$, $p < .001$).

Catch trial performance was generally low ($M = 27\%$, $SD = 24.88$), although scores were significantly higher for Chess Players ($M = 39\%$, $SD = 28.21$) than for Novices ($M = 18\%$, $SD = 15.17$), $t(48) = 3.33$, $p = .002$, $d = 0.93$.

One-sample t-tests revealed that displacements were significantly smaller than zero when the target was presented toward the top of the chessboard ($M = -3.07\%$, $SD = 3.25$, $t\ (49) = -6.68$, $p < .001$, $d = .95$) and toward the bottom of the chessboard ($M = -3.10\%$, $SD = 3.25$, $t\ (49) = -6.75$, $p < .001$, $d = .95$), indicating an overall downward response bias.

A two-way mixed ANOVA was conducted to examine the effect of Chess Expertise (Player, Novice) and Target Location (Top, Bottom) on displacements. We predicted that in the Chess Players' group, perceptual judgments would be displaced toward the expected trajectory; that is, in the direction of the target chess piece, situated at either the top or bottom of the chessboard, whereas no such bias should occur in the Novice group. The analysis revealed no main effect of Expertise ($F\ (1, 48) = 0.317$, $p = .576$, $hp^2 = 0.007$), no main effect of Target Location ($F\ (1, 48) =$

0.013, p = .909, hp^2 = 0.00), and no interaction (F (1, 48) = 0.00, p = .996, hp^2 = .000).

Figure 4.4.

Distribution of scores from the Checkmate Puzzles pre-test.

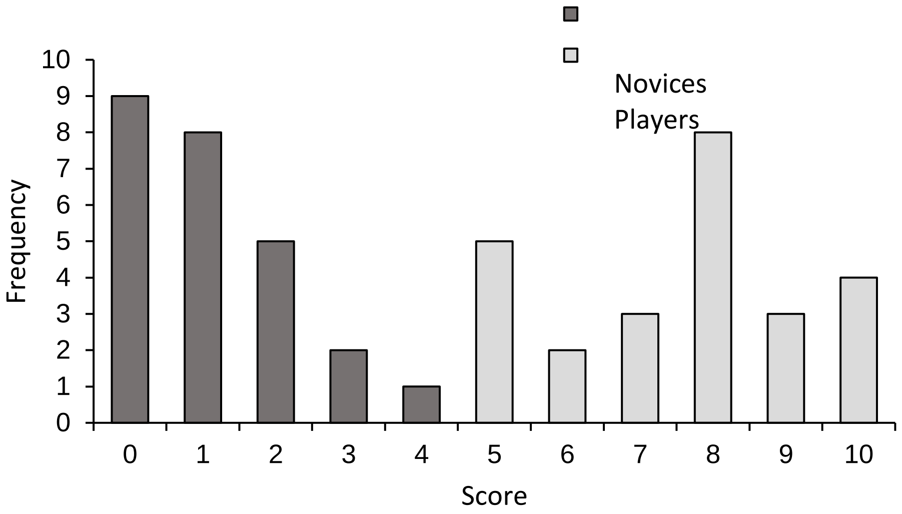
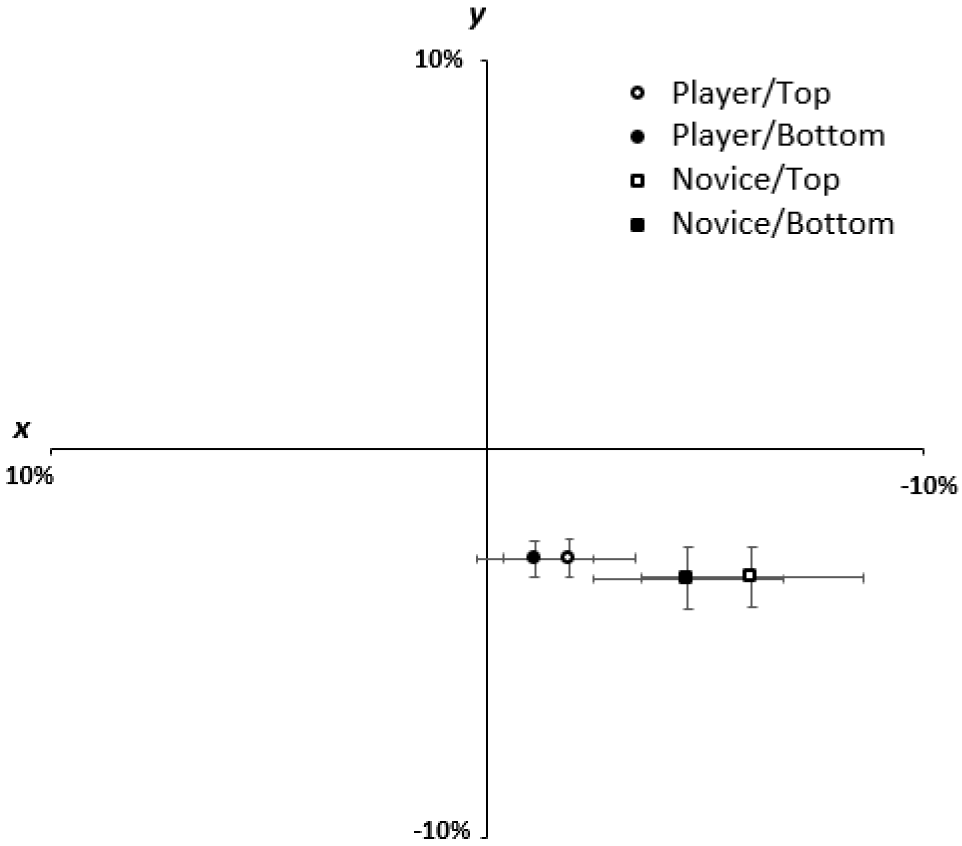


Figure 4.5.

Experiment 3 results.



The interaction between task condition (Player/Novice) and target-object location in Experiment 3. The difference between the hand's real final position and participants' selected final position (in percentage coordinates) is plotted for the X and Y axis. The centre of each plot represents the real final position on any given trial (0px difference on each axis). Error bars show the standard error.

Discussion

The current chapter investigated whether action goal predictions are guided by participants' high-level expertise. Chapter 2 provided evidence that action predictions are derived, at least in part, from object affordances. Expert chess players, through repeated exposure to the objects and rules of chess, develop domain-specific 'knowledge structures' which allow them to quickly and efficiently recognise individual chess pieces, as well as their associated functions (Bilalić et al., 2010). Such expertise would enable top-down predictions regarding what might occur next within a game of chess. We therefore hypothesised that chess players' top-down predictions, generated based on their prior knowledge of the objects in play and the rules of their movements, would bias the perception of an observed action towards the target chess piece. In contrast, novice players, who lack the the required knowledge structures, would have no expectations as to the object the actor might be reaching for, and would therefore exibit no perceptual bias.

We showed two groups of participants – skilled chess players or novice players – a series of short videos of a hand reaching towards a chessboard, set up in a checking configuration. Part-way through the action, the hand disappeared and participants indicated the index finger’s final position on a touchscreen. We predicted that prior knowledge of the rules of chess would lead to a predictive bias in the perception of a subsequent action towards the target chess piece. However, we failed to find evidence for a predictive bias within the current paradigm.

One reason for the absence of any observable effects may relate to the problems we had recruiting "expert" chess players online. We had originally intended to recruit highly skilled players via national chess clubs. However, the response to recruitment was exceptionally slow and uptake low. Our alternative approach was to screen participants via MTurk, using a 'Checkmate Puzzles' pre-test, comprised of 10 'mate-in-one' test questions of varying difficulty. Despite the use of the pre-test, analysis of catch trial performance suggested that participants at both high and low levels of expertise did not engage with the task. Nevertheless, it is possible that participants in the skilled group simply lacked the level of expertise required to automatically activate action representations afforded by the objects when passively viewing the chess configurations in the main experiment. This raises the question, what level of expertise are required for action predictions to be generated automatically? The international chess Elo scale is an interval scale with a theoretical mean of 1500 and standard deviation of 200 (Elo, 1978). Beginners typically have a rating of around 500, while Grand Master have ratings over 2500. Players with a rating of 2000 Elo points or more are classed as experts. The use of Elo ratings would be a more accurate way to determine a player's ability level. Previous studies have used eye-tracking to determine whether experts focus on the relevant chess pieces of each configuration compared to novices (Bilalić et al., 2010). A lab-based study utilising this technique would not only allow us to assess whether or not participants are paying attention to the task stimuli, but would also indicate if experts focus on the relevant target object compared to novices, whose eye movements one would expect to be more random.

Another possible explanation is the complexity of the task stimuli. Given that the task was run online via touchscreen devices, it is possible that the stimuli were not clear enough to view on smaller screens. The experiment was only operable on Apple devices running iOS, with a minimum screen resolution of 1024 x 768 pixels. This eliminated the possibility of participants using iPhones, and meant that the smallest useable device was an iPad mini with a screen size of 7.9" (a comprehensive list of iOS devices and their respective screen resolutions can be found at: https://www.ios-resolution.com). While this screen size may have been sufficient for our earlier online experiments, in which a perceptual bias was observed (see Chapter 2), it may not be sufficient for the current task, which requires more intricate object recognition. One solution for this, besides using a larger a screen, would be to use simpler checking configurations. Kiesel and colleagues (2009) tested whether chess expertise improve perceptual processing to an extent that allows complex visual stimuli to unconsciously bias behaviour. Participants - expert or novice players - were asked to judge whether a 3 x 3 chess configuration entailed a checking configuration. Each display was preceded by a masked prime configuration that either represented a checking or non-checking configuration. Results showed that chess players, but not novices, were faster to respond when the prime and target displays were congruent (both checking or both nonchecking) rather than incongruent. This suggests that chess experts can process chess configurations even when presented unconsciously. Even so, the stimuli used in Kiesel et al.'s experiment, and many others investigating chess expertise (e.g., Bilalić et al., 2010; Bitensky et al., 2014; Charness, 2001; Reingosld et al., 2001), are two-dimensional images that represent, while bearing little resemblance to, real-life objects. In contrast, the three-dimensional characteristics of the real objects used in our experiment, and the context in which they occur, would likely take longer to process. The use of simplified checking configurations, with fewer pieces in play, could help to overcome this problem.

Finally, it should be considered that the unnaturalistic perspective of the chessboard may not have allowed players sufficient time to process the board configuration and identify the target chess piece. Chess matches or puzzles, whether in person or online, are typically presented from a first-person perspective. Here, the chess board and the action were observed from a third-person perspective, arguably making it more difficult for the observer to anticipate the next move. This design was not entirely through choice, but was constrained by the limited resources available during the COVID-19 pandemic. Unfortunately, it was not possible to record new first-person action sequences.

The current chapter aimed to test the degree to which action goal predictions are guided by object expertise. However, using the current paradigm, we were unable to detect any observable differences between skilled and novice chess players. We suggest that this can be explained by participants' poor engagement with the online task and the low ability level of the chess players. Although there was a clear difference in the ability levels of the two groups, as shown by the pre-test scores, this difference may not have been sufficient to yield significant differences in action judgements, as the chess players were not true "experts". The unnatural perspective of the chessboard might have increased the difficulty of the task and further weakened the effect.

Chapter 5: Do object affordances inform motor action preparation when viewing goal-directed actions?

Previously, we demonstrated that action prediction is driven, at least in part, by the affordances of available objects. In Chapter 2, we found that the presentation of a prime object led to a predictive bias in the perception of a subsequent action towards a functionally related target object. This finding supports predictive processing views that action perception is guided by goal predictions, based on prior knowledge about the context in which the action occurs (Friston et al., 2011; Kilner et al., 2007). In Chapters 3 and 4, we investigated the degree to which this predictive bias was influenced by either knowledge about the actor, or expertise in the particular object. The present chapter changes focus to explore the low-level mechanisms of these anticipatory representations, namely by investigating the possible involvement of the motor system in the anticipatory representation of action goals associated with objects.

Action understanding is known to involve the decoding of both the visual kinematics of the observed action and the action goal. The action understanding account of mirror neurons assumes that the motor system automatically activates when observing the actions of others, and that this direct matching allows for inferences regarding the goal of an action (Rizzolatti et al., 2001). In other words, the actor's goal is embedded in the visual kinematics of their actions. However, it has since been argued that action simulation alone might not be sufficient to unambiguously identify the goals of others' actions, since most actions can be performed to achieve a number of different goals, and different actions can be performed to achieve the same goal (Jacob & Jeannerod, 2005; Kilner et al., 2007).

Other theories suggest that kinematic processing is guided by prior activation of predictions about the actor's goal, which are driven by non-motoric contextual information, such as language (Hudson, Bach, et al., 2018), gaze (Bayliss et al., 2007) or available objects (Bach et al., 2014). Recent predictive processing models suggest that inferred goals generate top-down predictions of what will be perceived next, biasing perception towards the expected action (Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016). This perceptual bias has been demonstrated using a modified representational momentum paradigm by ourselves (see Chapter 2) and others (Hudson, Bach, et al., 2018; Hudson, McDonough, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; McDonough et al., 2019).

Evidence suggests that actions are not only planned around goals, but are also perceived as being goal-directed (Decroix & Kalénine, 2019; Wurm & Schubotz, 2017). DeCroix and Kalénine (2019) investigated whether visual attention would be preferentially drawn to the visual kinematics (i.e., hand grip) or action goal information (i.e., orientation of the object) when processing others' actions. In a visual search task, participants were shown four pictures of an object-directed action with a particular object (e.g., a pencil). The object-directed actions could display a "correct grip and correct goal", a "correct grip only", a "correct goal only", or both "incorrect grip and incorrect goal". For example, using a power gip to write with an upright pencil is atypical, but does not prevent writing. In contrast, using a precision grip to write with an upside-down pencil does not allow writing, but the grip configuration is typical. Thus, the grip and goal varied independently from one another. Participants were asked to select the picture that displayed the correct action according to the typical use of the object. In this example, the correct target would display a precision grip with an upright pencil. Gaze movements were used to evaluate the extent to which visual attention was preferentially driven towards grip or goal-related information before the identification of the target. Results showed that visual attention towards the goal-related distractor initially increased in comparison to grip-related distractors, then decreased. This is in line with predictive approaches of action understanding, which assume that observers first make a prediction about the actor's goal before verifying this prediction using the visual kinematics of the action.

It has been suggested that object knowledge, including how objects are used and what they are used for, is a primary contributor to how goals are inferred and actions are interpreted (Bach et al., 2014). Gibson's (1979) influential theory of affordances suggests a tight link between perception and action whereby people not only perceive the overt physical properties of objects, but also the possible actions they can afford, based on the motor capabilities of the person. The concept of *affordances* has since been widely expanded on and the term is used more broadly to describe the properties of an object that elicit the automatic activation of mental representations for action within a perceiver (Ellis & Tucker, 2000; Phillips & Ward, 2002). Here, *properties* refers to the manipulable aspects of an object that are associated with its typical use. For example, the handle of a mug for drinking (Chong & Proctor, 2020). In this view, affordances are represented in the brain such that the representation of a visual object includes not only a description of its visual properties, but also encodings of the actions associated with its canonical use (Ellis & Tucker, 2000).

Humans represent objects in terms of both the goals that can be achieved with them (*function* knowledge) and the specific actions required to achieve these goals (manipulation knowledge). This knowledge makes a contribution to action understanding by allowing observers to infer the goals someone wants to achieve with an object (via function knowledge) and, crucially, to *predict* the actions this person would need to carry out to achieve these goals (via manipulation knowledge) (Bach et al., 2015). Nicholson, Roser and Bach (2017) conducted an fMRI study in which participants watched everyday instrumental actions (e.g., posting a letter) while attending to either the movements performed, the objects used, or the actions' goal, while visual stimulation was kept identical. They found activation in the goal-identification task overlapped directly with object- but not movement-related activation. This suggests that goal understanding draws on the same neural networks as object identification. Furthermore, movement-related regions were activated only when the goal was unclear. Thus, while the goal task activated regions involved in object identification, it did not activate the motor regions engaged by the movement task. Taken together, these findings suggest that objects, rather than movements, provide the key information about the goals of others' actions.

It has been widely demonstrated that there is a direct link between object perception and motor performance. The mere sight of an object activates a representation of the action that can be performed on it, even in the absence of explicit intentions to act (Chao & Martin, 2000; Craighero et al., 1998; Ellis & Tucker, 2000; 2004; Phillips & Ward, 2002; Tucker & Ellis, 1998, 2001). Early evidence of an affordance effect stemmed from Tucker and Ellis's (1998) research using a stimulus-response compatibility (SRC) paradigm. A typical SRC task involves testing whether behavioural responses are faster when the spatial location of the target stimulus is compatible, relative to incompatible, with the required response (e.g., a left-side stimulus is compatible with a left, but not a right, hand response, and is therefore responded to faster; Hommel, 1997). In their seminal study, Tucker & Ellis (1998) found that the handle orientation of a saucepan, though irrelevant to the task, could influence participants to give faster responses when the response hand and the handle orientation were matched. In further experiments, participants identified high or low auditory tones by mimicking a precision grip (PG) or whole-hand grip (WHG), while simultaneously viewing real objects that would typically be grasped using a pinch or WHG. They found an interaction between the required response and the grip afforded by the observed object. PG responses tended to be faster on trials displaying objects compatible with a PG than those compatible with a WHG. Conversely, WHG responses tended to be faster on trials displaying objects compatible with a WHG than those compatible with a PG (Ellis & Tucker, 2000). These behavioural results, along with many others (Borghi et al., 2007; Bub et al., 2008; Tipper et al., 2006; Tucker & Ellis, 2001) suggest that when we see an object, our motor system prepares for the action that object affords, independent of intentions, giving a reaction time advantage to the congruent motor response.

Neurophysiological studies have shown that observing objects activates possible actions to perform with them (Chao & Martin, 2000; Grafton et al., 1997; Grèzes et al., 2003; Valyear et al., 2007). Passive observation of objects, with no intention to interact with them, elicits activation in frontal and parietal regions associated with motor processes (Chao & Martin, 2000; Grafton et al., 1997; Grèzes & Decety, 2002; Grèzes et al., 2003; Kourtis et al., 2018; Valyear et al., 2007, 2012). Moreover, an fMRI study investigating the neural correlates of the behavioural affordance effects demonstrated by Tucker and Ellis (2000) found a strong correlation between the size of the affordance effect and neural activity in anterior parietal, dorsal premotor and inferior frontal cortex. These findings suggest that passive viewing of objects not only involves the processing of its visual properties and identity, but also induces a representation of actions that are afforded by the specific features of the object.

Evidence for the modulation of the motor system when passively viewing objects also comes from studies using transcranial magnetic stimulation (TMS) over motor cortex to elicit motor evoked potentials (MEPs) in specific muscles as quantification of cortico-spinal excitability at the time of stimulation (Bestmann & Duque, 2016). The size of the MEP recorded from a specific muscle is known to vary with increasing cortical preparation for relevant motor acts (Makris et al., 2011). Using this paradigm, Buccino and colleagues (2009) recorded MEPs from the right hand of right-handed participants while viewing objects oriented to the left or right, with a whole or broken handle. Results showed that MEPs were larger when the objects were oriented to the right side, but only when the handle was complete. This suggests that a right-handed action was being planned in response to the right orientation of the object.

Combining the behavioural methods of Ellis and Tucker (2000) and the TMS protocol used by Buccino et al. (2009), Makris, Hadar & Yarrow (2011) presented participants with pictures of objects that afforded either a PG or a WHG. Participants either made speeded PG or WHG responses to an orthogonal visual stimulus (Experiment 1) or passively observed the objects while receiving TMS (Experiment 2). MEPs were recorded from the hand muscles associated with either a PG (first dorsal interosseous; FDI) or WHG (abductor digiti minimi; ADM, and FDI). Results showed that reaction times were faster when the type of response (PG or WHG) and type of object ('pinchable' or 'graspable') were congruent, although this was only the case when the response stimulus was presented soon (400 ms) after the object image, but not at longer intervals (800 ms or 1200 ms). In partial support of these behavioural results, Experiment 2 found that "pinchable" objects produced larger FDI (index finger) MEPs compared to "graspable" objects. However, there were no significant differences for ADM (little finger) MEPs when viewing "graspable" objects compared to "pinchable" objects.These data imply that the physical properties of objects automatically activate specific motor codes. However, the effect is rapid and short-lived.

