Sensory prediction mechanisms in action

Daniel Yon

Birkbeck, University of London

Thesis submitted to the Faculty of Science at the University of London for the degree of Doctor of Philosophy

I, Daniel Yon, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgments

My first acknowledgment is to my supervisor Clare Press, whose insight and clarity of thought over the past four years have been matched only by her warmth and generosity. I feel very lucky to have received an apprenticeship as intellectually stretching and good-humoured as this one, and am thankful for all of your support.

I have had the good fortune to share these years with a number of remarkable people. Of particular note is my lab partner Rosy Edey, whose accomplishments as a psychologist are matched only by her virtuoso karaoke performances. Thanks are also due to Jono Batten, Nick Berggren, James Chard, Manu Ducrocq, Rebecca Nako, Jenny Glennon, Michael Papasavva and Sinead Rocha, without whom Birkbeck would have been considerably less bright. I was also lucky to find myself with a ready-made set of ‘aunts’, ‘uncles’ and ‘cousins’ across King’s and City who made PhD life all the richer, and I am particularly grateful to Richard Cook for our stimulating discussions and his continuing encouragement.

The research reported herein has benefitted from input from a great many others, notable among them Sam Gilbert, Richard Ivry and Floris de Lange. Harish Patel was an enormous help, as were the staff at the FIL. There is a strong chance I would never have considered a PhD had I not encountered Charles Spence and his infectious sense that lots of interesting answers were only one experiment away. I also am very grateful for the support of the ESRC, and I am beholden to Michael Thomas for the steps he took to secure the award.

I am indebted to my family and friends, especially Rebecca Nohl for her continued interest and enthusiasm. I owe the greatest debt of all to Rosa Schiller Crawhurst, whose wit, wisdom, love and support have made the past years what they were. The highs have been all the headier and the lows much easier to bear with you by my side to share them.

My final acknowledgment is to my late father, Peter Yon. Though you may not have pored over every page of this thesis, I would never have been able to accomplish this (or much else) without all that you gave me. I dedicate this thesis to my abiding memory of this kind and gentle man, who is deeply loved and sorely missed.

Some of the work reported in this thesis has been published in the following papers:

Chapter 2 – Yon, D., & Press, C. (2017). Predicted action outcomes are perceptually facilitated before cancellation. *Journal of Experimental Psychology: Human Perception and Performance,* 43(6), 1073-1083.

Chapter 5 – Yon, D., & Press, C. (in press). Sensory predictions during action support perception of imitative reactions across suprasecond delays. *Cognition.*

Chapter 6 – Yon, D., Edey, R., Ivry, R.B. & Press, C. (2017). Time on your hands: Perceived duration of sensory events biased towards concurrent actions. *Journal of Experimental Psychology: General, 146*(2), 182-193.

During my PhD I have also contributed to the following papers:

Edey, R., Yon, D., Cook, J., Dumontheil, I. & Press, C. (2017). Our own action kinematics predict the perceived affective states of others. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1263-1268.

Yon, D., & Cook, R. (2016). The imitation game: can newborn babies mimic their parents? *The Conversation.*

Yon, D., & Press, C. (2014). Back to the future: synaesthesia could be due to associative learning. *Frontiers in Psychology,* 5:702.

Abstract

When we produce an action we generate predictions about the sensory consequences that are likely to ensue. This thesis tests a series of claims about the functional contribution these predictions make to perception, the role that such predictions play in processing the reactions of others, and the range of sensory inputs that these prediction mechanisms operate over. Chapter 1 outlines the theoretical background to each of these claims, alongside the previous literature that motivates subsequent experiments.

The first three empirical chapters focus on claims about the functional role of sensory predictions during action: that they act to ‘cancel’ perception of expected action outcomes. Chapter 2 investigates this hypothesis in the context an intensity judgement task, Chapter 3 tests the hypothesis in the context of a signal detection task and Chapter 4 assess how predictions generated during action influence multivariate measures of visual brain activity, recorded via functional magnetic resonance imaging.

Chapter 5 investigates the claim that sensory predictions during action support the processing of imitative reactions in others. Two psychophysical experiments are reported which investigate whether sensory predictions generated during action have temporal properties needed to support processing of others’ reactions.

Chapter 6 investigates whether sensory predictions generated during action influence the ‘when’ - as well as the ‘what’ - of perception. Four psychophysical experiments investigate whether the temporal features of executed actions are incorporated into duration perception. Chapters 7 and 8 report preliminary investigations into the mechanism underlying these effects. Chapter 7 assesses whether these influences arise through a mechanism that is primarily tuned to biological action outcomes. Chapter 8 investigates whether these effects arise as a result of statistical learning about the relationship between actions and outcomes.

Chapter 9 summarises the studies presented in the thesis, and outlines their implications for thinking about sensory prediction during action.

Contents

List of figure legends and page numbers

Chapter 1: Introduction

This thesis is concerned with two developments in psychology, and the relationship between them. First, the idea that expectations exert a top-down influence on perceptual processing has gained traction across cognitive science. A range of theoretical and empirical work challenges the old orthodoxy that what we perceive is driven in a purely 'bottom-up' fashion. Following a venerable tradition (Helmholtz, 1860), it is instead argued that perception unfolds through the combination of inherently ambiguous sensory evidence with prior knowledge of the environment in which we find ourselves. In recent decades, this predictive processing approach has been supported by a range of behavioural and neural data showing that perceptual judgements – and the neural mechanisms supporting them – are influenced by expectations (Clark, 2013).

Second, there is a growing appreciation of the interactions between perception and action. The classical division of labour in experimental psychology assumed that perception and action proceeded largely independently of one another, with the motor system serving solely as the output buffer for other cognitive processes occurring 'upstream' (Neisser, 1967). This division has been undermined by a large body of work that has revealed that perceptual and motor processes interact to optimise both the control of action and our representation of the sensory world.

These separate theoretical developments – that expectations and actions influence perception – may seem superficially similar; both suggest that perception does not proceed purely in a bottom-up fashion, but is instead shaped by other sources of top-down information. However there are significant contrasts between theories developed to explain how action predictions influences perception and those developed more generally to account for effects that expectations have on perceptual performance. Given these contrasts, this thesis therefore first investigates more closely the nature of the predictive mechanisms deployed during action and whether they really contrast with mechanisms deployed outside of action contexts. Specifically, it considers the claim that top-down *motor* predictions have a functionally dissociable influence on perceptual processing (functional-specificity).

Second, using theoretical developments in both fields, it investigates the inputs over which motor predictions operate. It asks whether motor predictions influence our ability to process the actions of other agents as well as our own action effects (agent-specificity), and whether predictions generated during action influence both the 'what' and 'when' of perceptual estimates, investigating predictive motor contributions to time perception (modality-specificity). This Introduction outlines the theoretical background to each of these strands, drawing out explicit comparisons between models developed in the action literature and those developed to account for predictive effects in other contexts. The Introduction closes with an outline of how the identified open questions are investigated in the empirical chapters that follow.

1.1. What are 'top-down effects' and 'expectations'?

This thesis examines both theories developed to explain top-down predictive *motor* contributions to perception and those developed in sensory cognition that aim to account for influences of top-down expectations on perceptual processing *in general*. I will begin by explaining what is meant by the terms 'top-down' and 'expectations' in the context of this thesis.

Psychologists tend to describe influences as 'top-down' when the operation of a psychological process is shaped by prior knowledge. This is contrasted with 'bottom-up' processes which are driven primarily by current input (Eysenck, 1998). While the distinction between top-down and bottom-up processes partly reflects a traditional view on information processing – with simpler sensory input mechanisms receiving input at the 'bottom' that is passed to higher-order functions at the 'top' – it is also influenced by how neuroscientists have described the hierarchical organisation of functions in the cortex (e.g. Mechelli, Price, Noppeney, & Friston, 2003). Indeed recent theoretical work describing the functional cytoarchitecture of the cortex describes 'top-down' effects as influences which are mediated by descending neural projections into deep cortical layers of 'lower' brain areas and 'bottom-up' effects are those which are mediated by ascending neural projections into the superficial cortical layers of 'higher' brain areas (Friston, 2005). However, the psychological distinction between top-down and bottom-up effects is not at heart about which kind of neural projection mediates the influence, and it is in the psychological sense – of prior information influencing processing of current input (Shea, 2015) – that the term is used throughout this thesis.

Top-down effects of this kind have been the focus of considerable theoretical and empirical work in sensory cognition, with researchers distinguishing different top-down effects based on the kinds of representation that shape perceptual processing. For example, Summerfield and Egner (2009; 2016) distinguish effects of top-down expectation from top-down attention: expectation effects are driven by representations of what is likely to occur (e.g. a statistical association) whereas attentional effects are driven by representations about what is *relevant* in a given environment (e.g. task goals). These definitions of 'top-down' and 'expectation' (the latter of which is used synonymously with 'prediction') are used in the present thesis. Importantly, the definition thus drawn often excludes effects on perceptual processing caused by the structure of the body and brain - which some consider to be 'expectations' acquired over the course of evolution (Plotkin, 1997).

With these definitions in mind we turn to the specific ways that sensory predictions generated by the motor system have been hypothesised to operate, and where this contrasts with models developed to explain top-down influences of expectation general.

1.2. Functional-specificity: Do sensory predictions during action influence perception in a special way?

It has been appreciated for some time that effective action control depends on anticipating the consequences of our movements (James, 1890). However, while the role of sensory prediction in action selection and execution is well-established (Greenwald, 1970; Hommel, Musseler, Aschersleben, & Prinz, 2001), in recent decades interest has developed in how predictions generated by the motor system influence the perception of action outcomes. This section outlines the predominant model used to explain perceptual prediction during action (the Cancellation model – Section 1.2.1), and will contrast this model with those used to describe effects of prediction outside of action (Section 1.2.2). It will then describe an alternative 'domain-general' account (the OPPOSE model) which is consistent with both literatures but does not assume a special role for motor prediction (Section 1.2.4).

1.2.1. Prediction during action 'cancels' perception of expected outcomes

Research into sensory prediction during action has largely followed the Cancellation model proposed by Wolpert and colleagues (Blakemore, Wolpert, & Frith, 2000; Wolpert, Ghahramani, & Jordan, 1995). Under this model, when an action is initiated a forward model of the motor system predicts the sensory consequences that the movement will produce. These predictions are sent to sensory brain regions, and are subtracted from the sampled input (i.e. expected sensory units are suppressed; Bays & Wolpert, 2007), leading to attenuated perception of expected action outcomes. It is argued that such a 'cancellation' mechanism is functionally adaptive, as it allows actors to process preferentially unexpected events that are more likely to require learning or a novel response. For example, if when lifting a cup of coffee actors attenuate processing of expected sensory input (e.g. touch on the finger tips, sight of the moving cup) relative to unexpected input (e.g. sight of spilling coffee), they will be better placed to perform corrective actions to avoid spillage or learn about the dynamics of the sensory environment (e.g. perhaps the cup is lighter than anticipated).