In a follow-up study (Makris et al., 2013), participants viewed real “pinchable” and “graspable” objects while receiving TMS over their left hemisphere hand motor area, and MEPs were recorded from the FDI and ADM at three different SOAs (150, 300 and 450 ms) relative to the visual onset of the prime object. Analyses showed significant interactions between type of muscle and prime object, both of which were in the direction of concordant effects of priming upon precision (FDI) and whole-hand (ADM) grip responses. For the 300 ms SOA, FDI MEPs were significantly larger for objects associated with a PG than for “graspable” objects, with the corresponding reverse trend suggesting larger ADM MEPs for “graspable” objects comparing to “pinchable” objects, although this did not quite reach statistical significance (p = 0.16). However, for the 450 ms SOAs, FDI MEPs were significantly larger for objects affording a PG compared to WHG objects, and ADM MEPs were significantly larger for objects affording a WHG compared to PG objects.

Others have found early (150 ms) and dissociable muscle synergies evoked by the presentation of action-object pairs priming a PG or WHG affordance (Bartoli et al., 2014). Participants were presented with a video of a hand mimicking a PG or WHG, followed by a static image of an object offering either a PG or WHG affordance. Single-pulse TMS was triggered 150 ms from the onset of the image and

MEPs were recorded from the abductor pollicis brevis (APB; thumb muscle) and ADM of the right hand. They found increased muscle activation in the APB when observing a PG affordance, whereas observation of a WHG affordance increased motor recruitment of the ADM. Thus, action-object observation appears to elicit an early and specific pattern of motor activation, matching the affordances of the objects.

Taken together, the evidence suggests that (a) humans identify their own actions and the actions of others as goal-directed, and (b) objects automatically generate compatible motor plans, giving a reaction time advantage to the congruent motor response. If we make predictions about others' actions based on our inferences about their goals, and objects are a source of these inferences, observing an actor pick up an object (e.g. a cigarette) should lead us to make a prediction about what they intend to do next (e.g. smoke it). The following experiments aimed to explore the possible involvement of the motor system in the anticipatory representation of action goals associated with objects.

In our initial TMS experiment, Experiment 4a, participants observed an actor grasp an object (the 'prime'; e.g. a cigarette) with their left hand before ambiguously reaching towards two other objects in front of them with their right hand. One of these objects was functionally related to the prime (e.g., a lighter), while the other was functionally unrelated (e.g., a notepad). Measuring motor-evoked potentials (MEPs) in the right hand of the observer at the point the actor reached out towards the two objects, we expected to find anticipatory muscle activation for the grip afforded by the functionally related object (i.e. the lighter). Specifically, we expected ADM (little finger) and FDI (index finger) muscle activation when the functionally-related target was a larger object (e.g., a notepad), as these muscles are typically recruited for a WHG. Conversely, when the functionally-related target was a small object (e.g., a lighter), we expected activation in the FDI, but *not* the ADM, muscle because a PG does not typically recruit the little finger. As we were unable to implement this lab-based experiment due to COVID-19 restrictions, we adapted the experimental design for use online as a behavioural study.

In Experiment 4b, participants heard a high or low tone at the end of each action sequence. Using their keyboard, their task was to respond with the index finger of one hand when they heard a high tone, or *both* the index and little finger of the other hand when they heard a low tone. We predicted that participants would be faster to respond when the functionally related target object required a grip that was congruent with the correct key-press response. For example, when the functionally related object required a WHG (e.g., a notepad), reaction times should be faster when participants heard a low tone, versus a high tone, as this response required the same muscles as recruited for a WHG.

In a simplified version of this task, Experiment 4c presented participants with a static image of a small PG object (e.g. a key) or a larger WHG object (e.g. a cup). As before, they heard a high or low tone and were asked to respond using either their index finger or their index and little finger together. We predicted that participants would be faster to respond when the object required a grip that was congruent with the correct key-press response (e.g., cup – low tone). This is because observing an object activates the muscles recruited when using that object. In other words, viewing an object that requires a WHG (e.g. a mug) activates the muscles recruited when using that object (i.e. the little finger and index finger to grip the mug).

In Experiment 4d, we replaced the previous PG objects with more specific *'pinch grip'* objects (a matchstick, penny, or chess piece). Participants were presented with a static image of a small pinch-grip object or a larger WHG object. They then heard a low tone, to which they responded as quickly as possible using their little finger. Occasionally, they heard a different sound (white noise) which required them to respond with their index finger. We predicted that participants would be faster to respond to the tone when the object presented required a WHG rather than a pinch grip. This is because a WHG recruits the muscles in little finger, whereas a pinch grip does not. This activation therefore facilitates the little-finger response to the tone. In contrast, we expected to see no such facilitation when the object required a PG, as this grip only requires the index finger.

5.1 Experiment 4a: Planned TMS experiment

What follows in an outline of the experiment that was initially intended for this stage of the project. However, due to the restrictions imposed by the university's COVID-19 measures, we were unable to execute this lab-dependent study, switching instead to an online approach utilising behavioural measures.

The proposed study aimed to investigate the way in which object knowledge contributes to our understanding of others' actions. Specifically, we sought to explore the possible involvement of the motor system in the anticipatory representation of action goals associated with objects.

In this lab-based experiment, participants observed an actor grasp an object (the 'prime'; e.g. a cigarette) with their left hand before ambiguously reaching towards two other objects in front of them with their right hand. One of these objects was functionally related to the prime (e.g., a lighter), while the other was functionally unrelated (e.g., a notepad). If action perception relies on forward models predicting the future course of others' actions, as predictive coding models suggest (Sartori, Bucchioni, et al., 2013; Schippers & Keysers, 2011), and these predictions are based on object affordances (Bach et al., 2014), then observing someone pick up an object, like a cigarette, should generate a prediction about what they intend to do next (e.g. smoke it). When observing a subsequent reaching action that is ambiguous, we would therefore expect motor facilitation for the grasp afforded by the object most suitable to achieve the predicted goal.

Small objects, such as pens, tend to be handled with a PG, holding the object between the index finger and thumb. In contrast, larger objects, such as bottles, are held with a WHG between the palm and all of the fingers. Prior research suggests that the mere sight of an object facilitates the optimal grip response (Bartoli et al., 2014; Ellis & Tucker, 2000; Makris et al., 2013). It is not the grasping in general that is facilitated, but the specific grasp appropriate to the viewed object.

Using single-pulse transcranial magnetic stimulation (TMS) to induce motor-evoked potentials (MEPs) in the hand of the observer at the point the actor reaches out towards the two objects in front of them, we expected to find anticipatory muscle activation for the grip afforded by the functionally related object. Specifically, we predicted FDI (index finger) muscle activation when the functionally-related target object required a PG, and both ADM (little finger) and FDI activation when the related object required a WHG. These findings would demonstrate the activation of affordances in the total absence of any requirement to act.

5.2 Methods

Participants

Results of a priori power analysis (https://jakewestfall.shinyapps.io/pangea/) showed that a sample of 40 is required to achieve 80% power with a medium effect size (*d* = 0.6).

Design

The experiment used a within-subjects design with two conditions: the Goal condition (two functional objects placed on the table - one target and one distractor), and the Control condition (target and distractor objects replaced with two pieces of fruit). Target objects in the Control condition were matched to the Goal condition for grip type (i.e. one PG object and one WHG object).

MEPs were recorded from the ADM and FDI muscles of the right hand (of right-handed participants) during the observation of video clips showing goal-directed action sequences. Each participant's baseline corticospinal excitability was assessed by acquiring 10 MEPs while passively watching a white fixation cross on a black screen at the beginning and end of each session. Peak-to-peak amplitudes of the MEP from both the ADM and FDI muscles were measured and averaged separately for each condition. Our dependent variables were the MEP ratios, calculated using the participant's individual mean MEP amplitude during pre- and post-testing sessions as baseline values (MEP ratio = $MEP_{obtained}/MEP_{baseline}$).

Stimuli

Experimental stimuli consisted of 16 action sequences (resolution: 1920 x 1080), each consisting of 33 frames, proceeding in two frames jumps, each with a duration of 60 ms. The first 13 frames showed the actor reaching to grasp an object ("the prime," – either a cigarette or a pencil) with their right hand. The following 20 frames showed the actor reaching ambiguously towards two other objects with their left hand. One of these objects was functionally related to the prime (“the target,” e.g., cigarette – lighter), while the other was functionally unrelated (“the distractor,” e.g., cigarette – notepad). One object in each object pair required a PG and the other a WHG. Eight control trials were created by digitally replacing the target and distractor objects with two pieces of fruit, one big (an apple) and one small (a grape). A second set of stimuli was created by mirror-reversing all 24 videos so the actor was reaching with the opposite hand.

Procedure

The experiment was administered using MATLAB. Each experimental trial (see Figure 5.1) began with a fixation cross in the centre of the screen for 2000 ms. The first frame of the video was presented as a static image, showing the actor with one hand in a prone position on the table. After 2000 ms, the action sequence began and the actor reached out with the other hand. The thirteenth frame of the action sequence, depicting the actor grasping the prime object, was frozen for 2000 ms to allow participants time to process the prime object. MEPs from the ADM and FDI muscles of the right hand (of right-handed participants) were recorded at the final frame of the video.

TMS Simulation and MEP recording

Neural activity was modulated using single-pulse transcranial magnetic stimulation (TMS) and measured using electromyography (EMG). TMS stimulation was applied via a MagStim 200 stimulator with a MagStim D70^2 figure 8 coil, air cooled and equipped with a force sensor. Coils were attached and controlled using the Axilum Robotics TMS-Robot and Localite Neuronavigation software, providing an automated procedure, with continual control of the position, orientation and pressure contact of the stimulation coil on the head. The TMS coil was positioned in correspondence with the optimal scalp position (OSP), defined as the position at which the stimulation of a slightly suprathreshold intensity consistently produces the largest MEP from both the ADM and FDI muscle. Each participant's resting motor threshold (rMT) was determined as the minimum intensity that reliably induced MEPs ($\geq 50\ \mu V$ peak-to-peak amplitude) in the relaxed muscle in 5 out of 10 consecutive trials (Rossini et al., 1994). Stimulation intensity during the recording session was 110% of the rMT.

EMG was recorded using the ActiveTwo system by BioSemi using active electrodes. Flat-Type active electrodes were used to measure EMG from the first dorsal interosseous (FDI) and abductor digiti minimi (ADM) muscles of the right hand.

Data analysis

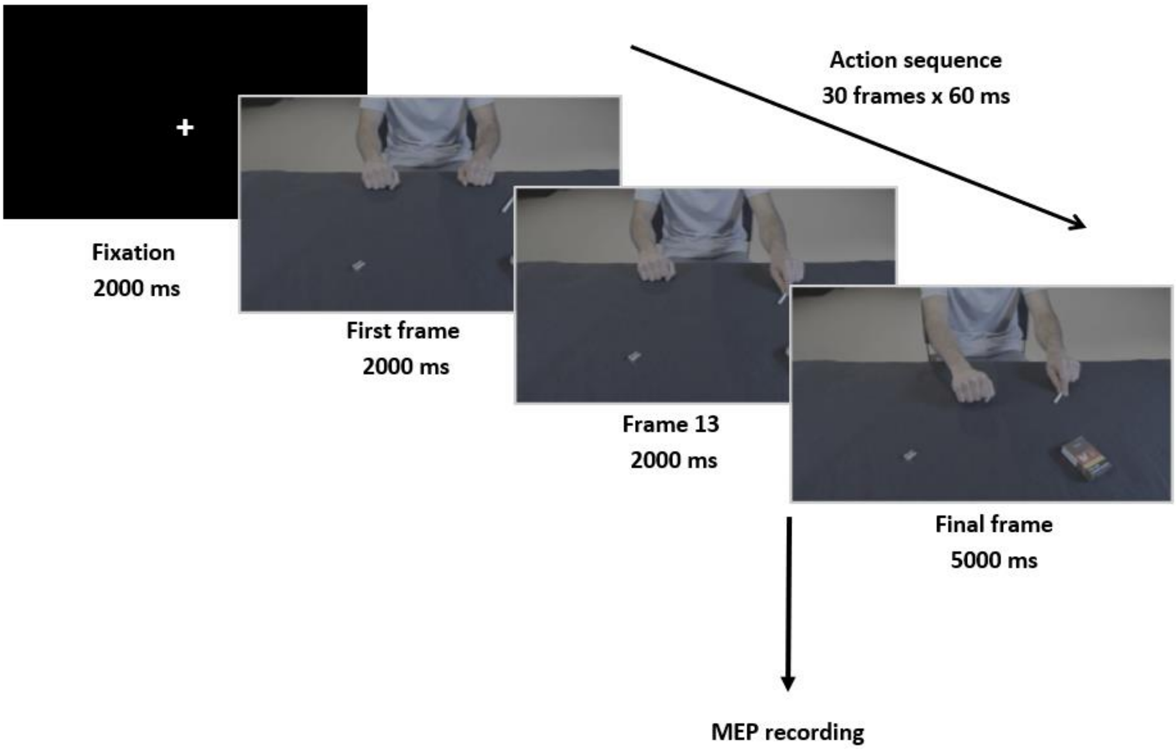
MEP amplitudes deviating more than 2 SDs from the mean for each condition, and trials contaminated by muscular pre-activation were excluded as outliers. Mean MEP ratios were calculated from the ADM and FDI muscles for each condition (Goal, Control) and each target grip (PG, WHG). A repeated-measures ANOVA was calculated on the MEP ratios (FDI and ADM), with Condition (Goal, Condition) and Target Grip (PG, WHG) as our independent variables. To allow for this factorial analysis, the two items in the control condition were dummy coded as either targets or distractors. In half of the control trials the ‘target’ was presented on the left side of the table, and in half it was presented on the right side.

Hypotheses

We predicted that we should observe a Condition x Target Grip interaction, such that an effect of target grip is present only in the experimental condition but not in the control condition. This is verified using pairwise comparisons on FDI and ADM MEP ratios for each Target Grip within each condition. As FDI is recruited for both PG and WHG, no MEP modulation is expected in terms of the target grip. However, in the experimental condition, we expected greater ADM activation when the target object afforded a WHG.

Figure 5.1.

Trial sequence (Experiment 4a).



5.3 Experiment 4b

The previous study aimed to explore the involvement of the motor system in the anticipatory representation of action goals associated with objects. As we were unable to implement this lab-based experiment due to COVID-19 restrictions, we adapted the experimental design for use online as a behavioural study, using RTs as an index for motor preparation. This approach not only aimed to address the same research question as in Experiment 4a, but also provided the opportunity to test whether reaction times can be reliably used to measure affordances.

In this online experiment, participants watched a series of short videos. Each video sequence began with an actor reaching to grasp an object ("the prime," – either a cigarette or a pencil) with one hand before reaching towards two other objects using their other hand. One of these objects was functionally related to the prime (“the target,” e.g., cigarette – lighter), while the other was functionally unrelated (“the distractor,” e.g., cigarette – notepad). One object in each object pair required a PG and the other a WHG. The video ended with the actor's hand mid-motion, in a neutral grip, at an equal distance between the two potential target objects. On the final frame of each video, participants heard a high or low tone. Using the keyboard, participants' task was to respond with the index finger of one hand when they heard a high tone, or *both* the index and little finger of the other hand when they heard a low tone. If a PG recruits the index finger and a WHG recruits *both* the index and little finger, the activation of an action representation elicited by the functionally related object should facilitate the congruent motor response. Similar studies (e.g., Ellis & Tucker, 2000; Makris et al., 2011) have measured grip responses using a specifically designed grasp-mimicking device. Due to the constraints of online testing, we were unable to implement this same design. In addition, we included a control condition in which the target and distractor objects were replaced by two pieces of fruit (one big and one small). If the affordance theory stands, an effect of grip type should only be found for objects with action significance.

Our primary prediction was that participants would be faster to respond when the functionally related target object required a grip that was congruent with the correct key-press response. For example, when the functionally related object requires a precision grip (e.g., a pencil sharpener), reaction times should be faster when participants hear a high tone, compared to a low tone, and must respond with their index finger, as this is the muscle recruited by a PG. In other words, when the actor grasps the prime object (e.g. a cigarette), the observer makes a prediction about the actor's intention (to smoke it using the lighter). This prediction facilitates execution of the grip recruited when acting to achieve that goal (e.g. a PG to pick up the lighter).

Further to this, we mirror-reversed each of the video sequences to show the actor reaching with the opposite hand, to test whether motor facilitation was driven by an imitative "mirroring" mechanism, whereby participants are faster to respond when the observed hand is congruent with the executed hand, or a spatial compatibility effect (the 'Simon effect'; Simon, 1969), whereby RTs are faster when the executed hand and target location are congruent (e.g., right-hand, right stimulus).

5.4 Method

Details of the experimental design and analysis were pre-registered on AsPredicted.org (available at: https://aspredicted.org/4vn2h.pdf)

Participants

Participants were 59 right-handed volunteers (24 male, 35 female; mean age: 31 years, range: 20 to 50 years), recruited online via Prolific Academic. Participants were given £3 for taking part.

Stimuli

Experimental stimuli consisted of 16 action sequences (resolution: 1920 x 1080), each consisting of 33 frames, proceeding in two frames jumps, each with a duration of 60 ms. The first 13 frames showed the actor reaching to grasp an object ("the prime," – either a cigarette or a pencil) with their right hand. The following 20 frames showed the actor reaching ambiguously towards two other objects with their left hand. One of these objects was functionally related to the prime (“the target,” e.g., cigarette – lighter), while the other was functionally unrelated (“the distractor,” e.g., cigarette – notepad). One object in each object pair required a PG and the other a WHG. Each video ended with the actor's hand mid-motion, in a neutral grip, at an equal distance between the two potential target objects, (see Figure 5.2). Eight control trials were created by digitally replacing the target and distractor objects with two pieces of fruit, one big (an apple) and one small (a grape). A second set of stimuli was created by mirror-reversing all 24 videos so the actor was reaching with the opposite hand.

Procedure

The experiment was administered using Inquisit Web. Each trial (see Figure 5.2) began with a fixation cross in the centre of the screen for 2000 ms. The first frame of the video was presented as a static image, showing the actor’s hand in a prone position on the table. After 2000 ms, the action sequence began. The thirteenth frame (out of 20) of the action sequence, depicting the actor grasping the prime object, was frozen for 2000 ms to allow participants time to process the prime object. On the final frame, a high (1000 Hz) or low (500 Hz) tone played for 200 ms.

Using the keyboard, participants pressed the ‘A’ key with their left index finger when hearing the high tone, or the ‘G’ and ‘L’ keys together with their right hand when hearing the low tone (see Figure 5.3). Immediate feedback was given by presenting either ‘correct’ or ‘error’ in green or red text, respectively, in the centre of the screen for 1500 ms. The final frame of the action sequence remained on the screen until a response was given, or until 5000 ms had elapsed.

Each experiment presented four randomised iterations of *Target Type:* The target associated with each prime object was either large (WHG), small (PG), or unrelated (fruit), *Tone:* In half of the trials participants heard a high tone and in half they heard a low tone, *Hand:* In half of the trials the actor reached out using their left hand and in half they reached out using their right hand, *Target Location:* In half of the trials the target object was presented on the left-hand side of the table, and in half it was presented on the right-hand side. This produced 96 experimental trials. The experiment contained sixteen catch trials. Catch trials were identical to experimental trials, except that after responding to the tone, participants were presented with a multiple-choice question: *“Given the object he has just picked up, which other object is he most likely to pick up?”* Participants were given a choice for four possible answers, two of which were the target and distractor objects for that trial, and responded using the A, G, H or L key. The question remained on the screen until a response was given, or 20 seconds had elapsed. Participants completed the experiment in two blocks of 56 trials. Half of participants began with the left-hand reach videos and half with the right-hand reach videos. Trials were randomly presented within each block.

Training procedure

Before beginning the main experiment, participants first completed a training phase, comprised of two blocks, to familiarise them with the task. The first training block (10 trials) was designed to get participants used to hearing the tones and using the keyboard. Each trial began with a fixation cross on the screen for 1500 ms, followed by a high (1000 Hz) tone or a low (500 Hz) tone, with a duration of 200 ms. They were instructed that they should respond by pressing the ‘A’ key with their left index finger when hearing the high tone, and the ‘G’ and ‘L’ keys together with their right hand when hearing the low tone, keeping both hands on the keyboard for the duration of the task. Immediate feedback was given by presenting either ‘correct’ or ‘error’ in green or red text, respectively, in the centre of the screen. The fixation cross remained present until a response was given, or until 10 seconds had elapsed.

The second training block (4 trials) was designed to familiarise participants with the visual stimuli and catch trial questions. Each trial began with a fixation cross in the centre of the screen for 2000 ms, followed by an action sequence. Participants were then presented with a multiple-choice question: *“Given the object he has just picked up, which other object is he most likely to pick up?”* and asked to respond using the A, G, H or L key. The question remained on the screen until a response was given, or 20 seconds had elapsed.

Data analysis

Incorrect responses were removed from the analyses, as were responses faster than 200 ms or slower than two standard deviations from each participant's overall mean RT. Mean RTs, in milliseconds, were calculated from the remaining data for each participant, in each condition. For low-tone trials, on which the correct response required two simultaneous key-presses, an average RT of the two keys was calculated. Condition means were entered into three 2 x 2 within-subjects ANOVAs. Our independent variables were: (1) *Object Condition:* 'Goal'; two objects placed on the table - one target and one distractor, 'Control'; target and distractor objects replaced by two pieces of fruit. Target objects in the control condition were matched to the Goal condition for grip type (i.e. one PG and one WHG). (2) *Grip Congruency:* the grip recruited when using the 'target' object, and the response required by the tone were either congruent or incongruent. (3) *Hand Congruency:* the hand the actor reached out with and the required response hand were either the same or different. (4) *Spatial Congruency:* the response hand and target location were either compatible or incompatible.

Participants were excluded from the analyses if their accuracy in the catch trials was less than 70% suggesting they were not fully processing the prime or target objects. Four participants did not provide complete data sets. This left us with a sample size of 29. However, as our power analysis indicated that a sample size of 51 was required to yield sufficient statistical power, analyses were run both before and after exclusions.

Table 5.1.

Object pairs for Experiment 4b.

Goal Condition			Control Condition		
Prime	**Target**	**Distractor**	**Prime**	**Target**	**Distractor**
Cigarette	Lighter (PG)	Notepad (WHG)	Cigarette	Grape (PG)	Apple (WHG)
Cigarette	Cigarette packet (WHG)	Pencil Sharpener (PG)	Pencil	Apple (WHG)	Grape (PG)
Pencil	Pencil Sharpener (PG)	Cigarette packet (WHG)			
Pencil	Notepad (WHG)	Lighter (PG)			

***Figure 5.2*.**

Trial sequence (Experiment 4b).

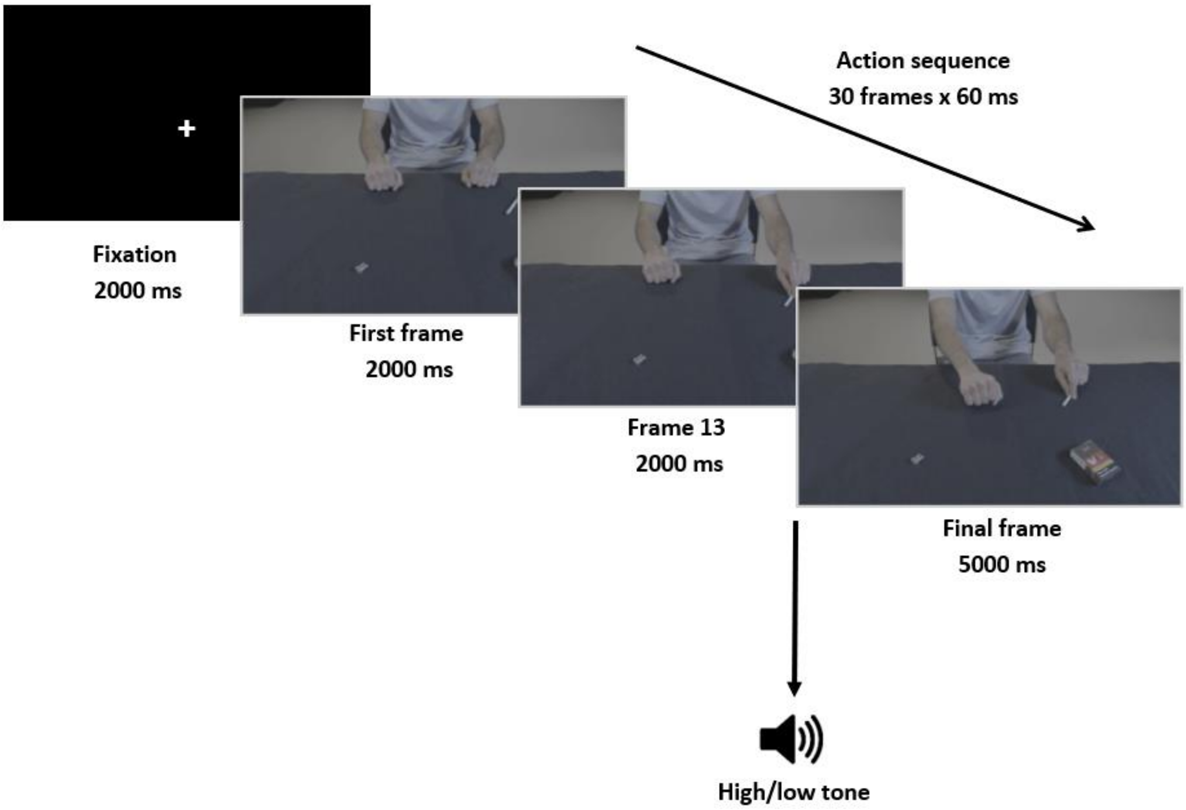
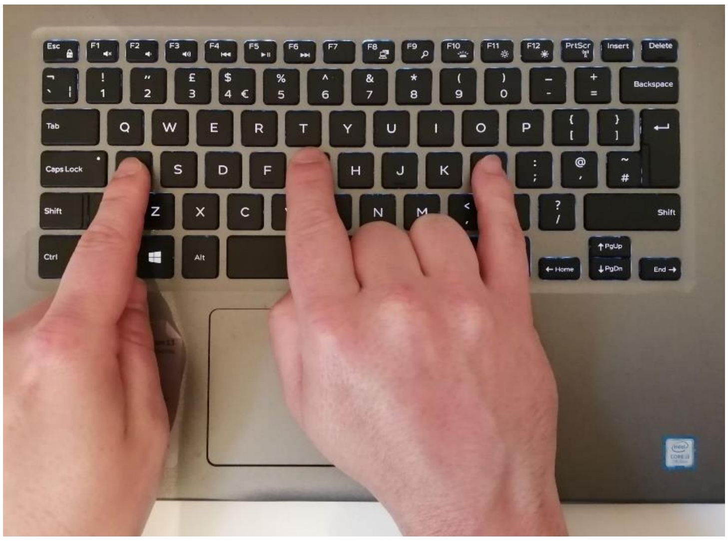


Figure 5.3.

Image accompanying written instructions (Experiment 4b).



5.5 Results

Following the protocol set out in our experiment pre-registration, participants were removed from our analyses if their accuracy on catch trials fell below 70%. However, this left us with an inadequate sample size (*N* = 29) to yield sufficient statistical power. We therefore conducted the same analyses using the full data set (*N* = 55).

Object Condition x Grip Congruency

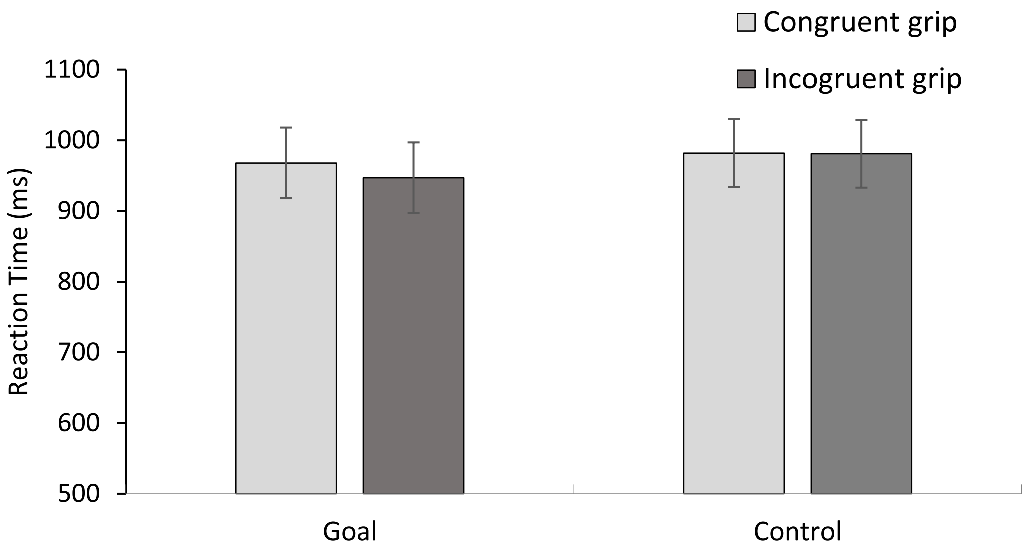
Our primary prediction was that participants would be faster to respond when the functionally related target object required a grip that was congruent with the correct key-press response. To address this hypothesis, a 2 x 2 repeated-measures ANOVA was calculated on mean RTs, with Object Condition (goal/control) and Grip Congruency (congruent/incongruent) as our independent variables. To allow for this factorial analysis, the two items in the control condition were dummy coded as either targets or distractors. For trials in which the prime object was a cigarette, the grape was classified as the 'target', while the apple was the 'distractor'. For trials in which the prime object was a pencil, the apple was classified as the 'target', while the grape was the 'distractor'. The analysis revealed no effect of Object Condition; *F* (1, 28) = 0.14, *p* = .709, hp^2 = .005, or Grip Congruency; *F* (1, 28) = 0.13, *p* = .719, hp^2 = .005, and no interaction *F* (1, 28) = 0.75, *p* = .786, hp^2 = .005, (see Figure 5.4).

The same analyses conducted using the full data set yielded similar results; no main effect of Condition, *F* (1, 54) = 0.80, *p* = .375, hp^2 = .015; no main effect of

Grip Congruency, *F* (1, 54) = 0.64, *p* = .426, hp^2 = .012; and no interaction, *F* (1, 54) = 1.74, *p* = .193, hp^2 = .031.

Figure 5.4.

Mean reaction times in Goal and Control conditions (Experiment 4b). Error bars represent standard errors.



Hand Congruency x Grip Congruency

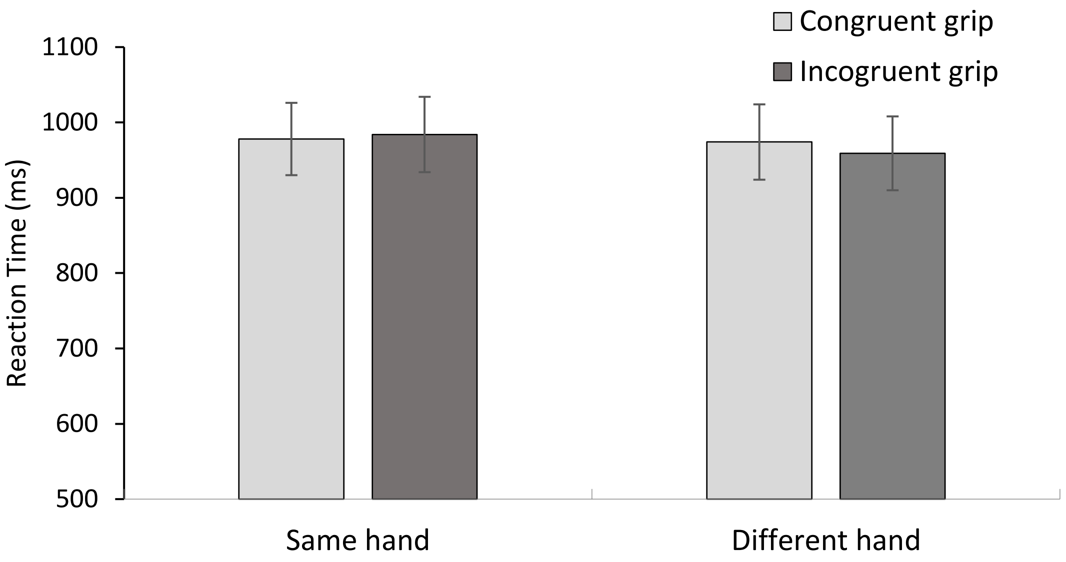
A 2 x 2 repeated-measures ANOVA was calculated to explore the interaction between Grip Congruency (congruent/incongruent) and Hand Congruency (same hand/different hand). An effect of Grip Congruency in the same-hand condition, but not the different-hand condition, would suggest that motor facilitation is driven by an imitative "mirroring" mechanism. However, the analysis revealed no effect of Grip Congruency, *F* (1, 28) = 0.023, *p* = .881, hp^2 = .001, or Hand Congruency, *F* (1, 28)

= 0.563, *p* = .459, hp^2 = .020, and no interaction, *F* (1, 28) = 0.020, *p* = .890, hp^2 = .001 (See Figure 5.5).

Similar results were obtained using the full data set: No main effect of Hand Congruency, *F* (1, 54) = 0.001, *p* = .981, hp^2 = .000; no main effect of Grip Congruency, *F* (1, 54) = 0.126, *p* = .724, hp^2 = .002; and no interaction, *F* (1, 54) = 0.027, *p* = .871, hp^2 = .000.

Figure 5.5.

Mean reaction times in Same Hand and Different Hand conditions (Experiment 4b). Error bars represent standard errors.



Spatial Compatibility x Grip Congruency

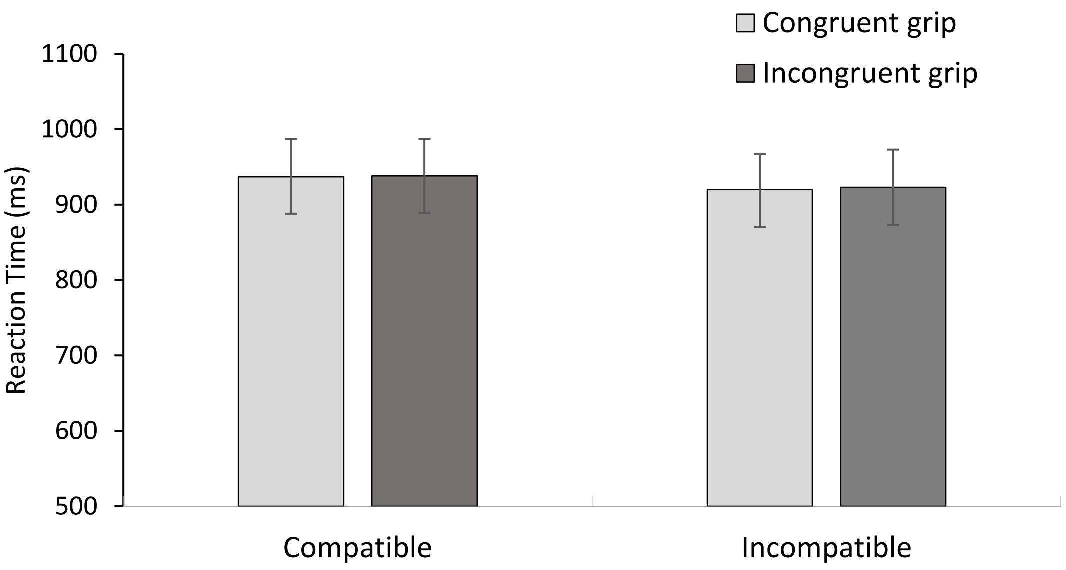
Finally, we conducted a 2 x 2 repeated-measures ANOVA to test whether a Grip Congruency effect might reflect a simple spatial compatibility effect whereby

RTs are faster when the executed hand and target location are compatible (e.g., right-hand, right stimulus) compared to when they are incompatibly (e.g., right-hand, left stimulus; 'The Simon Effect'; Simon, 1969). However, the analysis revealed no effect of Grip congruency, *F* (1, 28) = 1.32, *p* = .260, hp^2 = .045, or Spatial Compatibility, *F* (1, 28) = 0.02, *p* = .889, hp^2 = .001, and no interaction, *F* (1, 28) = 0.01, *p* = .920, hp^2 = .000, (see Figure 5.6).

Similar results were found using the complete data set; no main effect of Spatial Compatibility, *F* (1, 54) = 0.063, *p* = .803, hp^2 = .001; no main effect of Grip Congruency, *F* (1, 54) = 0.64, *p* = .429, hp^2 = .012; and no interaction, *F* (1, 54) = 0.01, *p* = .919, hp^2 = .000.

Figure 5.6.

Mean reaction times in Spatially Compatible and Spatially Incompatible hand-target conditions (Experiment 4b). Error bars represent standard errors.



Additional analyses

In both practice catch trials, the target object was always presented on the same side of the table as the hand the actor was using to reach with (see Figure 5.7 for example). It could therefore be argued that following the arm's initial trajectory, as they reach towards the centre of the table, they are in fact reaching toward the *distractor* object. It is therefore possible that participants assumed the actor was reaching for the object placed on the opposite side of their body.

Analysis of the practice block revealed that 18 participants responded correctly on both practice catch trials. Of these participants, 11 went on to score greater than 70% accuracy on catch trials, suggesting that some participants may have been basing their answers on the action, rather than the objects. To further investigate, a paired-samples *t*-test showed that participants had significantly higher accuracy on catch trials where the target object was located on the opposite side of the table to the hand the actor was reaching out with, t (54) = -13.43, $p < .001$ (see Figure 5.8).

We also investigated whether this bias influenced reaction times by recoding the target object based on the trajectory of the actor's initial reaching movement. For example, in Figure 5.7 the target object would be the notepad, rather than the lighter as in the original analysis. If participants' attention is focussed on the action rather than the objects, we would expect RTs to be faster when the grip recruited by the new target object is congruent with the correct key-press response. For example, if the arm's initial movement was towards the notepad, participants should be faster to respond to the low tone, rather than the high tone. However, a paired-samples *t*-test found no difference in reaction times when the new target object was positioned on the opposite side of the table, *t* (28) = -0.30, *p* = .766 (see Figure 5.9). A similar result was obtained using the complete data set, *t* (55) = 0.545, *p* = .588.

Figure 5.7.

Example catch trial stimuli (Experiment 4b). White cross indicates centre-point of screen.

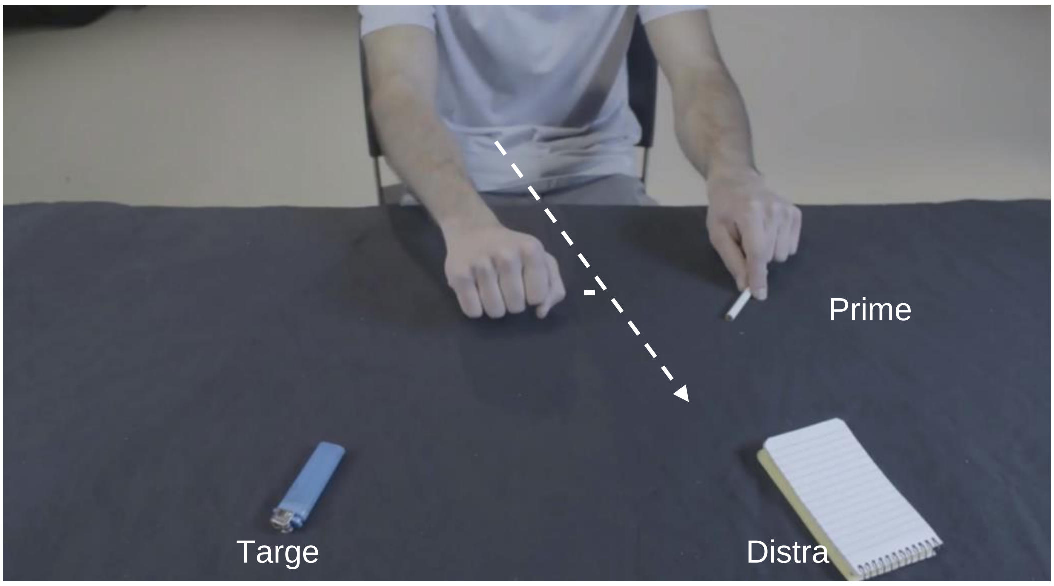


Figure 5.8.

Mean accuracy (%) for Congruent and Incongruent Hand-Target trials (practice block; Experiment 4b). Error bars represent standard errors.