The Cancellation model provides an explanation for the observation that it is difficult to tickle oneself (Weiskrantz, Elliott, & Darlington, 1971), and has drawn support from a range of behavioural studies that show events predictable on the basis of an executed action are perceived as less intense. For example, participants rate self-produced brushing sensations as less ticklish than those produced by a robot (Blakemore, Frith, & Wolpert, 1999) and self-produced taps are rated as less intense than equivalent forces produced by a machine (Bays, Wolpert, & Flanagan, 2005). Similar effects are seen outside of tactile domains; Weiss, Herwig and Schütz-Bosbach, (2011) report that self-produced tones appear quieter than those produced by an experimenter, while dot motion congruent with a keypress appears slower than motion in an incongruent direction (Dewey & Carr, 2013). Signal detection tasks have also suggested that observers are less sensitive to low contrast arrows and Gabor patches when their orientations are congruent with an executed action (Cardoso-Leite, Mamassian, Schütz-Bosbach, & Waszak, 2010; Müsseler & Hommel, 1997a; 1997b).

The Cancellation model has also drawn support from neuroimaging studies that find predictable action outcomes are associated with reduced activity in sensory brain regions. In touch, self-produced tactile sensations elicit reduced blood oxygenation level dependent (BOLD) responses in secondary somatosensory cortex when compared to externally-produced sensations (Blakemore, Wolpert, & Frith, 1998; Shergill et al., 2013; 2014), while in vision Stanley and Miall (2007) find attenuated BOLD responses in primary visual cortex (V1) when participants observe gestures congruent with those they are executing. Similar attenuation effects are seen in higher visual brain areas implicated in action perception (such as the superior temporal sulcus) when participants view actions congruent with their own movements (Kontaris, Wiggett, & Downing, 2009; Leube et al., 2003). In all cases, effects are interpreted as evidence for a ‘cancellation’ mechanism – with sensory predictions acting to suppress expected sensory inputs (Bays & Wolpert, 2007).

1.2.2. *Prediction outside of action facilitates perception of expected outcomes*

The influence that motor predictions are hypothesised to exert on perceptual processing contrasts with the functional role assigned to prediction by sensory cognition models outside of action. Following Helmholtz (1860), a number of theorists have emphasised that the sensory environment is inherently noisy, and that the patterns of stimulation that impinge on sensory receptors are insufficient to form reliable estimates of the outside world. This problem can be finessed by incorporating prior expectations into our perceptual estimates (Summerfield & de Lange, 2014). For example, when walking in heavy fog we may see ambiguous figures in the distance. Knowledge of our sensory context (e.g. that we are in a remote forest or an urban car park) is hypothesised to furnish our perceptual systems with expectations about likely features of the environment (e.g. that we can expect to encounter trees or cars, respectively) that aid interpretation of the ambiguous input, such that we are more likely to perceive what we expect. It is argued to be adaptive to bias perception in line with our expectations, as expected sensory events are by definition more likely to occur.

The idea that perception is formed as a synthesis between sensory evidence and expectations is the key claim of 'Bayesian Brain' approaches, which argue that human observers combine prior knowledge and sampled evidence in a manner that approximates the norms of Bayesian inference (Yuille & Kersten, 2006). Hierarchical predictive coding (Friston, 2005; Rao & Ballard, 1999) - a particularly influential form of the Bayesian Brain hypothesis - suggests that the brain attempts to create a generative model of its environment by predicting the activity of neural populations at lower levels of the processing hierarchy (e.g. projections from V2 predict the activation of populations in V1) and passing back prediction errors to refine subsequent predictions. However, irrespective of the particular implementation of predictive processing to which one subscribes, it is typically assumed that expectations are incorporated into and improve perception. In other words, we are more, rather than less likely, to perceive what we expect - in contrast with Cancellation models in action.

This idea draws support from studies which show valid expectations have a facilitatory influence on perceptual performance (Bar, 2004). Detection studies show that participants are more accurate when detecting stimuli that are congruent with expectations. For instance, Sekuler and Ball (1977) found that participants were better able to detect visual motion when the stimulus was compatible with a probabilistic cue, while Palmer (1975) reports that presenting observers with a particular sensory context (e.g. a kitchen) leads to more accurate detection of objects one would likely encounter within it (e.g. a loaf of bread). Studies using continuous flash suppression have also revealed that expectations can accelerate the entry of predicted stimuli into conscious awareness (Pinto, van Gaal, de Lange, Lamme, & Seth, 2015). Recent evidence also suggests that expected stimuli may appear phenomenally more intense. Han and VanRullen (2016) report that the apparent luminance of a grey disk is increased when it appears in the context of a three dimensional shape compared to the context of a random lines. Moreover a number of well-known illusions can also be explained as biases induced by prior knowledge. For example, in the hollow face illusion, participants are likely to perceive a concave face as convex, which may reflect the fact that convex faces are a much more frequent (and therefore more probable) feature of the environment (Gregory, 1970).

It is worth noting that both in the action and sensory cognition literatures, 'expectation' is manipulated in a variety of ways. Typically in the action literature, 'expected' sensory outcomes are those which are 'congruent' with action – i.e., matching action outcomes such as seeing an index finger move when one is moving one's index finger. Congruent events are by their nature likely to be expected given that they are the conditionally most probable consequence of a given movement (Shea, 2015), and agents have typically had vast amount of experience of these contingencies when learning to control their own actions (e.g. Rochat, 1998). Similar types of manipulation - where expectations are assumed to have been generated on the basis of a contingency experienced outside the lab - are sometimes seen in the sensory cognition literature, e.g., where bread is more probable in the context of a kitchen (Palmer, 1975) or where shapes are probable than random lines (Han & VanRullen, 2016). However, sensory cognition studies also often introduce new correlations between arbitrary events (e.g. where the frequency of an auditory tone predicts the orientation of a Gabor patch or a direction of dot motion; (Kok, Brouwer, van Gerven, & de Lange, 2013; Kok, Jehee, & de Lange, 2012) that allows participants to develop expectations over the course of an experiment. This thesis assumes that congruency effects that exploit expectations acquired over a lifetime's experience and expectation effects acquired as a result of recent laboratory learning reflect the operation of similar processes. Indeed, the logic of the present thesis assumes that congruent events are indeed more expected than incongruent events and this assumption appears uncontroversial. However, the broader implications of this assumption are considered in greater detail in Section 9.3.