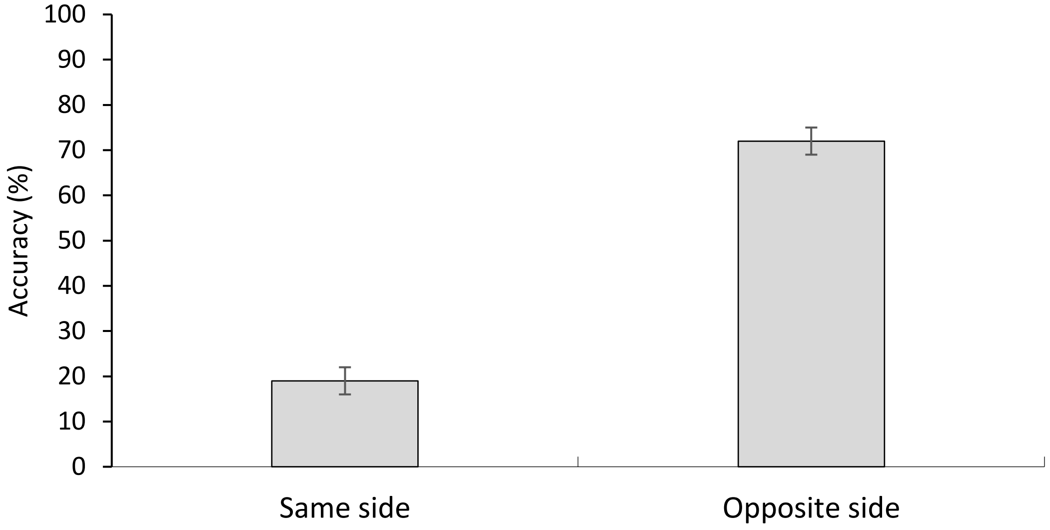
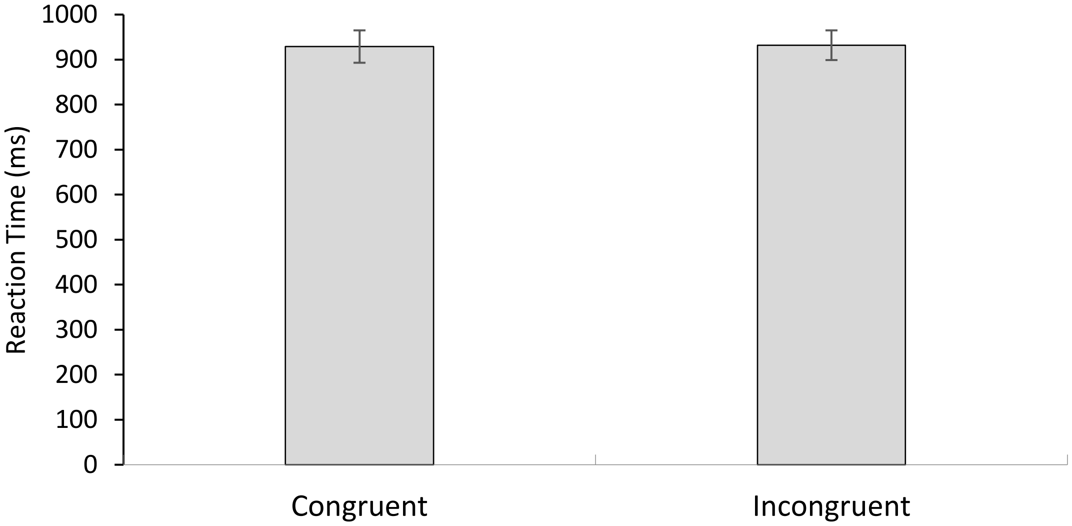


Figure 5.9.

Mean reaction times for Congruent and Incongruent grip responses (additional analyses; Experiment 4b). Error bars represent standard errors.



5.5.1 Discussion

The present study tested for the anticipatory activation of the motor system when viewing object-directed actions. In this online study, participants watched a series of short action sequences of an actor reaching to grasp an object with one hand before reaching towards two other objects with their other hand; one of which was functionally related and another which was functionally unrelated. One object required a PG and the other a WHG. At the end of each action sequence, participants heard a high or low tone. Using the keyboard, their task was to respond with the index finger of one hand when hearing a high tone, or *both* the index and little finger of the other hand when hearing a low tone.

We predicted that participants would be faster to respond when the functionally related object required a grip that matched the required response (e.g., if the functionally related object was a pencil sharpener, reaction times were expected to be faster when participants heard a high tone, compared to a low tone). We also tested whether motor facilitation was driven by an imitative "mirroring" mechanism, whereby participants would be faster to respond when the observed hand was congruent with the response hand, or a spatial compatibility effect, whereby reaction times would be faster when the executed hand and target location were congruent (e.g., right-hand, right stimulus).

We found no evidence for a grasping affordance effect whereby the presentation of a functionally relevant object activates a tendency to grasp it with the corresponding hand. This was the case both when the observed and executed actions were performed using the same hand (i.e. mirroring the observed action) and when the observed and executed actions were performed using different hands. Neither did we find evidence for a spatial compatibility effect, observing no difference in reaction times when the response hand and target location were congruent or incongruent.

Stimulus-response compatibility paradigms have been widely used in the study of object affordances (e.g., Kourtis & Vingerhoets, 2015, Grezes et al., 2003). In particular, the 'Simon effect' (Simon, 1969) demonstrates a correspondence between the (irrelevant) spatial location of the stimulus and the (relevant) location of the response. For example, responses compatible with the handle orientation of a target object are executed faster and more reliably than incompatible responses(Craighero et al., 1998; Phillips & Ward, 2002; Tucker & Ellis, 1998). While some have attributed these findings to a grasping-affordance affect whereby the "intrinsic properties of the object" automatically activate a tendency to grasp it with the corresponding hand (Grezes et al., 2003; Tucker & Ellis, 1998), others suggest that they are due to a simple spatial correspondence. For example, Phillips and Ward (2002) found that the left or right handle orientation of a saucepan influenced participants to give faster responses when the response hand and the handle orientation were matched. However, whether the handle faced toward or away from the participants had no influence on the compatibility effect, as would be expected if compatibility effects were due to action potentiation of the most afforded hand. Furthermore, when participants performed the same task with their hands crossed such that the left key was pressed using the right hand and the right key with the left hand, the spatial compatibility effect was even larger. Compatibility effects have also been obtained with left and right pedal responses using the feet (Phillips & Ward, 2002; Symes et al., 2005). These findings are therefore more consistent with a spatial compatibility account than true affordance effects. We accounted for the occurrence of spatial compatibility effects by testing for faster RTs when the response hand and target location matched (e.g., right-hand, right stimulus). However, we were unable to detect even a simple Simon effect within the current paradigm.

One possibility is that the actor's 'ambiguous' reaching action towards the target and distractor objects was not actually perceived as ambiguous. The initial trajectory of the reaching action was towards the centre of the table. It could therefore be assumed that if the action sequence were to continue, the actor would reach across the table towards the object positioned contralateral to the actor's reaching arm. Additionally, the objects were presented at a left or rightward orientation suitable for the actor to comfortably grasp if reaching from this angle (see Figure 5.7). Consequently, the position of the objects may have provided a cue for the required action judgement. However, further analyses were unable to demonstrate that participants were biased towards either the target or distractor object based on their orientation. Although we had intended to replicate our previously planned experiment (see Experiment 4a) as closely as possible our design was perhaps too complex. The decision to have participants respond using both hands was to ensure that their fingers remained in the correct place throughout the task and that they did not respond using any other fingers (e.g., their middle finger rather than their index finger). However, we failed to counterbalance left and right hand responses – the index finger response was always with the left hand, and the dual index-little finger response with the right hand. This is problematic when one considers that participants were right-hand dominant, which could lead to interactions between the responding hand and the hand being observed.

Furthermore, the 'neutral' grip of the actor's reaching hand might have been perceived as the beginning stages of a WHG, whereas the prime object was always a small PG object. Thus, both grip representations could have been activated on every trial.

Overall, it is probable that responses were influenced by a combination of grip compatibility, spatial compatibility, and action trajectory. However, the presence of so many contributing factors, as outlined above, in addition to the attentional problems associated with poor catch trial performance, likely weakened the power of our analyses to detect an effect within the current paradigm. Experiment 4c aimed to demonstrate a simple grasping affordance effect by removing all contextual and spatial cues.

5.6 Experiment 4c: Simple Affordance Task

The previous study was unable to demonstrate the activation of grasping affordances when viewing object-directed actions. This was likely due to the interaction of several attention-directing cues; grip compatibility, spatial compatibility and action kinematics, in addition to other methodological problems as outlined in the previous discussion. Here, we simplified the previous task in order to establish whether a grasping affordance effect can be demonstrated by visually presenting a series of man-made objects, removing all contextual and spatial cues.

The affordance effect demonstrated by Tucker and Ellis (1998, 2001) and others (e.g., Phillips & Ward, 2002) may occur as a consequence of the action-relevant properties of objects, which automatically generate motor codes based on the actions most likely associated with them. However, it has also been suggested that the presence of a handle makes the object asymmetrical, thus inducing an attentional bias toward the most salient part of the object, and facilitating congruent hand responses (Anderson et al., 2002). Here, we attempted to control for this issue by including objects without any attention-directing cues, such as handles.

Here, in a simplified version of our previous experiment, we presented participants with a static image of a small PG object (e.g. a key) or a large WHG object (e.g. a cup). As before, they heard a high or low tone and were asked to respond using either their index finger or their index and little finger together, respectively. The aim of this experiment was not to directly test the hypothesis that object knowledge primes motor representations, as in our previous experiment. Rather, aimed to test whether it is possible to develop a RT-based measure of action affordances, that would subsequently be used to test this hypothesis in the absence of TMS.

We predicted that participants would be faster to respond when the object required a grip that was congruent with the correct key-press response (e.g., cup – low tone). If viewing an object activates a tendency to grasp it, as the grasping affordance theory suggests, seeing an object should activate the muscles recruited when using that object, thus facilitating a congruent response. For example, seeing a larger object, like a cup, should activate the FDI (index finger) and ADM (little finger) muscles, as these are the muscles recruited for a WHG.

5.6.1 Method

Participants

Participants were 83 right-handed volunteers (43 male, 40 female; mean age: 37.02 years, range: 22 to 69 years), recruited online via Amazon Mechanical Turk. Participants were given £2 for taking part.

Design

The experiment used a repeated-measures design with the independent variable *Grip Congruency;* the grip recruited when using the observed object, and the response required by the tone were either congruent or incongruent. Our dependent variable was the mean reaction time (RT), in milliseconds, for each condition.

Stimuli and procedure

Experimental stimuli consisted of 6 images of manmade 'active' objects, presented at ~ 300 x 300 pixels in the centre of the screen against a black background. Half of the target objects were ‘small’ objects requiring a PG (key, pen, cigarette) and half were ‘large’ objects requiring a WHG (bottle, cup, tennis ball).

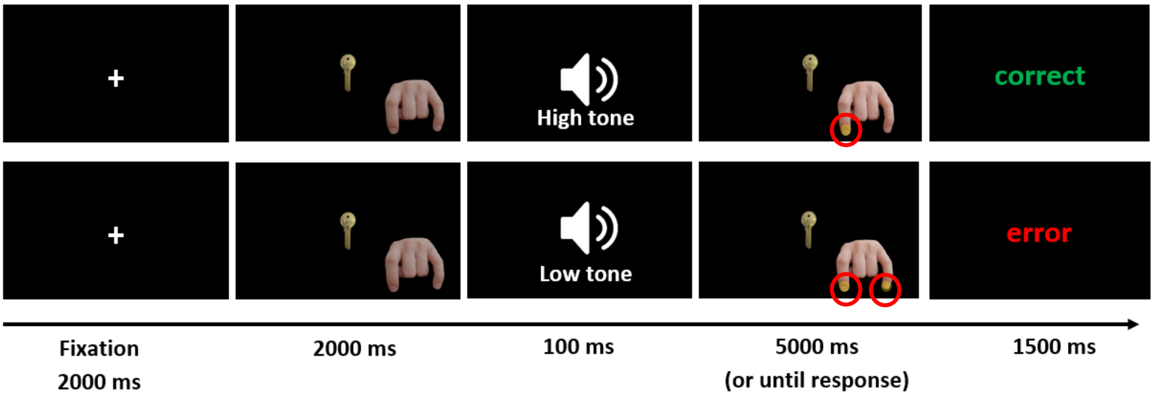
Before beginning the experimental task, participants were shown the objects in a sequence, for a duration of 5 seconds each, and told to imagine using this object. For example, if presented with a picture of a cup, they should imagine picking it up and drinking from it.

Each trial (see Figure 5.10) began with a fixation cross in the centre of the screen for 2000 ms, followed by the object stimulus. After 2000 ms, a high (1000 Hz) or low (500 Hz) tone played for 200 ms. Using the keyboard, participants pressed the ‘G’ key with their right index finger when hearing the high tone, or the ‘G’ and ‘L’ keys together with their index and little finger when hearing the low tone. Participants responded using only their right hand to reduce the effect of handedness bias. Immediate feedback was given by presenting either ‘correct’ or ‘error’ in green or red text, respectively, in the centre of the screen for 1500 ms. The object remained on the screen until a response was given, or until 5000 ms had elapsed. A mirror image of a hand was displayed in the lower-right side of the screen for the duration of each trial and highlighted the correct finger response after each tone.

Participants completed two blocks of 48 trials. Each object was presented 6 times. In half of the trials participants heard a high tone and in half they heard a low tone. The experiment also contained 24 catch trials. Catch trials were identical to experimental trials, except that after responding to the tone, participants were presented with a multiple-choice question: *"Which object did you just see?"* Participants were given a choice for four possible answers and asked to respond using the A, S, D or F keys with their left hand. The question remained on the screen until a response was given, or 20 seconds had elapsed. Trials were randomly presented in each block. Participants completed 12 practice trials to familiarise them with the experimental procedure before beginning the main experiment.

Figure 5.10.

Trial sequence (Experiment 4c).



Panel A depicts a trial in which the grip recruited by the object (PG) and the response required by the tone (index finger only) are congruent, Panel B depicts a trial in which the grip recruited by the object (PG) and the response required by the tone (index and little finger) are incongruent.

5.6.2 Results

Participants were excluded from the analyses if their accuracy on catch trials was less than 70%, suggesting they were not fully processing the prime or target objects. Four participants did not provide complete data sets. This left us with a sample size of 52.

Incorrect responses were removed from the analyses, as were responses faster than 200 ms or slower than two standard deviations from each participant's overall mean. Mean RTs, in milliseconds, were calculated from the remaining data for each participant, in each condition. For low-tone trials, on which the correct response required two simultaneous key-presses, an average RT of the two keys was calculated.

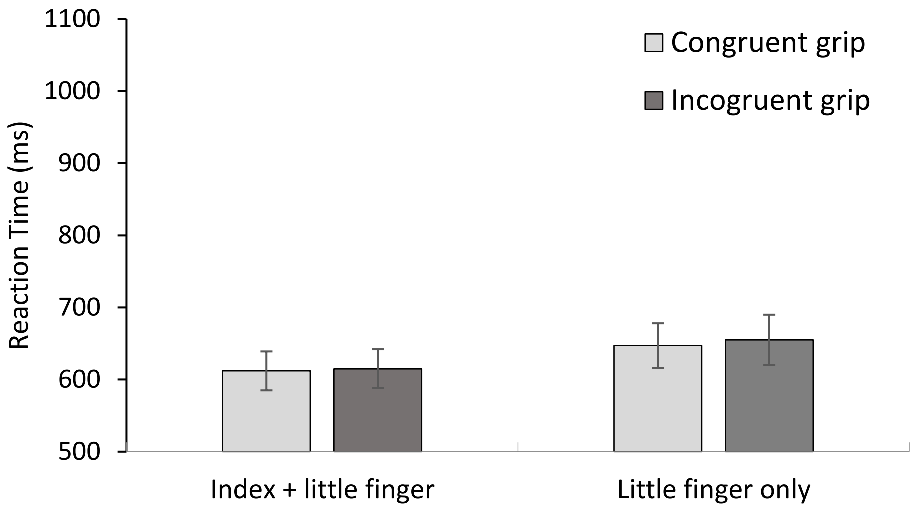
A paired-samples *t*-test found no significant difference in RTs between Congruent (*M* = 612, *SD* = 195) and Incongruent (*M* = 615, *SD* = 193) trials, *t*(51) = -0.78, *p* = .440, *d* = 0.11). Further analyses revealed that participants were faster to respond to High Tones (*M* = 578, *SD* = 164) than to Low Tones (*M* = 651, *SD* = 229, *t* (51) = -6.24, *p* < .001, *d* = 1.03.

Little finger (ADM) only:

We repeated the analysis using only RTs from the 'L' key, which participants pressed using their little finger (ADM). As it is only WHG responses that recruit the ADM muscle, whereas *both* WHG and PG responses recruit the FDI muscle, it is possible that FDI RTs weaken the overall effect. However, a paired-samples *t*-test using only little-finger responses found no significant difference in RTs between Congruent (*M* = 647, *SD* = 220) and Incongruent (*M* = 665, *SD* = 253) trials, *t*(51) = -0.91, *p* = .368, *d* = 0.27 (see Figure 5.11).

Figure 5.11.

Mean reaction times for Dual and Single key-press responses (Experiment 4c). Error bars represent the standard errors.



5.6.3 Discussion

The present study tested whether visually presented man-made objects activate grasping affordances, as measured using choice reaction times, when no other contextual or spatial cues are present. Participants were presented with a static image a small PG object (e.g., key) or a larger WHG object (e.g., bottle), which they were instructed to imagine using. Their task was to respond to a simultaneously presented audio tone, using their index finger when hearing a high tone, or their index and little finger *together* when hearing a low tone.

Overall, participants tended to be faster responding to high tones compared to low tones. This makes sense given that the high tone only required a single key press, whereas the low tone required two simultaneous key presses. There is also evidence that participants tend to find it easiest to pair a high tone with a precision grip and a low tone with a power grip, with responses being executed fastest for precison grips (Ellis & Tucker, 2000). The basis of this 'mapping effect' may be the association of small objects with high tones and large objects with low tones (e.g., mice and elephants; see Ellis & Tucker, 2000). However, it is the relationship between the visual objects and the responses that is of primary interest in this experiement.

We predicted that participants would be faster to respond when the visual object required a grip that was congruent with the correct grip response (e.g., cup – low tone - WHG). However, we were unable to find any evidence for grip activation based on the affordances of the visual objects. The compatibility effects demonstrated in previous work, such as Tucker and Ellis' (1998) well-known study, have been taken as evidence that observing objects activates affordances, and that affordances are activated *automatically*. However, more recent studies have shown that their activation is modulated by the task and context (Girardi et al., 2010; see Borghi & Riggio, 2015, for a review). Yu and colleagues (2014), for example, failed to replicate compatibility effects when participants were not explicitly instructed to imagine *picking up* the pictured objects.

Each of the experiments so far in this chapter have been based on studies which measured MEPs in the FDI and ADM muscles (index finger and little finger, respectively) as a test of PG versus WHG activation (Makris et al., 2011; Sartori et al., 2012; Sartori, Bucchioni, et al., 2013). In particular, we have focussed on differences in ADM activation, as this finger is activated for a WHG but not a PG. We also included what many would consider to be typical 'precision grip' objects such as a pen or key. However, there are several different variations of a PG, each of which activates different hand muscles. For example, the tip-to-tip connection of the fingers in a 'pinch grip' can be used for the manipulation of small objects such as a sewing needle, whereas holding a pen utilises a 'tripod grip', which requires flexing *both* the index and little finger (Cavina-Pratesi et al., 2018; Gentilucci et al., 2003) It could therefore be suggested that while our previous studies used 'small' and 'large' objects which are typically categorised as PG and WHG objects respectively, the ADM could have been primed in both conditions, thus weakening the effect. Furthermore, half of the trials required an index finger only response, and half required *both* an index and little finger response. Therefore, the FDI was also being activated on every trial.

In Experiment 4d, we refined the previous experiment by replacing the previous 'precision grip' objects with specific 'pinch grip' objects and instructed participants to imagine *picking up* that object, rather than using it.

5.7 Simple Affordance Task 2: Experiment 4d

In this experiment, we replaced the previous 'precision grip' objects (key, pen, cigarette) with specific 'pinch grip' objects (matchstick, chess piece, penny) and instructed participants to imagine *picking up* that object, rather than using it. For example, if presented with a small object like a penny, they should imagine grasping it between their thumb and index finger.

The previous study required participants to respond using only their index finger in the PG condition and *both* their index and little finger simultaneously in the

WHG condition, thus activating the index finger in both conditions and potentially weakening the affordance effect. Therefore, in the current study, participants responded using only their little finger when they heard a tone, and occasionally with their index finger when they heard white noise (white noise trials were to ensure participants had their hands positioned with their little finger on the 'L' key, rather than simply responding to the tone with another finger). We predicted that participants would be faster to respond to the tone when they saw a WHG object than a PG object. This is because observing an object activates the muscles recruited when using that object. In other words, viewing an object that requires a WHG (e.g. a mug) activates the muscles recruited when using that object (i.e. the little finger and index finger to grip the mug). This activation then facilitates the little-finger response to the tone. In contrast, we should see no such facilitation when the object requires a PG, as this grip only requires the index finger.

5.7.1 Method

Participants

Participants were 123 right-handed volunteers (66 male, 57 female, XX; mean age: 40.38 years, range: 22 to 72 years), recruited online via Amazon Mechanical Turk. Participants were given £2 for taking part.

Design

The experiment used a repeated-measures design with the independent variable *Grip Type;* observed objects required either a PG or a WHG. Our dependent variable was the mean reaction time (RT), in milliseconds, for each condition.

Stimuli and procedure

Experimental stimuli consisted of 6 images of manmade objects, presented at ~ 300 x 300 pixels in the centre of the screen against a black background. Half of the target objects were ‘small’ objects requiring a PG (penny, matchstick, chess piece) and half were ‘large’ objects requiring a WHG (bottle, cup, tennis ball).

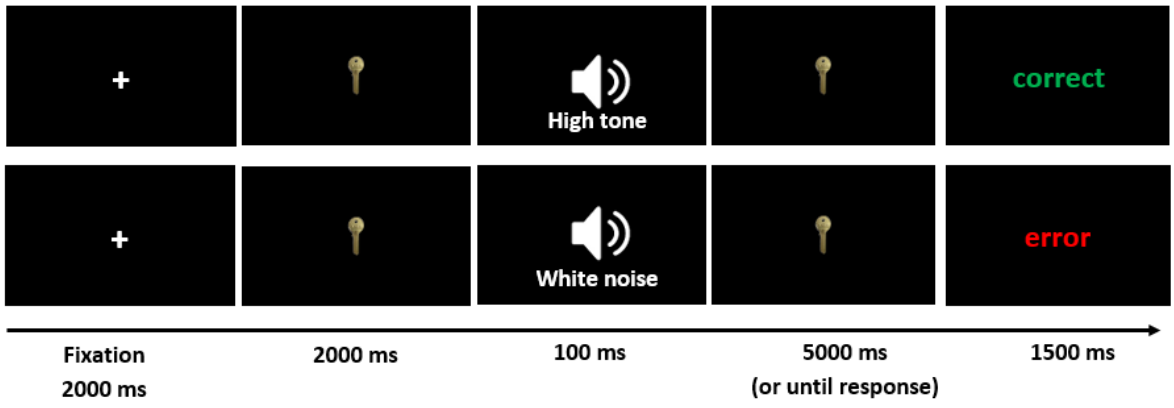
Before beginning the experimental task, participants were shown the objects in a sequence, for a duration of 5 seconds each, and told to imagine picking up that object. For example, if presented with a small object like a penny, they should imagine grasping it with their thumb and index finger.

Each trial (see Figure 5.12) began with a fixation cross in the centre of the screen for 2000 ms, followed by the object stimulus. After 2000 ms, participants heard either a low (500 Hz) tone or a white noise stimulus (0 dBFS, 44.1 kHz) for 200 ms. Using the keyboard, participants pressed the ‘L’ key with their right little finger when hearing the low tone, or the ‘G’ key with their index finger when hearing the white noise. Immediate feedback was given by presenting either ‘correct’ or ‘error’ in green or red text, respectively, in the centre of the screen for 1500 ms. The object remained on the screen until a response was given, or until 5000 ms had elapsed.

Participants completed three blocks of 30 trials. Each object was presented 12 times. White noise trials accounted for 20% of experimental trials. The experiment also contained 18 catch trials. Catch trials were identical to experimental trials, except that after responding to the tone, participants were presented with a multiple-choice question: *"Which object did you just see?"* Participants were given a choice for four possible answers and asked to respond using the A, S, D or F keys with their left hand. The question remained on the screen until a response was given, or 20 seconds had elapsed. Trials were randomly presented in each block. Participants completed 10 practice trials to familiarise them with the experimental procedure before beginning the main experiment.

Figure 5.12.

Trial sequence (Experiment 4d).



Panel A depicts a trial in which participants heard a high tone and were required to respond using their little finger (this would be a 'correct' response). Panel B depicts a trial in which participants heard white noise and were required to respond using their index finger (any other response resulted in an 'error' message).

5.7.2 Results

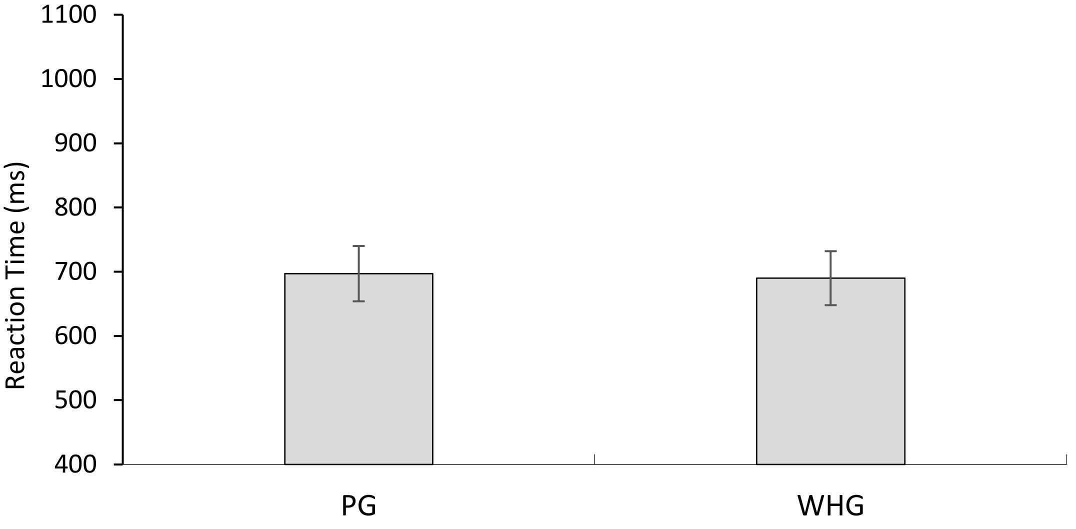
Participants were excluded from the analysis if their accuracy in the catch trials was less than 70% suggesting they were not fully processing the prime or target objects. Twelve participants did not provide complete data sets. This left us with a sample size of 99.

Incorrect responses were removed from the analysis, as were responses faster than 200 ms or slower than two standard deviations from each participant's overall mean. Mean RTs, in milliseconds, were calculated from the remaining data for each participant, in each condition.

A paired-samples *t*-test was conducted to test whether participants were faster to respond on the low tone trials when the object stimulus required a WHG compared to a PG. We found no significant difference in RTs between PG (*M* = 669 ms, *SD* = 427 ms) and WHG trials (*M* = 690 ms, *SD* = 414 ms), *t*(98) = 1.04, *p* = .302, *d* = 0.12 (see Figure 5.13).

Figure 5.13.

Mean reaction times for PG and WHG objects (Experiment 4d). Error bars represent standard errors.



5.7.3 Discussion

Experiment 4d sought to establish whether a grasping affordance effect could be demonstrated when visually presenting a series of man-made objects, removing all contextual and spatial cues. We refined the method used in Experiment 4c by replacing the previous PG stimuli with more specific 'pinch grip' objects and instructed participants to imagine *picking up* that object, rather than using it. Our earlier experiments required participants to respond using only their index finger in the PG condition and *both* their index and little finger simultaneously in the WHG condition, thus activating the index finger in both conditions and potentially weakening the affordance effect. We therefore only measured and analysed responses from the little finger in response to the audio tone.

We predicted that participants would be faster to respond to the audio tone when the object presented required a WHG rather than a PG, because only a WHG recruits the muscles in the little finger. Thus, the activation of this finger should facilitate the response to the audio tone. However, in contrast to these predictions, we observed no difference between PG and WHG trials.

5.7.4 General Discussion

Our earlier experiments tested for a behavioural bias in participant's perception of actions towards expected goals. The aim of the current chapter was to further test whether these biases occur not only in overt behaviour, but also if they are unconscious and automatic.

Initially designed as a TMS study (Experiment 4a), we sought to investigate the possible involvement of the motor system in the anticipatory representation of action goals associated with common everyday objects by measuring anticipatory muscle activation in the hands of participants when viewing objects. As we were unable to implement this lab-based experiment due to COVID-19 restrictions, we instead adapted our original experiment for use online as a behavioural study, using RTs as an index for motor preparation.

Experiment 4b tested for the anticipatory activation of grasping affordances when viewing object-directed actions. Participants viewed a series of ambiguous object-directed reaching actions while tasked with responding to a high or low audio tone using either their index finger only (to mimic a PG), or their index and little finger simultaneously (to mimic a WHG). We predicted that responses would be faster when the relevant object required a grip that matched the required response (e.g., a pencil sharpener requires a PG, which facilitates the index-finger only response).

Experiment 4c removed the action component of the previous task, as well as all other attention-directing cues (e.g., handles), and tested for the anticipatory activation of grasping affordances when viewing each object in isolation. We predicted that participants would be faster to respond when the object required a grip that matched the correct response (e.g., a cup requires a WHG, which facilitates the index and little finger response).

Experiment 4d refined the methods and stimuli used in Experiment 4c. We replaced the previous 'precision grip' objects (e.g., key) with more specific 'pinch grip' objects (e.g., penny), to minimise activation of the ADM (little finger) muscle on these trial and measured responses to a single audio tone using the little finger only, to tap into differences in ADM muscle activation.

We were unable to find evidence supporting the anticipatory activation of grasping affordances when viewing object-directed actions (Experiment 4b) or when viewing objects alone, in the absence of any contextual or spatial cues (Experiment 4c and 4d). The absence of a grasping affordance effect in Experiment 4b likely reflects the complexity of the task and the multiple conflicting action representations that may have been active in such a complex task. For example, the "neutral" grip of the actor's reaching hand could easily have been perceived as the early formation of a WHG, while the actor's other hand was always seen grasping a small PG object, thus activating both types of grip on every trial. The trajectory of the reaching action toward the centre of the table might also have implied that the actor was intending to reach for the object placed contralateral to the reaching arm. To further complicate matters, we asked participants to respond using both hands. The advantage of this was that it ensured participants' fingers remained on the correct keys throughout the task and that they could not respond using any other fingers (e.g., their middle finger rather than their index finger). The disadvantage of this approach was that it could have created interactions between the response hand and the hand being observed.

With these considerations in mind, Experiments 4c and 4d adopted a simplified paradigm to assess whether grasping affordances could be measured using reaction times. Once this has been established, we could then use reaction times test our original hypothesis that object knowledge primes motor representations. However, neither of these experiments were able to detect grip activation based on the affordances of the visual objects.

Our results are contrary to those of previous studies which have demonstrated grasping affordance effects when viewing real three-dimensional objects (Ellis & Tucker, 2000), two-dimensional coloured pictures of objects (Buccino et al., 2009a; Makris et al., 2011) and two-dimensional black and white drawings of objects (Phillips & Ward, 2002). The disparity in results could be explained by methodological differences. As in Ellis & Tucker's (2000) study, participants in our experiments were asked to respond to an orthogonal stimulus (an auditory tone) while viewing real objects that afforded either a PG or WHG. In Ellis and Tucker's experiment, as well as similar studies (e.g., Makris et al., 2011), responses were made using an apparatus which mimicked a PG or WHG. Due to the constraints of online-testing, we were unable to measure grip responses in this manner. Instead, we asked participants to respond using either their index finger only, or their index and little finger simultaneously. We reasoned that because these were the hand muscles recruited by a PG and WHG, respectively, the activation of action representations elicited by the objects should result in the facilitation of a subsequent motor response. In Experiments 4b and 4c, this approach proved problematic in that the typical use of some of the broadly defined 'precision grip' objects used in these tasks arguably recruits *both* the ADM (little finger) and FDI (index finger) muscles. For example, when holding a pen as if to write, both the index and little finger are flexed. Thus, the ADM muscle may have been activated on these trials. We attempted to partially rectify this problem by using more specific PG objects in Experiment 4d. However, the use of a more appropriate grip response device in a controlled lab setting would undoubtedly yield more accurate results.

Another possible explanation for this discrepancy in results is the longer stimulus-onset asynchrony (SOA) of 2000 ms between the onset of the object images and the presentation of the audio stimulus in our experiments. In a similar experiment, Makris, Hadar and Yarrow (2011) varied the onset time between the prime and target stimulus and found that affordance effects were present for a maximum period of ~600 ms post-stimulus and were largely absent after that. Ellis and Tucker (2000) also reported RT effects with a SOA of 700 ms. Thus, with a SOA of 2000 ms, any response activation in our experiments could have dissipated before the presentation of the tone.

In contrast, other studies have observed an *increase* in affordance effects with increasing SOAs, although the maximum SOA in these studies was still only 1200 ms; considerably shorter than our SOA of 2000 ms (Phillips & Ward, 2002; Vingerhoets et al., 2009). From an evolutionary perspective, the immediate activation and fast decay of prepared actions may provide an adaptive advantage. On the one hand, survival in a hostile environment may require immediate action by rapidly activating movements which are already planned. On the other hand, the build-up and maintenance of motor plans which might never be initiated would be both effortful and metabolically costly (Makris et al., 2011). Our decision to use a longer SOA was initially because we were interested in how knowledge of one object influenced motor responses to a second object. This required sufficient time for participants to process both objects, embedded in the context of an action. We therefore maintained the same SOA in our simplified paradigm, with the intention of moving back to this kind of design once we had established affordances could be reliably measured using reaction times. However, in Experiment 4d, participants were explicitly asked to imagine themselves using each object. One might therefore assume that this would extend the activation of this action representation. Future studies should use a range of SOAs to further understand the time course of response activation and dissipation.

The majority of early affordance studies have used stimulus-response compatibility (SRC) paradigms (e.g., Craighero et al., 1998; Phillips & Ward, 2002; Tucker & Ellis, 1998). For example, Tucker and Ellis (1988) found a compatibility effect between the location of the handle of the object (left/right) and that of the key press (left/right). These results suggest that handles evoke affordances, even if the task does not require to pay attention to them. This has been taken as evidence that observing objects activates affordances, and that affordances are activated *automatically*. However, the issue of automaticity has been widely debated and cannot be concluded without ruling out other explanations such as the spatial location of the object's salient features and social context (for a review, see Borghi &

Riggio, 2015). We were unable to find evidence in support of the automatic activation of object affordances, within the current paradigm. Surprisingly, we were also unable to find spatial compatibility effects (see Experiment 4b).

Tipper, Paul and Hayes (2006) found that the action state of objects influenced the activation of affordances. Active objects, with which current action is implied, produce larger affordance effects than passive objects, with which no action is implied. Similarly, Yoon and colleagues found that affordance effects were stronger when an agent was shown holding the objects, and disappeared altogether when the objects were not shown from the first-person perspective. This suggests that observers are sensitive to whether objects are positioned correctly for their own actions. The objects in Experiment 4b were positioned from the opposite viewpoint of the participant and therefore not positioned for a right-hand action. This could explain why we did not find a grasping affordance effect. It is possible that a further study in which the objects are depicted from an egocentric viewpoint would give rise to affordance effects.

Across 4 experiments, we aimed to assess the role of the motor system in representing the observed movements with respect to potential object affordances. In Experiment 4a, we aimed to assess this using TMS but were unable to run this study due to COVID-19 restrictions. Experiment 4b used a similar design, to see if this question could be addressed using reaction time measures. The lack of any clear differences in this study may be largely attributable to the complex design, and the multiple action representations portrayed in the videos. Following on from this, Experiments 4c and 4d used progressively simpler designs, to test the extent to which object affordances can be reliably measured using reaction time. The absence of clear effects in these studies suggests that the RT methods we developed were not sensitive enough to differentiate actions afforded by the target objects.

Chapter 6: General Discussion

6.1 Summary of aims

Understanding the goals and intentions of others is crucial for social interaction. However, this can be challenging when the goals and intentions that drive others' behaviour are not immediately obvious. For this reason, humans are particularly sensitive to cues in the environment that covey the goals of others' and the possible actions they might perform. Cues such as facial expressions (C. Frith, 2009; Johnston et al., 2010), gaze direction (Bayliss et al., 2007; Castiello, 2003; Pierno et al., 2006) and language (Hudson, bach et al., 2018; Hudson, Nicholson, Ellis, et al., 2016; Todd et al., 2011). The affordances of available objects have also been shown to facilitate top-down predictions about others' goals (Ambrosini et al., 2011; Cardellicchio et al., 2013; Schubotz et al., 2014). However, previous studies have tended to focussed on low-level action representations triggered by objects, such as grip size (Buccino et al., 2009a), or have explicitly provided high-level goal information (Hudson, Bach, et al., 2018; Hudson, Nicholson, Ellis, et al., 2016). This present thesis had three aims: (1) to test whether high-level action goals based on object knowledge can bias action perception, (2) to investigate the degree to which this perceptual bias can be influenced by high-level person knowledge, or by expertise in particular objects, (3) to explore the low-level mechanisms underlying the anticipatory representation of action goals associated with objects.

6.2 Overview of findings

Chapter 2 investigated whether objects activate higher-level action goals. Using a modified representational momentum paradigm adapted from Hudson et al. (Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpsons, et al., 2016;Hudson, Bach, et al., 2018), we tested whether the visual presentation of a prime object (e.g., a hammer) would lead to a predictive bias in the perception of a subsequent action towards a functionally related target object (e.g., a nail).

Participants watched videos of a hand reaching for or withdrawing from a single target object (Experiment 1a) or a target object and a distractor object (Experiment 1b). Midway through the action, the hand disappeared and participants were asked to judge the index finger’s final position on the touchscreen. Prior to action onset, we implicitly manipulated the observer’s knowledge of the actor’s intention by presenting a static image of a functionally related or functionally unrelated prime object. As expected, the actor's hand was consistently misperceived as being closer to the target object when it was functionally related to the prime (e.g. hammer/nail) than when it was functionally unrelated (e.g. hammer/cigarette). This bias was present only for reaching actions, but not withdrawing actions (Experiment 1a) and persisted even when the functionally-related target object was simultaneously presented with an unrelated distractor object (Experiment 1b). Importantly, this effect appeared to be specific to intentional actions and disappeared when the actor's hand was replaced by a moving ball following the same path (Experiment 1b). These findings suggest that the perceptual bias toward the target object was driven by knowledge of the actor's intention, derived from the affordances of the available objects.

Chapter 3 sought to determine the degree to which these perceptual biases could also be influenced by person knowledge. Specifically, the gender of a previously presented face stimulus. Using the same representational momentum paradigm as in Experiment 1b, we replaced the prime object with a photo of a male or female face, and replaced the target and distractor objects with object rated as being 'highly masculine' (e.g., a Swiss army knife) or 'highly feminine' (e.g., a lipstick). To draw participants' attention to the gender of the face, participants were occasionally asked to identify whether the face they had seen was male or female (Experiments 2a and b) or explicitly asked which object they thought the actor was reaching for (Experiment 2c). When participants were asked to report the gender of the face they had just seen (Experiment 2a and 2b), no perceptual biases were detected, even when the prime stimuli were replaced with a new set of faces that had been previously rated as either 'highly masculine' or 'highly feminine' (Ma et al., 2015). The absence of any observable effects could be attributed to weak associations in memory between gender and objects, resulting in the limited activation of gender stereotypes, and thus diminishing the magnitude of the priming effect (Dijksterhuis et al., 2000). In Experiment 2c we sought to strengthen this association by directly asking participants which object they thought the actor was reaching for. Interestingly, when asked to make these explicit judgements, participants exhibited a tendency to perceive the actor's hand as disappearing closer to the *distractor* object than to the target object (e.gs., male face – lipstick). In contrast, when making explicit action judgements, participants indicated an expected reach toward the *target* object (e.g., male face – Swiss army knife). Thus, there appeared to be a dissociation between the effects of implicit and explicit information.