Interestingly, fMRI studies looking at the effect of prediction on perceptual processing find that events predictable on the basis of a contextual cue are associated with reduced BOLD activity in sensory brain areas (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010), mirroring effects that have been reported to support the Cancellation action models (e.g. Blakemore et al., 1998). However, these reductions in univariate BOLD activity are not typically argued to reflect a suppression of expected sensory signals, but a sharpening of population responses. For example Kok, Jehee and de Lange (2012) used fMRI to investigate how expectations alter univariate and multivariate measures of visual brain activity. The authors found that when the orientation of a grating was validly predicted by a prior probabilistic cue, univariate activity in early visual cortex was reduced (relative to trials with invalid cues). However, multivariate pattern analyses (MVPA) of the same data revealed superior classification of stimulus orientation when expectations were valid. Such results – along with the behavioural findings of facilitatory influences of expectation in similar contexts – suggest that expectations act to enhance the quality of underlying sensory representations and that univariate reductions in signal do not reflect cancellation.

There is thus a stark contrast between the theoretical models and empirical results relating to sensory prediction during action, and those relating to predictive influences outside of action contexts. On the one hand, a cancellation of events consistent with sensory predictions is argued to privilege the processing of *unexpected* events during action due to their behavioural relevance - evidenced by a range of studies revealing that predicted action outcomes appear less intense, and are harder to detect. On the other hand, theorists interested in expectation more generally emphasise the role of sensory prediction in privileging the processing of *expected* events, improving the reliability of our perceptual estimates by incorporating our prior knowledge. This is evidenced by a number of studies showing precisely the opposite effects – that predictable outcomes are easier to detect and appear as more intense. Arranged in this way, a functionally-specific explanation for motor predictions appears to provide the best account of the data.

1.2.3. Sensory prediction during action facilitates perception of expected outcomes

However, there are theoretical and empirical issues with this apparent disjunct between prediction during action and prediction of other kinds. Theoretically, it is not clear that the adaptive arguments apply differently in action and non-action contexts. In principle, it is just as important in both types of context to be sensitive to expected and unexpected outcomes. To use the previous example, if an observer in a remote forest comes across a parked car (an unexpected event) it is still adaptive to devote resources that will promote learning or planning a new course of action. Comparably, if we are attempting to drink a cup of coffee in a noisy or ambiguous environment (e.g. a dark kitchen before sunrise) we will generate more veridical percepts of the ongoing action if we increase the weight given to expected sensory signals (e.g. the sight of the moving hand) over unexpected ones.

Empirically, a number of findings also suggest that prediction during action may not influence the percept so differently from prediction in other contexts. For example, trained pianists are biased to hear an ambiguous sequence of rising or falling tones in line with keypresses that typically produce either a rising or falling sequence (Repp & Knoblich, 2007), and participants are biased to report ambiguous clockwise/anticlockwise dot motion in line with a concurrent hand movement (Wohlschläger, 2000). Related effects are obtained under binocular rivalry, where observed hand actions (Di Pace & Saracini, 2014) or rotating spheres (Maruya, Yang, & Blake, 2007) are more likely to dominate the observer's percept when they are congruent with ongoing hand movements. Such results are reminiscent of those seen in predictive contexts outside of action e.g. in illusions where perception is biased toward expected outcomes (Gregory, 1970).

1.2.4. *The OPPOSE model: Optimising Perception of Predicted Outcomes and Surprising Errors*

The results reviewed thus far suggest a paradox in the way that predictions generated during action influence perception. On the one hand, expected action outcomes are typically perceived as less intense than unexpected outcomes – consistent with the idea that predictions have been 'cancelled' from perception (Blakemore et al, 1998). On the other hand, predictions generated during action can bias ambiguous inputs towards expected outcomes (e.g. Wohlschläger, 2000) – consistent with the idea that predicted sensory inputs are 'facilitated' by expectation, in line with models of predictive processing developed outside of action contexts. These latter findings suggest that contrast between prediction in action and non-action domains may not be as clear cut as previously assumed.

It is possible to explain both facilitation and cancellation effects on perception without assuming the motor predictions operate differently from sensory predictions generated in other contexts. Recently Yon and Press (2017) suggested that both kinds of effect could be explained if it is assumed two processes optimise perception during action – one which increases the weight given to expected information, and another which is deployed when surprising errors are detected. Here this hypothesis is called the OPPOSE model (Optimising Perception of Predicted Outcomes and Surprising Errors). Under the OPPOSE model a primary prediction process during action is hypothesised to operate identically to prediction mechanisms described in other areas of sensory cognition (Summerfield & de Lange, 2014; Yuille & Kersten, 2006).