Chapter 4 aimed to assess the degree to which perceptual biases could be influenced by an individual's own expertise in a particular object. Adapting the representational momentum paradigm used in chapters 2 and 3, we presented two groups of participants – skilled chess players or novice players – with a series of short videos showing a hand reaching toward a chessboard, which was set up in a checking configuration. As in our previous experiments, the hand disappeared mid-action and participants indicated the index finger’s final position on a touchscreen. We predicted that expertise in chess (i.e., knowledge of the rules of movement associated with individual chess pieces) would generate a top-down predictions regarding which chess piece the actor would reach for. However, contrary to our predictions, chess knowledge did not exert an observable influence on perceptual judgements. This lack of clear differences could be explained by the low overall ability level of participants in the 'skilled' chess players group. Despite our best efforts to recruit highly skilled chess players from national chess clubs, uptake was low, leading us to screen for chess expertise via a custom 10-question pre-test. Although there was a statistically significant difference in pre-test scores between the Skilled and Novice players, it could be argued that the level of chess knowledge found within our 'Chess Players' group was too weak to influence action perception.

The experiments in chapters 2, 3 and 4 tested whether action goal predictions elicit perceptual biases. Chapter 5 tested whether such predictions also elicit *motor* biases, derived from the affordances of available objects. Experiment 4a initially aimed to assess this using TMS, hypothesising anticipatory muscle activation in the hand of the observer for the grip afforded by the expected target object. However, in response to the university's COVID-19 restrictions, we adapted our original study design to see if we could answer the same question behaviourally, using reaction time measures. In Experiment 4b, participants watched a series of short videos showing an actor reaching to grasp an object (the "prime") with one hand before reaching towards two other objects with their other hand; one object that was functionally related to the prime, and another that was functionally unrelated to the prime. One object required a precision grip (PG) and the other a whole-hand grip (WHG). At the end of each video, participants heard a high or low tone. Using the keyboard, their task was to respond with the index finger of one hand when hearing a high tone, or *both* the index and little finger of the other hand when hearing a low tone. We predicted that participants would be faster to respond when the functionally related object required a grip that matched the required response. However, we found no evidence for a grasping affordance effect whereby the presentation of a functionally relevant object facilitated a motor response with the corresponding hand muscles. The lack of clear differences in this study may be largely attributable to the complexity of the task, and the multiple action representations portrayed in the videos. With this in mind, Experiment 4c used a simplified paradigm to first establish whether a grasping affordance effect could be demonstrated using reaction time based measures. Only when this has been established can we use this method to behaviourally test our original hypothesis that object knowledge primes motor representations.

In Experiment 4c, participants were presented with an image of a small PG object (e.g., key) or a larger WHG object (e.g., bottle), which they were instructed to imagine using. The task was to respond to a simultaneously presented audio tone, using their right index finger when hearing a high tone, or their right index and little finger *together* when hearing a low tone. We predicted that participants would be faster to respond when the object required a grip that matched the correct response (e.g., cup - low tone - WHG). However, we were unable to find any evidence of grip activation based on the affordances of the visual objects. In this experiment, we removed any contextual or spatial cues, such as handles, to reduce the complexity of the task and to rule out the possibility of inducing an attentional bias toward the salient features of the objects. We did not, however, fully consider the vast array of human hand actions, which can vary extensively depending on the end goal of the action (e.g., using an object or moving it) or the features of the target object (e.g., size, weight, or orientation; (Cavina-Pratesi et al., 2018). To categorise objects as either "precision grip" or "whole-hand grip" objects might have been far too simplistic given the many subcomponents of human hand actions. For example, when using a key, the key is held between the flexed thumb and middle phalanx of the index finger. In contrast, when using a pen, the pen is held between the thumb and the index and middle finger. These, along with the ring and little finger, are flexed at the metacarpophalangeal joint, slightly flexed at the proximal interphalangeal joint, and extended at the distal interphalangeal joint (Jones & Lederman, 2006).

In experiment 4d, we replaced the previously defined 'precision grip' objects with "pinch grip" objects. That is, objects which would typically be picked up via a tip-to-tip connection between the thumb and index finger (e.g., a penny). Rather than asking participants to imagine *using* each object, we instructed them to imagine *picking up* each object. In the previous studies, participants had responded using their index finger in both PG and WHG conditions, potentially weakening the affordance effect. Here, we asked participants to respond using only their little finger when hearing a single tone. We hoped that these adjustments would strengthen the affordance effect, facilitating reaction times when participants viewed WHG objects, compared to PG objects. Contrary to this prediction, no differences were observed between PG and WHG trials.

6.3 Predictive processing and action perception

Traditional views of social perception describe action understanding as a bottom-up process whereby the motor system automatically activates when observing the actions of others. These low-level motor activations propagate upwards to higher levels in the observer's motor hierarchy, allowing for inferences regarding the likely goal of an action (Gazzola & Keysers, 2009; Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). More recent theories have highlighted the involvement of top-down processing in action perception; we first use non-motoric information, such as context, to predict the possible goal of an action. Our motor system then simulates this action to assess the degree to which the assumed goal matched the behaviour of the individual (Csibra, 2008; Kilner et al., 2007). More specifically, object-based views of action understanding highlight the importance of objects in predicting the potential goal of an upcoming action (Bach et al., 2014). The experiments in this thesis aimed to investigate the way in which object knowledge might contribute to action understanding.

If social perception is predictive and prior knowledge of others' intentions shapes our perception of their actions, when an action is ambiguous, our perception should be biased towards our prediction. This perceptual bias has been demonstrated using a modified representational momentum paradigm (J. Freyd & Finke, 1984; Hubbard, 1993), originally designed to measure top-down predictive influences in non-social motion perception, to test for similar predictive effects in social perception (Hudson, Bach, et al., 2018; Hudson, McDonough et al., 2018;

Hudson, Nicholson, Ellis, et al., 2016; Hudson, Nicholson, Simpson, et al., 2016; McDonough et al., 2019). Hudson, Nicholson, Ellis, and colleagues (2016) showed participants short videos of an actor either reach for or withdraw from an object. Prior to action onset, participants heard the actor state either “I’ll take it” or “I’ll leave it”. The action disappeared midway through its trajectory and participants estimated its disappearance point. They found that participants perceived reaches further towards an object than they really were when they assumed that the actor wanted to pick it up, and further away when they assumed a withdrawal. This suggests that the perception of others’ behaviour is biased by our prior expectations of what they will do. Subsequent experiments built on these initial findings, showing similar effects when participants made predictions on the basis of object type, saying "take it!" or "leave it!" when presented with a 'safe' or 'painful' object prior to action onset (Hudson, Nicholson, Simpson, et al., 2016). We aimed to build on this prior research by testing whether high-level action goals based solely on object affordances can influence action perception in a similar way.

The results from Experiments 1a and 1b provide evidence for predictive processing during social perception. We found a representational momentum effect such that people consistently misperceived the vanishing point of a hand as being closer toward the target object than it really was when it was preceded by a functionally related prime object compared to a functionally unrelated prime. For example when the hand was seen reaching for a cup, the disappearance point was reported as being further along the trajectory if preceded by an image of a bottle, compared to an image of a hammer. This finding is in accordance with predictive processing accounts of social perception, which suggest that inferred goals generate top-down predictions of what will be perceived next, biasing perception towards the expected action (Bach et al., 2014; Friston et al., 2011; Kilner et al., 2007). This finding builds on previous work by Hudson and colleagues, in which actions goals were provided directly to participants prior to action onset, whether by hearing the actor directly state their intention (Hudson, Nicholson, Ellis, et al., 2016), or by marching the hand's grip to an available object (Ambrosini et al., 2011). Here, no explicit information was available. The actor's grip was aways matched to the target object (Experiment 1a) and the distractor object (Experiment 1b). Thus, the goal could only have been inferred from knowledge about the prime object presented at the start of each trial. It suggests that the perceptual bias toward the expected goal was driven by a prediction about the actor's intention, derived from prior knowledge about the objects.

An important finding in Chapter 2 was that perceptual displacements towards expected goals were present when observing reaching actions, but not withdrawing actions (Experiment 1a).These findings are consistent with prior research showing that that action prediction specifically occurs for meaningful actions towards objects (Stapel et al., 2012) but not withdrawals from them (Schenke et al., 2016). They offer further support to recent predictive models of social perception (Bach et al., 2014; Csibra, 2008; Kilner et al., 2007), which suggest that observers are constantly testing their predictions about others' goals and integrating them with incoming sensory information. When the goal and sensory input match, the prediction is processed fluently and perception becomes biased towards the expected movement. By contrast, unexpected actions elicit prediction errors and re-evaluations of prior expectations (Csibra, 2008; Hudson, Nicholson, Ellis, et al., 2016). Thus, a mismatch between the expected intention to grasp the object and the observed withdrawing action would be easily detected, and the representational momentum effect reduced.

Crucially, the RM effect towards the expected target object was specific to intentional reaching actions, but eliminated when the hand was replaced by a non-agentive ball following the same trajectory (Experiment 1b). This suggests that the predictive bias observed in Experiment 1b was not solely driven by the activation of object function knowledge (Collette et al., 2016), but was also driven by cues to intentionality derived from the objects' semantics, and predictions about the goal of the observed action.

6.4 Person knowledge and action goal prediction

Predictive-processing models of social perception assume that action predictions reflect not only overt contextual cues, but also higher-level knowledge about the actors themselves. Our assumptions about others' are integrated with situational factors, such as the objects available to achieve the current goal, and translated into predictions about their forthcoming actions (Bach & Schenke, 2017). In Chapter 3, we tested whether gender information biased action perception towards a 'highly masculine' or 'highly feminine' object. However, Experiment 2a and 2b revealed no observable biases in perceptual judgements toward either the object. One possible explanation for this is that the association between the faces and the objects was not strong enough to induce any detectable biases. It has been widely demonstrated that the activation of stereotypes influences ongoing behaviour (Blair & Banaji, 1996; Dijksterhuis et al., 2000; Dijksterhuis & Van Knippenberg, 1996, 1998; Macrae, Stangor, & Milne, 1994) and biases the interpretation of behaviour (Condry & Ross, 1985). Furthermore, the degree of stereotype activation influences the size of the behavioural effect such that the higher the degree of stereotype activation, the more pronounced the behavioural effects will be (Dijksterhuis & Van

Knippenberg, 1998). If participants had low exposure to gender stereotypes in their everyday lives, their associations between the gender and the objects would be limited. This idea is supported by previous research showing that Individuals who endorse stereotypes have, on average, stronger gender associations. For example, people who score low on sexism find It harder to list female stereotypes than people who score high on sexism (Dijksterhuis et al., 1999).

In addition to making perceptual judgements about the hand's disappearance point, in Experiment 2c, participants were explicitly asked to judge which object they thought the actor was reaching for. The purpose of this question was to strengthen the association between the faces and the objects. Unexpectedly, we found that while participants reported an expected reach toward the target object, their perceptual judgements shifted toward the distractor object. In other words, when primed with a male face, participants explicitly reported an expected reach toward the knife, but misperceived the hand as being closer toward the lipstick. It is difficult to explain these results within the context of predictive processing models of social perception, which assume that top-down information about other people (e.g., their goals or beliefs) is used to generate predicts about their most likely actions, which in turn biases the identification of actions towards these predictions (Schenke et al., 2016). Based on these assumptions, one would assume that predictions derived from deeply engrained social stereotypes would bias perception in the direction of the stereotype. One explanation for these contrast effects is that they are the result of reactance, particularly if participants did score low on sexism. By explicitly priming a link between the gender and the objects we might have made them more aware of *not* portraying sexist behaviours. In other words, while participants were overtly trying to behave in the way we expected, their behaviour actually reflected a reaction *against* our portrayal of sexism. Put differently, one could argue that there is an underlying implicit bias, but once it becomes explicit, it triggers a reactance to avoid behaving in a way that is consistent with the stereotype.

6.5 Expertise and action goal prediction

Predictive processing models suggest that social perception is informed not only by bottom-up signals from the environment, but also by top-down predictions based on our expectations about what those incoming signals will be (Bach & Schenke. 2017). Our first two experiments (Experiments 1a and b) demonstrated that such predictions can be derived from prior knowledge about objects. In Experiment 3 (Chapter 4), we further hypothesised that expertise in a particular object would generate top-down predictions about the most likely forthcoming actions, biasing perception towards this prediction. To test the effect of expertise on action prediction, we compared chess experts, who possess extensive experience with and knowledge about chess pieces and their relations, to novice players. However, we were unable to find evidence for a predictive bias using our modified representational momentum paradigm.

The objects used in Experiments 1a and 1b were common household objects which all participants could be considered "experts" in (e.g., most people use a cup of some sort on a daily basis). By comparison, the players in Experiment 3 could be described as having had relatively limited exposure to the objects, particularly if most of their chess experience has taken place on a computer rather than 'over-the-board' using tangible chess pieces. Although this difference is unlikely to effect the degree to which expertise influence the ability to quickly determine which piece will be moved, it highlights the importance of manipulation knowledge, in addition to function knowledge, when anticipating the actions of others'.

For example, during an over-the-board game of chess, knowledge about the chess pieces' functions within the game (based on the rules of chess), as well as the specific action components associated with their use (i.e. executing a move), are activated (Bilalic et al., 2011). In other words, we know where a piece can be moved (*function* knowledge), and that achieving this goal requires reaching across the chessboard and grasping the desired piece using a precision grip (*manipulation* knowledge). Over time, the associations between the visual aspects of the objects, the rules associated with their movement, and the motor response they produce are incorporated into an object representation, which is stored in memory (Bilalic et al., 2010, 2011; Borghi & Riggo, 2015). By contrast, two-dimensional chess pieces, viewed on a computer screen, carry no such action information. Rather than reaching and grasping, their movement is defined by the click of a mouse. Thus, online chess players are less likely to develop motor affordances associated with these objects.

It has been argued that function knowledge (what an object is used *for*) supports action *interpretation* by providing insight into the potential goal an action, whereas manipulation knowledge (*how* an object is used) guides action *prediction* by highlighting potential forthcoming actions (Bach et al., 2014). Thus, if motor affordances are necessary to predict forthcoming movements and bias action perception, we might expect to see this in chess experts with a high level of experience in over-the-board chess, versus computer-based chess.

6.6 Motor biases in action perception

The experiments in chapters 2, 3 and 4 tested whether action goal predictions elicit biases in action perception. However, predictive processing models would further hypothesise that predictions also elicit motor biases, based on the assumption that the expected actions of others are derived via the observer's own motor system (Csibra, 2008; Schenke et al., 2021). Due to the limitations imposed by the university's COVID-19 restrictions we were unable to run our original TMS study (Experiment 4a). We therefore aimed to test whether it was possible to develop a RT-based measure of affordances, that could subsequently be used to test our original hypothesis that object knowledge primes motor representations, in the absence of TMS. Similar studies (e.g., Bub & Masson, 2010; Ellis & Tucker, 2000; Makris et al., 2011) have measured grip responses behaviourally using a specifically designed grasp-mimicking device. However, to avoid face-to-face participant testing entirely, the experiments in Chapter 5 were administered online. Participants responded via their keyboard, using either their index finger, or their index and little finger simultaneously. As these are the finger muscles recruited by a PG and WHG, respectively, we reasoned that activation of action representations elicited by small or large objects should facilitate congruent motor responses.

The absence of any significant effects in Experiment 4b may be largely attributable to the complexity of the task and the multiple action representations portrayed in the videos. For example, on each trial, participants observed the actor reach to grasp the prime object (a cigarette or a pencil) with their first hand configured to perform a precision grip, while the 'neutral' grip of the second hand could arguably have been perceived as a whole-hand grip. Thus, both types of grip were arguably being activated on every trial. Participants also responded using both hands; the left hand for PG responses and the right hand for WHG responses, with the intention of ensuring participants' maintained the correct position on the keyboard throughout the task. However, this proved to be problematic given they were observing two different hands performing two different grips, and that participants were predominantly right-hand dominant.

Analysis of catch trial performance in Experiment 4b appeared to confirm that several participants were making action judgements based on the actor's arm movements rather than the objects. When explicitly asked to report which object the actor was most likely to pick up, participants were more likely to select the correct target object when it was presented contralateral to the reaching arm, versus when it was presented ipsilateral to the reaching arm. This suggests that the 'ambiguous' reaching action was in fact not ambiguous, but was perceived by many participants as following a trajectory toward the opposite side of the actor's body (see Figure 5.7). In addition, the left or rightward orientation of the objects on the table was such that it might have implied a reach to grasp with the contralateral hand. Indeed, Tipper, Paul and Hayes (2006, experiment 2) found that the action state of the object influences the activation of affordances. They had participants make key-presses in response to the shape of a door handle, which was oriented to the left or right. They found that active objects, with which current action is implied, produced larger affordance effects than passive objects, with which no action is implied. One might therefore expect that an implied reach toward the contralateral object, and the object orientation indicative of a reach-to-grasp action, would activate the corresponding grip for that object. However, we found no evidence of a spatial compatibility effect when the target object was recoded based on the trajectory of the reaching arm. For example, if the actor was reaching with their right arm, the target object was recoded as the object on the actor's lefthand side. Again, this could be due to the multiple action representations being activated within such a complex task.

6.7 Limitations and future directions

Conducting perception research online

Due to the limitations imposed by the university's COVID-19 restrictions we were unable to conduct any experiments following Experiment 1b in the lab. Nevertheless, this presented us with an opportunity to test the extent to which object affordances can be reliably measured remotely online. The majority of our online studies were conducted via Amazon Mechanical Turk, a popular platform for sourcing convenience samples for cognitive science research. In 2017, 24% of articles in *Cognition*, 29% of articles in *Cognitive Psychology*, 31% of articles in *Cognitive Science*, and 11% of articles in Journal of *Experimental Psychology: Learning, Memory, and Cognition* mention MTurk or another online marketplace (Stewart et al., 2017).

Online testing can present many challenges, including a lack of control over the precise parameters of stimulus presentation and a lack of experimenter supervision when participants are completing the studies. It does, however, open up the possibility of testing large numbers of participants in a comparatively short time (typically less than 24 hours), at a relatively low cost, going beyond the typical constraints of Western, Educated, Industrialised, Rich, and Democratic ('WEIRD', see Henrich, Heine & Norenzayan, 2010) pools of participants who form the basis of the vast majority of psychological research (Woods et al., 2015). Online participants are generally more representative of the population at large than those typically recruited for lab-based studies in that a broader age range and more equal distribution of males and females tend to sign up for studies (Mason & Suri, 2012; Woods et al., 2015). However, comparisons of US MTurk samples to representative samples suggest that participants tend to be more educated, but report lower incomes and are more likely to be unemployed. They are also less religious and more liberal than the population as a whole (Levay et al., 2016; Paolacci & Chandler, 2014). Nevertheless, MTurk allows researchers to recruit from specific sub-populations, selecting a range of criteria for recruitment such as native language, age range, sex and ethnicity. For example, we were able to specify that we wanted right-handed participants from the UK, who spoken English as a first language. We only recruited participants with a high level of experience (> 100 MTurk studies completed) and reputation on MTurk (> 95% task approval ratings); criteria that correlate with attentiveness in online tasks. High-reputation MTurk participants rarely fail attention check questions and provide higher quality data than low-reputation participants (Peer et al., 2014).

One of the concerns with online perception research is the accuracy of the stimulus timing, which often needs to be millisecond specific. In particular, the variability in the hardware used to generate the auditory stimuli in Experiments 4b – 4d could have posed a potential problem for both the timing and quality of the audio tones. For example, some participants may have used headphones during the task, while others may have used speakers. Plant and Turner (2009) found that computer speakers introduced a delay before audio presentation that ranged from 3.31 ms to 37 ms. A consequence of this temporal variability is that it is difficult to know from when exactly reaction times should be measured. Additionally, reaction times can vary depending on the brand of keyboard used by participants. Mean delays between key-presses and reported reaction times can range between 18.30 ms and 33.73 ms for PC computers (Plant & Turner, 2009), and between 19.69 ms and 39.56 ms for Apple computers (Neath et al., 2011). Thus, the use of many different keyboards can introduce a lot of variation in latencies.