Such a prediction mechanism would be expected to amplify expected sensory signals, accounting for findings where participants are biased towards perceiving (Wohlschläger, 2000) expected action outcomes. However, under this model it is assumed that when observers detect unexpected events during action (i.e. prediction errors), despite the primary mechanisms rendering such detection less likely than that of expected events, these outcomes subsequently become the target of secondary processes which facilitate the processing of surprising outcomes. These processes operate post-perceptually – i.e., following the perception of the surprising event. One candidate mechanism driving such a process is spatial attention, as eye-tracking paradigms have previously shown observers overtly attend to spatially or temporally surprising outcomes (Itti & Baldi, 2009). As spatial attention has been shown to increase apparent intensity (Carrasco, Ling, & Read, 2004), such an orienting mechanism would lead to a relative enhancement in the perceived intensity of unexpected action outcomes. These relative enhancements of unexpected events could generate cancellation effects – where expected action outcomes are perceived as relatively *less* intense (Blakemore et al., 1999; Weiss et al., 2011) in the absence of any process that actively suppresses expected inputs, or that operates before inputs are received. Under this OPPOSE model, it may therefore be possible to jointly optimise perception of events that are likely to occur and events that are likely relevant to behaviour due to their surprising nature.

The OPPOSE models makes two predictions about the conditions under which action should facilitate perception of expected outcomes and when cancellation should occur. First, the prediction process necessarily precedes any process that depends on prediction-error: therefore facilitation of expected action outcomes should be observed before later cancellation. This prediction was recently supported by the results of a visual task by Yon and Press (2017), where participants rated the intensity of congruent and incongruent action outcomes at different delays after movement: participants reported more intense percepts for congruent action outcomes when perception was probed 50 ms after movement, while more intense percepts for incongruent action outcomes were observed at a 200 ms delay.

The second prediction of the OPPOSE model is that the relative contribution of prediction and prediction-error mechanisms should depend on the strength of the sensory signal. In particular, cancellation effects are assumed to occur because surprising, unexpected events have been oriented towards – but an event cannot be surprising if it is not detected. Therefore when events are presented at threshold (as in signal detection experiments) experiments should generally find that predicted action outcomes are more readily detected, with less evidence of cancellation because unpredicted outcomes are not always detected. This prediction is not shared by cancellation theorists, who assume that predictions act to suppress sensory input directly and regardless of signal strength (Bays & Wolpert, 2007). However, a vast majority of studies looking at the influence of prediction during action have used suprathreshold stimuli and this prediction therefore remains largely untested.

The hypothesis that prediction during action operates equivalently to other forms of prediction may also suggest an alternative interpretation of cancellation results at the neural level. As noted previously, the fact that tactile

(Blakemore et al., 1998; Shergill et al., 2013, 2014) and visual (Kontaris et al., 2009; Leube et al., 2003; Stanley & Miall, 2007) brain activity is reduced for expected action events has been interpreted as evidence of cancellation. However, as described in Section 1.2.2, work looking at the effects of prediction outside of action has found that univariate reductions in BOLD signal can be accompanied by increases in multivariate classification accuracy – indicative of more informative underlying neural representations (Kok et al., 2012). To date similar multivariate measures have not been applied to the question of prediction during action, and therefore it is unclear whether the univariate reductions in sensory BOLD reported in previous studies reflect a suppression of expected input or a sharpening of underlying sensory representations. The former is predicted under the Cancellation model, though the latter would be found if motor prediction mechanisms operate equivalently to other kinds of prediction.

1.2.5. Summary and open questions

Influential models of motor prediction have suggested that expectations during action alter perception in the opposite fashion to that described by models of top-down prediction developed outside of action. However, current work is also consistent with a model where motor predictions operate equivalently to other kinds of prediction (the OPPOSE model; note that this model would be additionally expected to apply to prediction outside of action contexts). These models make a number of contrasting predictions. First, under Cancellation models, intensity judgements of suprathreshold stimulation should always be lower for events expected on the basis of action, relative to unexpected events.

However, as already outlined, the OPPOSE model predicts early high intensity judgements, followed by later low intensity judgements. These OPPOSE predictions were met in two experiments reported in Yon & Press (2017), and Chapter 2 presents an experiment that tests an alternative account of these effects (Experiment 3 in Yon & Press, 2017). Second, Cancellation models predict reduced detection of events expected on the basis of action, whereas the OPPOSE model predicts enhanced detection. These predictions are distinguished in Chapter 3. Finally, Cancellation models predict that the suppression of expected sensory signals will lead to reduced multivariate pattern classification of expected action outcomes, whereas the OPPOSE model predicts that expectations may 'sharpen' predicted signals, leading to superior multivariate pattern classification performance. These predictions are distinguished in Chapter 4.

1.3. Agent-specificity: Do sensory predictions influence how we perceive the actions of others?

Section 1.2 described the functional influence that prediction mechanisms during action exert on perceptual processes and considered whether this contrasts with the functional role assigned to prediction in other contexts. It therefore considers the domain-specificity of prediction mechanisms in action relative to predictions that operate based on other kinds of information (e.g. sensory context). The latter sections of this Introduction consider a different type of domain-specificity of prediction mechanisms during action – namely the range of inputs over which they operate. This section considers whether predictions during action only influence how we perceive our own action outcomes, or whether they may also influence how we perceive the reactions of other agents.

1.3.1. *The role of sensory prediction during social interaction*

A popular suggestion in recent decades has been that the motor system plays an important role in supporting social understanding and interaction by furnishing our ability to process the observed actions of others. A major impetus for theorising on this topic was the discovery of mirror neurons – a class of visuomotor units found in the ventral premotor and inferior parietal cortices of the macaque that respond both when the monkey executes an action, and when the monkey observes the same action performed by the experimenter (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996).

Different researchers have suggested that mirror populations in premotor areas are the neural instantiation of the 'forward model' hypothesised in predictive approaches to action control (Kilner, Friston, & Frith, 2007; Miall, 2003; Wolpert, Doya, & Kawato, 2003), and it has been hypothesised that the models of our own action can therefore be repurposed to support social understanding. Indeed, there is considerable evidence in humans that homologous motor brain regions are activated during passive observation of others' actions (Buccino et al., 2001; Oosterhof, Tipper, & Downing, 2013), with studies of neurological patients and 'virtual lesions' induced through transcranial magnetic stimulation suggesting that such motor activity makes a causal contribution to action recognition (Pazzaglia, Smania, Corato, & Aglioti, 2008; Pobric & Hamilton, 2006).

However, independently of proposals that motor structures contribute to action understanding when we passively observe others, other researchers have suggested that predictions made during active movement may finesse social interactions (Schütz-Bosbach & Prinz, 2007; Wolpert et al., 2003). These models have emphasised the fact that when we interact with others (e.g. when we wave to a friend) the perceptual consequences of our own actions are similar to those generated by our social partners. Given this similarity, the same mechanisms that generate predictions about our own movements may be useful for perceiving the imitative reactions of others. Though comparatively little work has investigated these ideas, Schütz-Bosbach and Prinz (2007) have speculated this may be 'the most important social function: it might render an individual selectively susceptible to *similar actions of conspecifics*'(p.354, emphasis added).

1.3.2. Are our own sensorimotor models suited to the prediction of others' actions?

The assumption that the consequences of our actions are perceptually similar to the reactions of others might seem reasonable given the broad similarities between the morphology of human bodies and kinematics with which they move (e.g., Fitts, 1954). However, this assumption overlooks substantial differences between how the consequences of our own actions and the reactions of others are typically processed. In particular, when we produce an action (e.g. a wave) the sensory consequences produced are always matching (a wave) and occur at predictable delays in the order of milliseconds, while the reactions of others are only sometimes matching (imitative) and occur at variable delays after our own movements in the order of seconds. If predictive models of our own actions are to serve the hypothesised social functions (Schütz-Bosbach & Prinz, 2007), such models would need to generate predictions with sufficient generality across time (Kunde, Weller, & Pfister, 2017).

A handful of studies in the tactile domain examining the temporal precision of prediction during action suggest that predictions may not operate across such suprasecond delays. For example, predictive effects on perceived tactile force are seen when action execution and sensory outcomes are simultaneous, but are absent when delays of ~300 ms are imposed between action and outcomes (Bays et al., 2005). Such effects are interpreted to suggest that the underlying predictions are temporally-specific, and moreover this subsecond precision has been argued to be necessary if the predictions are to be useful in supporting the rapid detection of errors and the initiation of corrective actions (Wolpert et al., 1995). However, we experience the reactions of others primarily through vision, and it is unclear whether visual predictions operate with the same subsecond tuning as described in touch, or whether they operate with a temporal profile that could support the prediction of others' reactions.

1.3.3. Summary and open questions

Models advanced in the social cognition literature have suggested that predictive models of our own actions may be used for predicting similar reactions of others – i.e., when we are imitated by a partner in social interaction. An important but untested assumption of these theories is that predictions generated during movement generalise across suprasecond delays, which may be difficult to reconcile with previous reports from the tactile domain that such mechanisms operate with subsecond precision. However, no work to date has directly considered this question within vision, and in Chapter 5 psychophysical experiments are reported which investigate whether predictive influences on visual intensity judgements are modulated by suprasecond action-effect delays.

1.4. Modality-specificity: Do sensory predictions during action influence temporal features of perception?

The previous section considered the generality of sensory predictions during action, describing how predictions may (or may not) generalise from our own actions to the reactions of others. The following section considers a different form of generality – namely whether predictive mechanisms influence *when,* as well as *what,* we perceive during action. A large body of work suggests that during movement we generate predictions about *what* will change in the environment – and that these expectations about the spatial properties or identity of outcomes influence perception. However, little work has considered whether predictions are also generated about the temporal features of action outcomes, such as *when* changes will occur and how they will evolve over time. The following section describes why such temporal integration could be adaptive (Section 1.4.1), before outlining different models of time perception that make contrasting predictions about the plausibility of such integration (Section 1.4.2.).

1.4.1. *Predicting 'when' during action*

As already outlined in Section 1.2.2, sensory cognition models highlight how perceptual systems face a considerable challenge in generating accurate estimates of the environment from sensory inputs corrupted by noise. One strategy that an observer might use to deal with this uncertainty is to use top-down information about the likelihood of different inputs. The adaptive argument outlined earlier applies equivalently to the temporal features of action and perception, which tend to be highly correlated. For example, when a cellist plays a note there is a strong correlation between the duration of the executed movement and the sensory consequences that are produced (duration of the produced note, duration of visual motion of the bow etc.) – and incorporating these predictive relationships into perception during action should generate more veridical estimates of time. However, while it is well-known that action execution can influence the subjective passage of time (Haggard, Clark, & Kalogeras, 2002; Merchant & Yarrow, 2016), it is not known whether the temporal features of movements are integrated into perceptual estimates in a comparable way to that which has previously been demonstrated for *what* information (e.g. where the perception of ambiguous rotating motion is biased towards concurrently executed hand rotations - Wohlschläger, 2000; see Section 1.2.3).

1.4.2. *Dedicated and intrinsic models of time perception*

The question of temporal integration between action and perception must necessarily consider some of the ongoing debates in the time perception literature. In particular, there is ongoing controversy concerning whether time perception depends on mechanisms specifically dedicated to representing temporal information, or whether it arises as an intrinsic property of networks that represent other features (Ivry & Schlerf, 2008). For example one 'dedicated' timing model, the cerebellar timing hypothesis, suggests that duration information important for perception and action is represented by populations of interval timers akin to hourglasses – where particular event durations activate specific representations (Ivry, 1996). In contrast, 'intrinsic' timing models, such as those offered by Buonomano and colleagues (Buonomano, 2000; Karmarkar & Buonomano, 2007) assume that time is not explicitly represented but arises as a property of neural processing dynamics in modality-specific brain regions. These models suggest for example that the duration of a visual stimulus presentation leads to characteristic changes in the spatial distribution of visual activity, and that judgements about the duration of a stimulus are made by learning to recognise these characteristic changes – not by accessing a metric of time.