The computer monitor can also be a problem in online reaction time tasks given that the web browser does not know when the monitor refreshes, so cannot synchronise the stimulus presentation with a given screen refresh. De Leeuw & Motz (2016) found that web-browser-based reaction times were overstated by around 25 ms compared to lab-based measures. This is a particular problem when large amounts of system resources are in use (Simcox & Fiez, 2014). We strived to overcome some of the problems associated with browser-based studies by running our experiments via Inquisit Web; an experimental software package that participants can temporarily download to their computer, which runs outside of participants' browser. The advantage is that the software executes the same code as is used for offline testing on a local device, although the machines running this code will of course vary more than a those in a well-managed lab. In comparison to Inquisit Lab, which only runs on desktop computers, Inquisit Web can also be downloaded to iOS or Android devices. This enabled us to run touchscreen-based studies (Experiment 1b – 3) on iPads. To eliminate the possibility of participants running these experiments on their iPhones, we instructed Inquisit Web to only run these experiments on iOS devices with a minimum screen resolution of 1024 x 768 pixels. Following these criteria, this meant that the smallest possible device was an iPad mini with a screen size of 7.9" (for a list of iOS devices and their respective screen resolutions, see: https://www.ios-resolution.com).

Another concern with online research is that participants do not take the same care and diligence as those in lab-based studies. However, the number of exclusions made based on catch trial performance in Experiment 1a (*N* = 12), which was conducted in the lab, was comparable that of Experiments 2a-c (mean number of exclusions = 5.6), which were conducted online. However, catch trial accuracy decreased considerably with more complex study designs, such as those reported in Chapters 4 and 5. In Experiment 4b, approximately half of the original sample were excluded based on catch trial performance (<60%). Thus, while we were able to obtain large sample sizes in a relatively short space of time, after excluding participants based on poor catch trial performance, this often left us with a considerably smaller data set on which to conduct our analyses. Online testing is therefore only suitable for simple experimental designs in which the instructions are easily to follow, with no need for clarification from the researcher.

Another limitation that should be acknowledged is the absence of a calibration phase at the beginning of our touchscreen experiments (Experiments 1 – 3). Variation in touchscreen responses can be attributed to the way an individual positions their hand when responding to task stimuli as well as their perception of the centre of the target. A calibration phase is used to maintain the accuracy and standardisation of coordinate measures by accounting for each participant’s spatial bias. A calibration phase typically involves presenting a target at various locations on the screen and asking participants to touch the perceived centre of the target with their index finger. The target remains on the screen until a judgement has been made. Calibration values are calculated in the same way as displacements for test trials (i.e. subtracting the selected screen coordinate from the real coordinate). Mean displacements in each experimental condition are then adjusted by subtracting the corresponding mean horizontal and vertical calibration values for each condition.

Large-scale online studies, combined with lab-based experiments offering finer control over the testing environment and stimuli, may be useful and economical for the future of perception research. Conducting exploratory research online enables the researcher to scope out interesting hypotheses while pruning out the alternatives that have little support (Woods et al., 2015). For example, conducting a series of online experiments in parallel with one another, in quick succession, allowed us to explore multiple avenues of interest, and to refine our experimental design, before spending a comparatively large amount of time and money testing in the lab. In Chapter 5, our aim was to first test for behavioural biases in participants' actions using online reaction time measures. These behavioural measures would then lay the foundation for a lab-based neurophysiological study. The advantage of this approach was that we could first test if biases occur in overt behaviour before moving to the lab to test whether these behavioural biases were unconscious and automatic. As we were unable to establish any affordance effects behaviourally, this allows us time to reconsider our design before testing in the lab.

Is the activation of action goals unconscious and automatic?

Our results show partial support for the hypothesis that action predictions, derived from object knowledge, subtly distort the perceptual representation of observed actions (Chapter 2). The latter part of this thesis aimed to test whether these biases occur not only in overt behaviour, but also if they are unconscious and automatic. One way of assessing this is by using TMS to measure anticipatory muscle activation when viewing object-directed actions (Experiment 4a). Prior research using TMS has shown that during action observation, the observer's motor system simulates under threshold the same muscles as those used in the observed action (Fadiga et al., 1995). When viewing objects, motor evoked potentials (MEPs) from the right hand were larger when viewing objects with a handle oriented to the right, compared to any other orientation. As the size of the MEP recorded from a specific muscle is known to vary with increasing cortical preparation for relevant motor acts (Rösler, & Magistris, 2008), this implies that a right-handed action was being planned to a greater extent in response to right-oriented objects (Buccino et al., 2009a). If objects guide action goal prediction in a manner that is both unconscious and automatic, one might expect larger MEPs in the hand muscles associated with the expected target object, versus an unrelated distractor object which requires a different grip.

An alternative way of testing the automaticity of action goal representations associated with objects would be to use an implicit measure of semantic processing, such as sensory attenuation. Reduced neuronal activity is often reported following repeated processing of the same or similar stimuli. Sensory attenuation occurs when the consequences of self-initiated actions are perceived as weaker, leading to weaker brain responses compared to passively perceived sensory stimuli (Cao & Gross, 2015). This repetition suppression effect has been observed using behavioural measures (Roussel et al., 2013; Schneider et al., 2008), EEG (Roussel et al., 2013) and event-related functional-imaging methods (Grill-Spector et al., 2006). For example, using EEG, Gruber & Müller (2002) found that repeated presentation of familiar visual objects was accompanied by a reduction in total gamma-band activity. Priming effects on event-related potentials (ERPs) have been observed in experiments showing attenuation of the N400 component (a negativity peaking approximately 400 ms after stimulus onset) when words were preceded by related sounds (van Petten & Rheinfelder, 1995). Similar effects have also been observed following repeated presentation of the same object in different sensory modalities (Schneider et al., 2008). These findings provide evidence that predicted outcomes are subject to sensory attenuation. Thus, if viewing an object automatically elicits a representation of its expected use, subsequent presentations of this object, be they visual or audible, should be attenuated. For example, if participants see a pair of scissors, they should pre-activate a representation of the outcome (e.g., cutting paper). If this is the case, when they subsequently hear the sound of paper being cut, both the neural response (auditory N1) and the subjective intensity of the sound should be reduced (i.e., sensory attenuation).

6.8 Conclusion

The findings from this thesis provide evidence that object knowledge biases action perception in the direction of expectations. In Chapter 2, we demonstrated that such inferences are derived, at least in part, from object affordances. Object-directed actions are judged as closer towards a target object when preceded by the presentation of a functionally-related object (e.g., Hammer – nail). This shows that people integrate action kinematic information with prior knowledge about objects, including how they are used and what they are used for, to derive the likely goal of the observed action. This in turn biases perception towards the expected goal. This finding supports predictive processing accounts of social perception, which suggest that action understanding does not rely solely on passively represented perceptual input, but involves actively predicting what will be perceived next, based on inferences about the intended goals of others.

We found no evidence that this perceptual bias could be influenced by prior knowledge about the gender of the actor (Chapter 3) or by participants' expertise with in the available objects (Chapter 4). The potential reasons for the lack of observable differences were discussed, along with suggestions for how we might test these ideas further.

In the final experimental chapter of this thesis (Chapter 5), we explored the potential for motor biases resulting from object-based goal predictions. Constrained by online testing methods, we sought to answer this question behaviourally, using reaction time measures as an index of motor preparation. However, the absence of any significant effects in these experiments suggests that (1) the reaction time measures we developed were not sensitive enough to differentiate the grips afforded by the objects, and (2) the tasks we developed were too complex, particularly for online research. Nevertheless, online perceptual psychology tasks may be a useful tool for exploring multiple avenues of interest prior to conducting larger (and more expensive) lab-based studies. As technology develops, some of the limitations associated with online experiments, such as differences in hardware and timing issues, will become less problematic.

References

Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, *30*(2), 179–188. https://doi.org/10.1007/s11031-006-9020-2

Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology*, *106*(3), 1437–1442. https://doi.org/10.1152/jn.00118.2011

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), 1–9. https://doi.org/10.1371/journal.pone.0067916

Anderson, S. J., Yamagishi, N., & Karavia, V. (2002). Attentional processes link perception and action. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1497), 1225–1232. https://doi.org/10.1098/rspb.2002.1998

Approaches to representational momentum : theories and models. (2010).

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. https://doi.org/10.3758/s13423-010-0029-x

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Frontiers in Human Neuroscience*, *8*(May), 254. https://doi.org/10.3389/fnhum.2014.00254

Bach, P., Nicholson, T., & Hudson, M. (2015). Response: No need to match: a comment on Bach, Nicholson, and Hudson’s “Affordance-Matching Hypothesis.” *Frontiers in Human Neuroscience*, *9*(December), 2015–2017. https://doi.org/10.3389/fnhum.2015.00685

Bach, P., & Schenke, K. C. (2017). Predictive social perception: Towards a unifying framework from action observation to person knowledge. *Social and Personality Psychology Compass*, *11*(7), 1–17. https://doi.org/10.1111/spc3.12312

Bach, P., & Tipper, S. P. (2006). Bend it like Beckham: Embodying the motor skills of famous athletes. *Quarterly Journal of Experimental Psychology*, *59*(12), 2033–2039. https://doi.org/10.1080/17470210600917801

Bargh, J. A., Chen, M., & Burrows, L. (1996). Crowd Psychology and Engineering: Designing for People or Ballbearings? *Journal of Person*, *71*(2), 230–244. http://doi.apa.org/journals/psp/71/2/230.html

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*(3), 437–449. https://doi.org/10.1037/0022-3514.43.3.437

Bartoli, E., Maffongelli, L., Jacono, M., & D’Ausilio, A. (2014). Representing tools as hand movements: early and somatotopic visuomotor transformations. Neuropsychologia, 61, 335-344.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*(3), 644–653. https://doi.org/10.1016/j.cognition.2006.07.012

Bestmann, S., & Duque, J. (2016). Transcranial Magnetic Stimulation: Decomposing the Processes Underlying Action Preparation. *Neuroscientist*, *22*(4), 392–405. https://doi.org/10.1177/1073858415592594

Bilalić, M. (2017). The neuroscience of expertise. *The Neuroscience of Expertise*, 1–

300. https://doi.org/10.1017/9781316026847

Bilalić, M., Langner, R., Erb, M., & Grodd, W. (2010). Mechanisms and Neural Basis of Object and Pattern Recognition: A Study With Chess Experts. *Journal of Experimental Psychology: General*, *139*(4), 728–742. https://doi.org/10.1037/a0020756

Bilalić, M., McLeod, P., & Gobet, F. (2009). Specialization effect and its influence on memory and problem solving in expert chess players. *Cognitive Science*, *33*(6), 1117–1143. https://doi.org/10.1111/j.1551-6709.2009.01030.x

Bilalić, M., McLeod, P., & Gobet, F. (2014). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*(3), 652–661. https://doi.org/10.1016/j.cognition.2008.05.005

Binkofski F., B. L. (2013). *Two Action Systems in the Human Brain*. *127*(2), 222–229. https://doi.org/10.1016/j.bandl.2012.07.007.Two

Bitensky, I., Mama, Y., & Algom, D. (2014). Efficacy of Priming: Evidence from Expert Performance. *Psychology*, *05*(17), 1923–1932. https://doi.org/10.4236/psych.2014.517195

Blair, I. V., & Banaji, M. R. (1996). Automatic and Controlled Processes in Stereotype Priming. *Journal of Personality and Social Psychology*, *70*(6), 1142–1163. https://doi.org/10.1037/0022-3514.70.6.1142

Borghi, A. M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., & Nicoletti, R. (2007). Are visual stimuli sufficient to evoke motor information? *Neuroscience Letters*, *411*(1), 17–21. https://doi.org/10.1016/j.neulet.2006.10.003

Borghi, A. M., & Riggio, L. (2015). Stable and variable affordances are both automatic and flexible. *Frontiers in Human Neuroscience*, *9*(JUNE), 1–16. https://doi.org/10.3389/fnhum.2015.00351

Boronat, C. B., Buxbaum, L. J., Coslett, H. B., Tang, K., Saffran, E. M., Kimberg, D. Y., & Detre, J. A. (2005). Distinctions between manipulation and function knowledge of objects: Evidence from functional magnetic resonance imaging. *Cognitive Brain Research*, *23*(2–3), 361–373. https://doi.org/10.1016/j.cogbrainres.2004.11.001

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating Action Understanding: Inferential Processes versus Action Simulation. *Current Biology*, *17*(24), 2117–2121. https://doi.org/10.1016/j.cub.2007.11.057

Bub, D. N., & Masson, M. E. J. (2010). Grasping Beer Mugs: On the Dynamics of Alignment Effects Induced by Handled Objects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 341–358. https://doi.org/10.1037/a0017606

Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*(1), 27–58. https://doi.org/10.1016/j.cognition.2006.12.010

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic mannen: An FMRI study. *Social Neuroscience: Key Readings*, *9780203496*, 133–142. https://doi.org/10.4324/9780203496190

Buccino, G., Sato, M., Cattaneo, L., Rodà, F., & Riggio, L. (2009a). Broken affordances, broken objects: A TMS study. *Neuropsychologia*, *47*(14), 3074–3078. https://doi.org/10.1016/j.neuropsychologia.2009.07.003

Buccino, G., Sato, M., Cattaneo, L., Rodà, F., & Riggio, L. (2009b). Broken affordances, broken objects: A TMS study. *Neuropsychologia*, *47*(14), 3074–3078. https://doi.org/10.1016/j.neuropsychologia.2009.07.003

Butler, E. E., Ward, R., & Ramsey, R. (2015). Investigating the relationship between stable personality characteristics and automatic imitation. *PloS one, 10*(6), e0129651.

Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: Dissociations in apraxic and nonapraxic subjects. *Brain and Language*, *82*(2), 179–199. https://doi.org/10.1016/S0093-934X(02)00014-7

Buxbaum, L. J., Veramonti, T., & Schwartz, M. F. (2000). Function and manipulation tool knowledge in apraxia: Knowing “What For” but not “How.” *Neurocase*, *6*(2), 83–97. https://doi.org/10.1080/13554790008402763

Canessa, N., Borgo, F., Cappa, S. F., Perani, D., Falini, A., & Buccino, G. (2008). *The Different Neural Correlates of Action and Functional Knowledge in Semantic Memory : An fMRI Study*. *April*. https://doi.org/10.1093/cercor/bhm110

Cao, L., & Gross, J. (2015). Attention wins over sensory attenuation in a sound detection task. *PLoS ONE*, *10*(8), 1–8. https://doi.org/10.1371/journal.pone.0136585

Cardellicchio, P., Sinigaglia, C., & Costantini, M. (2013). Grasping affordances with the others hand: A tms study. *Social Cognitive and Affective Neuroscience*, *8*(4), 455–459. https://doi.org/10.1093/scan/nss017

Castiello, U. (2003). Understanding Other People's Actions: Intention and Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 416–430. https://doi.org/10.1037/0096-1523.29.2.416

Cavina-Pratesi, C., Connolly, J. D., Monaco, S., Figley, T. D., Milner, A. D., Schenk, T., & Culham, J. C. (2018). Human neuroimaging reveals the subcomponents of grasping, reaching and pointing actions. *Cortex*, *98*, 128–148. https://doi.org/10.1016/j.cortex.2017.05.018

Chainay, H., & Humphreys, G. W. (2002). Privileged access to action for objects relative to words. *Psychonomic Bulletin and Review*, *9*(2), 348–355. https://doi.org/10.3758/BF03196292

Chambon, V., Domenech, P., Pacherie, E., Koechlin, E., Baraduc, P., & Farrer, C. (2011). What are they up to? The role of sensory evidence and prior knowledge in action understanding. *PLoS ONE*, *6*(2). https://doi.org/10.1371/journal.pone.0017133

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*, *12*(4), 478–484. https://doi.org/10.1006/nimg.2000.0635

Charness, N. (2001). The Perceptual Aspect of Skilled Perfirmance in Chess, Eye Movement. *Memory & COgnition*, *29*(8), 1146–1152.

Cho, D. T., & Proctor, R. W. (2011). Correspondence effects for objects with opposing left and right protrusions. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 737–749. https://doi.org/10.1037/a0021934

Chong, I., & Proctor, R. W. (2020). On the Evolution of a Radical Concept: Affordances According to Gibson and Their Subsequent Use and Development. *Perspectives on Psychological Science*, *15*(1), 117–132. https://doi.org/10.1177/1745691619868207

Chong, T. T. J., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI Adaptation Reveals Mirror Neurons in Human Inferior Parietal Cortex. *Current Biology*, *18*(20), 1576–1580. https://doi.org/10.1016/j.cub.2008.08.068

Collette, C., Bonnotte, I., Jacquemont, C., Kalénine, S., & Bartolo, A. (2016). The development of object function and manipulation knowledge: Evidence from a semantic priming study. *Frontiers in Psychology*, *7*(AUG), 1–11. https://doi.org/10.3389/fpsyg.2016.01239

Condry, J. C., & Ross, D. F. (1985). Sex and Aggression: The Influence of Gender Label on the Perception of Aggression in Children. *Child Development*, *56*(1), 225. https://doi.org/10.2307/1130189

Correll, J., Park, B., & Judd, C. M. (2002). *The Police Officer ' s Dilemma : Using Ethnicity to Disambiguate Potentially Threatening Individuals*. *83*(6), 1314–1329. https://doi.org/10.1037//0022-3514.83.6.1314

Costantini, M., Ambrosini, E., Scorolli, C., & Borghi, A. M. (2011). When objects are close to me: Affordances in the peripersonal space. *Psychonomic Bulletin and Review*, *18*(2), 302–308. https://doi.org/10.3758/s13423-011-0054-4

Costantini, M., & Stapleton, M. (2016). *How the Body Narrows*. *1979*.

Cutting, J. E. (1982). Two Ecological Perspectives: Gibson vs. Shaw and Turvey.

The American Journal of Psychology, *95*(2), 199. https://doi.org/10.2307/1422466

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*(1), 20–33. https://doi.org/10.1037/0022-3514.44.1.20

Decroix, J., & Kalénine, S. (2019). What first drives visual attention during the recognition of object-directed actions? The role of kinematics and goal information. *Attention, Perception, and Psychophysics*, *81*(7), 2400–2409. https://doi.org/10.3758/s13414-019-01784-7

Dijksterhuis, A., Aarts, H., Bargh, J. A., & Van Knippenberg, A. (2000). On the Relation between Associative Strength and Automatic Behavior. *Journal of Experimental Social Psychology*, *36*(5), 531–544. https://doi.org/10.1006/jesp.2000.1427

Dijksterhuis, A., Macrae, C. N., & Haddock, G. (1999). When Recollective Experiences Matter : *Personality and Social Psychology Bulletin*, *25*(6), 766–774. http://psp.sagepub.com/content/25/6/766.full.pdf+html

Dijksterhuis, A., & Van Knippenberg, A. (1996). The knife that cuts both ways: Facilitated and inhibited access to traits as a result of stereotype activation. *Journal of Experimental Social Psychology*, *32*(3), 271–288. https://doi.org/10.1006/jesp.1996.0013

Dijksterhuis, A., & Van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology*, *74*(4), 865–877. https://doi.org/10.1037/0022-3514.74.4.865

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, *34*(4), 590–598. https://doi.org/10.1037/0022-3514.34.4.590

E. Tory Higgins. (1996). Knowledge activation: Accessibility, applicability, and salience. *Social Psychology: Handbook of Basic Principles*, *June*, 133–168.

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology*, *91*(4), 451–471. https://doi.org/10.1348/000712600161934

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*(6), 2608–2611. https://doi.org/10.1152/jn.1995.73.6.2608

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobstrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013–1027. https://doi.org/10.1037//0022-3514.69.6.1013

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences*, *8*(1), 33–39. https://doi.org/10.1016/j.tics.2003.11.004

Finke, R. A., Freyd, J. J., & Shyi, G. C. (1986). *Implied Velocity and Acceleration Induce Transformations of Visual Memory*. *I*(2), 175–188.