Though these models differ in many respects, an important distinction between dedicated and intrinsic models of time is the possibility of interactions in time perception across modalities. Under dedicated models, information about event duration in different modalities is represented explicitly and can therefore be combined. In contrast, under intrinsic models there are no temporal representations per se, and the information used to make temporal judgements differ considerably across modalities, precluding integration (Ivry & Schlerf, 2008).

However, little work has considered whether temporal features are integrated online in a comparable fashion to spatial features in the sensorimotor tasks described above (Section 1.2.3) – that is, it is unknown whether time perception in one modality is biased by expectations established through another modality, such as action. While such integration could be adaptive and may be possible under dedicated timing models, intrinsic models of time perception predict that such interactions should not occur.

1.4.3. Summary and open questions

A number of experiments suggest that spatial and identity features of perception (‘*what’* information) are influenced by action, and these effects can be explained through a predictive process – where perceptual estimates are biased towards expectations established through movement. Considerably less attention has been paid to whether temporal features of perception (*‘when’* information) are similarly influenced by action. While a biasing of time perception in line with expectations during action may be adaptive, existing models of time perception make divergent predictions about whether temporal features of action could be integrated into perceptual estimates. The possibility that perception of temporal features of sensory stimuli is biased towards those of executed actions is investigated in Chapter 6. Chapters 7 and 8 subsequently present preliminary experiments addressing the specific nature of the underlying mechanisms supporting these effects.

1.5. Thesis outline

Chapter 2- 4 investigate the claim that predictions during action have a functionally-specific influence on perceptual processing. Chapter 2 reports a discrimination experiment looking at influences of prediction during action on intensity judgements (Experiment 1), while Chapter 3 reports a signal detection experiment (Experiments 2) that investigates how prediction during action influences detection performance. Chapter 4 reports an fMRI experiment (Experiment 3) investigating how prediction during action influences univariate and multivariate measures of visual brain activity.

Chapter 5 investigates claims about the agent-specificity of predictions during action, and reports two psychophysical experiments (Experiments 4 and 5) that examine whether effects of action on the perception of predicted outcomes generalise across delays of the nature encountered during social interaction. Chapter 6 assesses claims about the modality-specificity of predictions during action and reports four psychophysical experiments (Experiments 6 – 9) that examine whether action influences how temporal features are perceived.

Chapters 7 and 8 report preliminary investigations into the specific nature and origins of the predictive mechanisms identified in Chapter 6. Chapter 7 reports a psychophysical experiment investigating the tuning of sensory predictions to biological action stimuli (Experiment 10) while Chapter 8 reports a laboratory training study that examines how sensory predictions may alter when contingencies between action and perception are manipulated (Experiment 11).

Chapter 2: Influences of action on intensity judgements

The Cancellation model suggests that perceptual processing of expected action outcomes is suppressed. Support for this hypothesis has come from experiments which report that expected action outcomes are perceived as less intense. However, the alternative OPPOSE model suggests that predictions generated during action bias sensory activity towards (and thereby increase the perceived intensity of) expected outcomes, with previous 'cancellation' effects likely reflecting later prediction-error dependent processes. Evidence for the OPPOSE model was recently found in a task where congruent action outcomes were judged to be more intense at short delays after action execution, and incongruent action outcomes were judged to be more intense at a later delay. Experiment 1 was conducted to extend these findings by investigating whether these effects reflect a perceptual or decisional bias. Participants produced finger actions while observing congruent or incongruent actions performed by an avatar hand. After a delay (50 or 200 ms) the observed finger increased in brightness, and participants rated its intensity relative to a comparison stimulus. Results revealed that congruent action outcomes were rated to be more intense at short delays after action execution. While this effect is difficult to reconcile with the Cancellation model, it accords with the predictions of the OPPOSE model and is consistent with the idea that expectation leads to enhanced activity in expected sensory units.

2.1. Introduction

In Chapter 1 two models were introduced describing how predictions generated during action should influence perception. The Cancellation model hypothesises that predicted action outcomes are ‘subtracted’ from perception by suppressing expected sensory activity at the input stage (Bays & Wolpert, 2007). In contrast, the OPPOSE model hypothesises that predictions generated during action act to amplify expected sensory inputs in a similar way to has been described for other kinds of expectation (Summerfield & de Lange, 2014). However, when observers detect unexpected events (i.e. prediction errors), these events benefit from an additional process dedicated to surprising outcomes e.g. the overt orienting of spatial attention (Itti & Baldi, 2009).

Both accounts are consistent with evidence that participants tend to rate events congruent with executed actions as less intense – such as findings that self-produced tactile sensations are less ticklish (Blakemore et al., 1999) or self-produced tones appear quieter (Weiss et al., 2011). Under the Cancellation model, relatively less intense perception of congruent action outcomes is taken as evidence for a direct suppression of expected sensory activity at the input stage (Bays & Wolpert, 2007). In contrast, under the OPPOSE model congruent action outcomes appear relatively less intense because incongruent outcomes have had their apparent intensity increased by prediction-error dependent processes.

However, the two models make divergent predictions about the timecourse of these effects. Under the OPPOSE model, the primary prediction process necessarily precedes a later process that depends on prediction-errors. As such, at short delays after action execution perceptual processing should only reflect the influence of the early prediction process, whereas influences of prediction-error dependent processes should be detectable at later delays. In an intensity judgement task, the OPPOSE model would therefore hypothesise that when perception is probed at short delays after action, congruent sensory outcomes would appear as more intense, given evidence that in non-action contexts predictable outcomes appear with greater intensity (Han & VanRullen, 2016). Alternatively when perception is probed at longer delays after action, incongruent sensory outcomes would appear with greater intensity. These predictions about the timecourse of predictive effects are not shared with the Cancellation model, which as noted above suggests that expected sensory activity is suppressed from the outset (Bays & Wolpert, 2007).