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

https://doi.org/10.1037/0278-7393.10.1.126

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

Friston, K., Square, Q., & Kilner, J. (2011). Europe PMC Funders Group Action understanding and active inference. *Biological Cybernetics*, *104*(1–2), 137–160. https://doi.org/10.1007/s00422-011-0424-z.Action

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Frith, C. D., & Frith, U. (2006). The Neural Basis of Mentalizing. *Neuron*, *50*(4), 531–534. https://doi.org/10.1016/j.neuron.2006.05.001

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, *19*(6), 1239–1255. https://doi.org/10.1093/cercor/bhn181

Gentilucci, M., Caselli, L., & Secchi, C. (2003). Finger control in the tripod grasp. *Experimental Brain Research*, *149*(3), 351–360. https://doi.org/10.1007/s00221-002-1359-3

Gergely, G., & Csibra, G. (2008). Action mirroring and action understanding : an alternative account. *Sensorymotor Foundations of Higher Cognition. Attention and Performance XXII*, 435–459. https://doi.org/10.1093/acprof:oso/9780199231447.003.0020

Girardi, G., Lindemann, O., & Bekkering, H. (2010). Context Effects on the Processing of Action-Relevant Object Features. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 330–340. https://doi.org/10.1037/a0017180

Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, *6*(4), 231–236. https://doi.org/10.1006/nimg.1997.0293

Gregory, R. L. (2005). The Medawar Lecture 2001 knowledge for vision: Vision for knowledge. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1231–1251. https://doi.org/10.1098/rstb.2005.1662

Grèzes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, *40*(2), 212–222. https://doi.org/10.1016/S0028-3932(01)00089-6

Grézes, J., Frith, C., Passingham, R. E., Grèzes, J., Frith, C., & Passingham, R. E. (2004). Brain Mechanisms for Inferring Deceit in the Actions of Others. *The Journal of Neuroscience*, *24*(24), 5500–5505. https://doi.org/10.1523/JNEUROSCI.0219-04.2004

Grèzes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, R. E. (2003). Objects automatically potentiate action: An fMRI study of implicit processing. *European Journal of Neuroscience*, *17*(12), 2735–2740. https://doi.org/10.1046/j.1460-9568.2003.02695.x

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23.

https://doi.org/10.1016/j.tics.2005.11.006

Gruber, T., & Müller, M. M. (2002). Effects of picture repetition on induced gamma band responses, evoked potentials, and phase synchrony in the human EEG. *Cognitive Brain Research*, *13*(3), 377–392. https://doi.org/10.1016/S0926-6410(01)00130-6

Halász, V., & Cunnington, R. (2012). Unconscious effects of action on perception. *Brain Sciences*, *2*(2), 130–146. https://doi.org/10.3390/brainsci2020130

Harel, A., Kravitz, D., & Baker, C. I. (2013). Beyond perceptual expertise: Revisiting the neural substrates of expert object recognition. *Frontiers in Human Neuroscience*, *7*(DEC), 1–12. https://doi.org/10.3389/fnhum.2013.00885

Hommel, B., Forschung, M., & Hommel, B. (1997). *Toward an Action-Concept Model of Stimulus-Response Compatibility*.

Hubbard, T. L. (1993). The effect of context on visual representational momentum. *Memory & Cognition*, *21*(1), 103–114. https://doi.org/10.3758/BF03211169

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin and Review*, *12*(5), 822–851. https://doi.org/10.3758/BF03196775

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, *44*(3), 211–221. https://doi.org/10.3758/BF03206290

Hudson, M., Bach, P., & Nicholson, T. (2018). You said you would! The predictability of other’s behavior from their intentions determines predictive biases in action perception. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(2), 320–335. https://doi.org/10.1037/xhp0000451

Hudson, M., McDonough, K. L., Edwards, R., & Bach, P. (2018a). Perceptual teleology: Expectations of action efficiency bias social perception. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1884). https://doi.org/10.1098/rspb.2018.0638

Hudson, M., McDonough, K. L., Edwards, R., & Bach, P. (2018b). Perceptual teleology: Expectations of action efficiency bias social perception. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1884). https://doi.org/10.1098/rspb.2018.0638

Hudson, M., Nicholson, T., Ellis, R., & Bach, P. (2016). I see what you say: Prior knowledge of other’s goals automatically biases the perception of their actions. *Cognition*, *146*, 245–250. https://doi.org/10.1016/j.cognition.2015.09.021

Hunnius, S., & Bekkering, H. (2010). The Early Development of Object Knowledge: A Study of Infants’ Visual Anticipations During Action Observation. *Developmental Psychology*, *46*(2), 446–454. https://doi.org/10.1037/a0016543

Iacoboni, M. (2009). Imitation, Empathy, and Mirror Neurons. *Annual Review of Psychology*, *60*(1), 653–670. https://doi.org/10.1146/annurev.psych.60.110707.163604

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., & Mazziotta, J. C. (2005). Grasping the intentions of others with one’s own mirror neuron system. *PLoS Biology*, *3*(3), 0529–0535. https://doi.org/10.1371/journal.pbio.0030079

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique.

Trends in Cognitive Sciences, *9*(1), 21–25. https://doi.org/10.1016/j.tics.2004.11.003

Jeannerod, M. (1995). *M. jeannerod*. 1419–1432.

Johnston, L., Miles, L., & Macrae, C. N. (2010). Why are you smiling at me? Social functions of enjoyment and non-enjoyment smiles. *British Journal of Social Psychology*, *49*(1), 107–127. https://doi.org/10.1348/014466609X412476

Jordan, J. S. (2001). *The Role of Perception in the Mislocalization of the Final Position of a Moving Target*. *September*. https://doi.org/10.1037/0096-1523.27.4.829

Joyce, K., Schenke, K., Bayliss, A., & Bach, P. (2016). Looking ahead: Anticipatory cueing of attention to objects others will look at. *Cognitive Neuroscience*, *7*(1–4), 74–81. https://doi.org/10.1080/17588928.2015.1053443

Kellenbach, M. L., Brett, M., & Patterson, K. (2003). Actions speak louder than functions: The importance of manipulability and action in tool representation. *Journal of Cognitive Neuroscience*, *15*(1), 30–46. https://doi.org/10.1162/089892903321107800

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. https://doi.org/10.1016/S0042-6989(00)00226-1

Kerzel, D. (2002). *Memory for the position of stationary objects : disentangling foveal bias and memory averaging*. *42*, 159–167.

Kiesel, A., Kunde, W., Pohl, C., Berner, M. P., & Hoffmann, J. (2009). Playing Chess

Unconsciously. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(1), 292–298. https://doi.org/10.1037/a0014499

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). *The mirror-neuron system : a Bayesian perspective*. *18*(6).

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience*, *29*(32), 10153–10159. https://doi.org/10.1523/JNEUROSCI.2668-09.2009

Kourtis, D., Vandemaele, P., & Vingerhoets, G. (2018). Concurrent Cortical Representations of Function- and Size-Related Object Affordances: An fMRI Study. *Cognitive, Affective and Behavioral Neuroscience*, *18*(6), 1221–1232. https://doi.org/10.3758/s13415-018-0633-1

Kozhevnikov, M., & Hegarty, M. (2001). *Impetus beliefs as default heuristics : Dissociation between explicit and implicit knowledge about motion*. *8*(3), 439–453.

Kret, M. E., & Fischer, A. H. (2018). Recognition of facial expressions is moderated by Islamic cues. *Cognition and emotion, 32*(3), 623-631.

Kret, M. E., & De Gelder, B. (2012). Islamic headdress influences how emotion is recognized from the eyes. *Frontiers in psychology, 3*, 110.

Levay, K. E., Freese, J., & Druckman, J. N. (2016). The Demographic and Political Composition of Mechanical Turk Samples. *SAGE Open*, *6*(1). https://doi.org/10.1177/2158244016636433

Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces:

The role of race categories. *Journal of Experimental Psychology: General*, *135*(4), 501–512. https://doi.org/10.1037/0096-3445.135.4.501

Loach, D., Frischen, A., Bruce, N., & Tsotsos, J. K. (2008). An attentional mechanism for selecting appropriate actions afforded by graspable objects. *Psychological Science*, *19*(12), 1253–1257. https://doi.org/10.1111/j.1467-9280.2008.02234.x

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. Journal of experimental social psychology, 30(4), 370-389.Makris, S., Hadar, A. A., & Yarrow, K. (2011). Viewing objects and planning actions: On the potentiation of grasping behaviours by visual objects. *Brain and Cognition*, *77*(2), 257–264. https://doi.org/10.1016/j.bandc.2011.08.002

Makris, S., Grant, S., Hadar, A. A., & Yarrow, K. (2013). Binocular vision enhances a rapidly evolving affordance priming effect: behavioural and TMS evidence. Brain and cognition, 83(3), 279-287.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon’s Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23. https://doi.org/10.3758/s13428-011-0124-6

Matheson, H. E., White, N. C., & McMullen, P. A. (2014). A test of the embodied simulation theory of object perception: Potentiation of responses to artifacts and animals. *Psychological Research*, *78*(4), 465–482.

https://doi.org/10.1007/s00426-013-0502-z

Matthew Hudson, Toby Nicholson, William A. Simpson, Rob Ellis, P. B. (2016). One step ahead: the perceived kinematics of others' actions are biased towards expected goals. *Journal of Experimental Psychology. General*, *145*(1), 1–7. https://doi.org/10.1017/CBO9781107415324.004

McBride, J., Sumner, P., & Husain, M. (2012). Conflict in object affordance revealed by grip force. *Quarterly Journal of Experimental Psychology*, *65*(1), 13–24. https://doi.org/10.1080/17470218.2011.588336

Mcdonough, K. L., Costantini, M., Hudson, M., Ward, E., & Bach, P. (2020). Journal of Experimental Psychology : Human Perception and Performance Affordance Matching Predictively Shapes the Perceptual Representation of Others ' Ongoing Actions. *Journal of Experimental Psychology: Human Perception and Performance*, *May*. https://doi.org/10.1037/xhp0000745

McDonough, K. L., Hudson, M., & Bach, P. (2019). Cues to intention bias action perception toward the most efficient trajectory. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-42204-y

Meagher, B. R. (2017). Judging the gender of the inanimate: Benevolent sexism and gender stereotypes guide impressions of physical objects. *British Journal of Social Psychology*, *56*(3), 537–560. https://doi.org/10.1111/bjso.12198

Moskowitz, G. B., & Olcaysoy Okten, I. (2016). Spontaneous Goal Inference (SGI). *Social and Personality Psychology Compass*, *10*(1), 64–80. https://doi.org/10.1111/spc3.12232

Motes, M. A., Hubbard, T. L., Courtney, J. R., & Rypma, B. (2008). A Principal

Components Analysis of Dynamic Spatial Memory Biases. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(5), 1076–1083. https://doi.org/10.1037/a0012794

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. https://doi.org/10.1016/j.cub.2010.02.045

Nicholson, T., Roser, M., & Bach, P. (2017). Understanding the goals of everyday instrumental actions is primarily linked to object, not motor-kinematic, information: Evidence from fMRI. *PLoS ONE*, *12*(1), 1–21. https://doi.org/10.1371/journal.pone.0169700

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Visuo-motor imagery of specific manual actions: A multi-variate pattern analysis fMRI study. *NeuroImage*, *63*(1), 262–271. https://doi.org/10.1016/j.neuroimage.2012.06.045

Oosterhof, N. N., Wiggett, A. J., Diedrichsen, J., Tipper, S. P., & Downing, P. E. (2010). Surface-based information mapping reveals crossmodal vision-action representations in human parietal and occipitotemporal cortex. *Journal of Neurophysiology*, *104*(2), 1077–1089. https://doi.org/10.1152/jn.00326.2010

Otten, M., Seth, A. K., & Pinto, Y. (2017). A social Bayesian brain: How social knowledge can shape visual perception. *Brain and Cognition*, *112*, 69–77. https://doi.org/10.1016/j.bandc.2016.05.002

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*, *23*(3), 184–188. https://doi.org/10.1177/0963721414531598

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*(4), 1023–1031. https://doi.org/10.3758/s13428-013-0434-y

Phillips, J. C., & Ward, R. (2002). S-R correspondence effects of irrelevant visual affordance: Time course and specificity of response activation. *Visual Cognition*, *9*(4–5), 540–558. https://doi.org/10.1080/13506280143000575

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. https://doi.org/10.1162/jocn.2006.18.12.2130

Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 839–850. https://doi.org/10.1037/0096-1523.22.4.839

Reingold, E. M., Charness, N., Schultetus, R. S., & Stampe, D. M. (2001). Perceptual automaticity in expert chess players: Parallel encoding of chess relations. *Psychonomic Bulletin and Review*, *8*(3), 504–510. https://doi.org/10.3758/BF03196185

Rholes, W. S. (1977). Category accessibility and impression formation by higgins, rholes and jones. *Journal of Experimental Social Psychology*, *4*, 141–154.

Riggio, L., Iani, C., Gherri, E., Benatti, F., Rubichi, S., & Nicoletti, R. (2008). The role of attention in the occurrence of the affordance effect. *Acta Psychologica*, *127*(2), 449–458. https://doi.org/10.1016/j.actpsy.2007.08.008

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., & Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiological Reviews*, *94*(2), 655–706. https://doi.org/10.1152/physrev.00009.2013

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

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). and Imitation of Action. *Neuroscience*, *2*(September), 1–10. https://doi.org/10.1038/35090060

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*(4), 264–274. https://doi.org/10.1038/nrn2805

Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. *Neuropsychologia*, *51*(5), 922–929. https://doi.org/10.1016/j.neuropsychologia.2013.02.005

Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in Black and White children’s perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology*, *39*(4), 590–598. https://doi.org/10.1037/0022-3514.39.4.590

Sartori, L., Betti, S., & Castiello, U. (2013). Corticospinal excitability modulation during action observation. *Journal of Visualized Experiments : JoVE*, *82*, 51001. https://doi.org/10.3791/51001

Sartori, L., Bucchioni, G., & Castiello, U. (2013). When emulation becomes reciprocity. *Social Cognitive and Affective Neuroscience*, *8*(6), 662–669. https://doi.org/10.1093/scan/nss044

Sartori, L., Cavallo, A., Bucchioni, G., & Castiello, U. (2012). From simulation to reciprocity: The case of complementary actions. *Social Neuroscience*, *7*(2), 146–158. https://doi.org/10.1080/17470919.2011.586579

Schenke, K. C., Wyer, N. A., & Bach, P. (2016). The things you do: Internal models of others' expected behaviour guide action observation. *PLoS ONE*, *11*(7), 1–22. https://doi.org/10.1371/journal.pone.0158910

Schenke, K. C., Wyer, N. A., Tipper, S. P., & Bach, P. (2021). Predictive person models elicit motor biases: The face-inhibition effect revisited. *Quarterly Journal of Experimental Psychology*, *74*(1), 54–67. https://doi.org/10.1177/1747021820945270

Schippers, M. B., & Keysers, C. (2011). Mapping the flow of information within the putative mirror neuron system during gesture observation. *NeuroImage*, *57*(1), 37–44. https://doi.org/10.1016/j.neuroimage.2011.02.018

Schneider, T. R., Debener, S., Oostenveld, R., & Engel, A. K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage*, *42*(3), 1244–1254. https://doi.org/10.1016/j.neuroimage.2008.05.033

Schneider, T. R., Engel, A. K., & Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Experimental Psychology*, *55*(2), 121–132. https://doi.org/10.1027/1618-3169.55.2.121

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*(JUN), 1–15. https://doi.org/10.3389/fpsyg.2014.00636

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. https://doi.org/10.1080/13506280143000331

Spratling, M. W. (2016). Predictive coding as a model of cognition. *Cognitive Processing*, *17*(3), 279–305. https://doi.org/10.1007/s10339-016-0765-6

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons. *Journal of Personality and Social Psychology*, *37*(10), 1660–1672. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.4255&rep=rep1&type=pdf%0Apapers3://publication/uuid/82E8F314-2114-4124-941F-1ED75786B42F

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. https://doi.org/10.1007/s00426-012-0423-2

Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in Cognitive Sciences*, *21*(10), 736–748. https://doi.org/10.1016/j.tics.2017.06.007

Symes, E., Ellis, R., & Tucker, M. (2005). Dissociating object-based and space-based affordances. *Visual Cognition*, *12*(7), 1337–1361. https://doi.org/10.1080/13506280444000445

Tipper, S. P., & Bach, P. (2011). The face inhibition effect: Social contrast or motor competition? *Journal of Cognitive Psychology*, *23*(1), 45–51. https://doi.org/10.1080/20445911.2011.451888

Tipper, S. P., Paul, M. A., & Hayes, A. E. (2006). Vision-for-action: The effects of object property discrimination and action state on affordance compatibility effects. *Psychonomic Bulletin and Review*, *13*(3), 493–498. https://doi.org/10.3758/BF03193875

Tipper, S. P., Weaver, B., & Houghton, G. (1994). Behavioural Goals Determine Inhibitory Mechanisms of Selective Attention. *The Quarterly Journal of Experimental Psychology Section A*, *47*(4), 809–840. https://doi.org/10.1080/14640749408401098

Todd, A. R., Molden, D. C., Ham, J., & Vonk, R. (2011). The automatic and co-occurring activation of multiple social inferences. *Journal of Experimental Social Psychology*, *47*(1), 37–49. https://doi.org/10.1016/j.jesp.2010.08.006

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), 830–846. https://doi.org/10.1037/0096-1523.24.3.830

Tucker, M., & Ellis, R. (2001). The Potentiation of Grasp Types During Visual Object Categorization. *Visual Cognition*, *8*(6), 769–800. https://doi.org/10.1017/CBO9781107415324.004

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica*, *116*(2), 185–203. https://doi.org/10.1016/j.actpsy.2004.01.004

Unkelbach, C., Forgas, J. P., & Denson, T. F. (2008). The turban effect: The influence of Muslim headgear and induced affect on aggressive responses in the shooter bias paradigm. *Journal of Experimental Social Psychology*, *44*(5), 1409–1413. https://doi.org/10.1016/j.jesp.2008.04.003

Valyear, K., Borghi, A., Craighero, L., Fadiga, L., Rizzolatti, G., Umiltà, C., Umana, F., & Parma, U. (n.d.). *Visuomotor priming Related papers*.

Valyear, K. F., Cavina-Pratesi, C., Stiglick, A. J., & Culham, J. C. (2007). Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage*, *36*(SUPPL. 2), 94–108. https://doi.org/10.1016/j.neuroimage.2007.03.031

Valyear, K. F., Gallivan, J. P., Adam McLean, D., & Culham, J. C. (2012). fMRI repetition suppression for familiar but not arbitrary actions with tools. *Journal of Neuroscience*, *32*(12), 4247–4259. https://doi.org/10.1523/JNEUROSCI.5270-11.2012

van Elk, M., van Schie, H. T., & Bekkering, H. (2009). Action Semantic Knowledge About Objects Is Supported by Functional Motor Activation. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1118–1128. https://doi.org/10.1037/a0015024

Van Overwalle, F., Van Duynslaeger, M., Coomans, D., & Timmermans, B. (2012). Spontaneous goal inferences are often inferred faster than spontaneous trait inferences. *Journal of Experimental Social Psychology*, *48*(1), 13–18. https://doi.org/10.1016/j.jesp.2011.06.016

van Petten, C. K., & Rheinfelder, H. (1995). Conceptual Relationships Between Words and Environmental Sounds: Event-related Brain Potentials Measures. In

Neuropsychologia (Vol. 33, Issue 4, pp. 485–508). http://www.sciencedirect.com/science/article/pii/002839329400133A

Vingerhoets, G., Vandamme, K., & Vercammen, A. (2009). Conceptual and physical object qualities contribute differently to motor affordances. *Brain and Cognition*, *69*(3), 481–489. https://doi.org/10.1016/j.bandc.2008.10.003

Vinson, N. G., & Reed, C. L. (2002). *Representational Momentum * Sources of Object-Specific Effects in Representational Momentum*. *9*(February), 1–18.

Westra, E. (n.d.). *Stereotypes , Theory of Mind , and the Action - Prediction Hierarchy*. 1–43.

Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: A tutorial review. *PeerJ*, *2015*(7). https://doi.org/10.7717/peerj.1058

Wurm, M. F., & Schubotz, R. I. (2017). What's she doing in the kitchen? Context helps when actions are hard to recognize. *Psychonomic Bulletin and Review*, *24*(2), 503–509. https://doi.org/10.3758/s13423-016-1108-4

Wyk, B. C. Vander, Hudac, C. M., Carter, E. J., Sobel, D. M., & Pelphrey, A. (2010). Action Understanding in the Superior Temporal Sulcus Region. *Psychological Science*, *20*(6), 771–777. https://doi.org/10.1111/j.1467-9280.2009.02359.x.Action

Yu, A. B., Abrams, R. A., & Zacks, J. M. (2014). Limits on action priming by pictures of objects. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1861–1873. https://doi.org/10.1037/a0037397

Yuille, A., & Kersten, D. (2006). Vision as Bayesian Inference : Analysis by Synthesis ? Introduction : Perception as inference. *Trends in Cognitive Sciences*, *10*(7), 301–308.