These predictions were probed in a recent set of psychophysical experiments by Yon and Press (2017). Here participants executed finger movements (index or middle) while observing synchronised lifts on an avatar hand, which could be congruent (same finger) or incongruent (opposite finger) with their own action. After action the observed finger increased in apparent brightness and participants rated the target's intensity relative to a comparison. Importantly, the target's brightness increase could occur at early (50 ms) or late (200 ms) delays after action, allowing the authors to probe perceptual processing at different timepoints while keeping the action-effect delay constant. These experiments revealed that when perception was probed at an early delay congruent action outcomes were rated as more intense than incongruent outcomes, whereas at the later delay the reverse pattern was observed – with incongruent outcomes appearing more intense. Moreover, in a second experiment neither effect was observed for non-action stimuli occurring at the same spatial locations, suggesting that underlying mechanisms were sensitive to the identity of the action outcomes rather than simple spatial features.

This pattern is consistent with the OPPOSE model's predictions, and is harder to explain under the Cancellation model. However, a limitation in the two experiments described is the use of a comparative judgement – participants were required to indicate which of two events (target or comparison) was more intense. When comparative intensity judgements are used, shifts in the point of subjective equivalence (PSE; see Section 2.2.1) across conditions can indicate differences in the perceived intensity of a stimulus. However, differences in PSE can also be driven by influences on decision processes (Schneider & Komlos, 2008). For example, participants may perceive all events with equivalent intensity, but differences in PSE may arise if participants are biased to select one response option. This may be a particular concern given that expectations may bias activity in decision circuits before a stimulus is even presented (Hanks, Mazurek, Kiani, Hopp, & Shadlen, 2011).

The limitations associated with comparative judgements can be overcome through the use of an equality judgement, where participants do not indicate which of two events was the more intense but whether their intensity was the same or different (Han & VanRullen, 2016; Schneider & Komlos, 2008). In an equality judgement task the value of the PSE is not affected by biases to select one response option (same or different). To investigate whether the effects reported by Yon and Press (2017 – Experiments 1 and 2) reflect perceptual or response biases we replicated the paradigm used, but required participants to make equality rather than comparative judgements. If effects were driven by biases to select one response alternative these should not be found when an equality judgement is used. In contrast, if effects occur at a perceptual locus they should persist under this procedural change.

2.2. Experiment 1

2.2.1. Method

Participants

Twenty-six participants (19 female, mean age 24.7 years, SD =4.1) were recruited from Birkbeck, University of London and paid a small honorarium for their participation. Four of these were replacements for participants who could not complete the perceptual discrimination (modelled PSEs were beyond the range of presented stimuli and/or acceptable psychometric functions could not be modelled to their responses- see below).

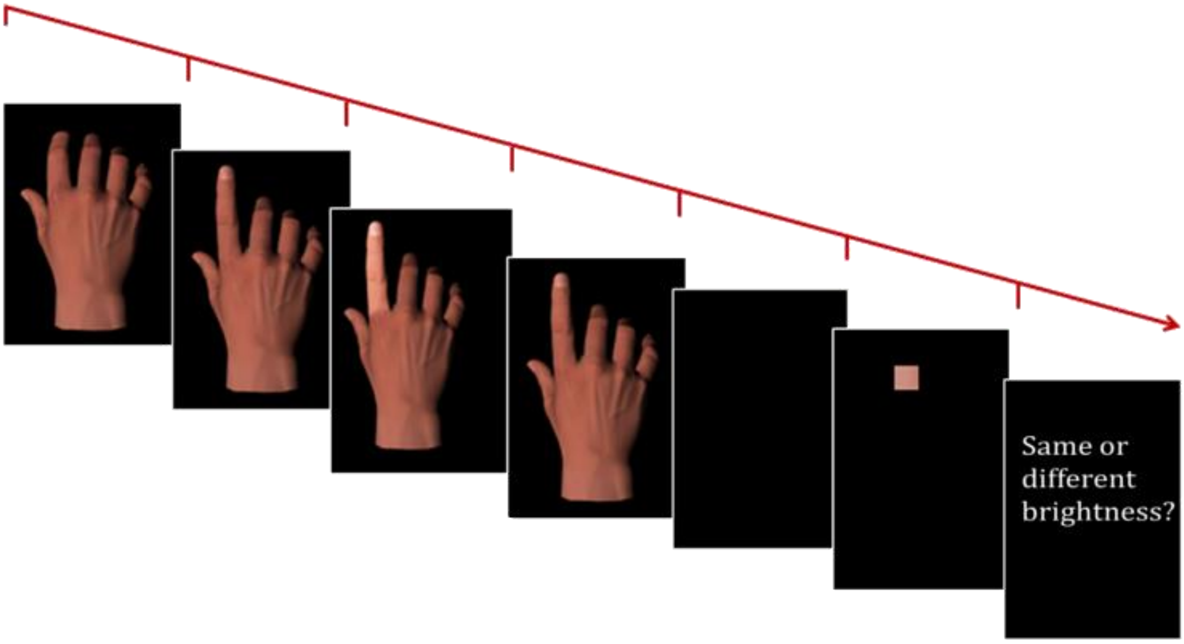

Figure 2.1: Timecourse of presentation and stimuli in Experiment 1

Procedure and Stimuli

The experiment was conducted in MATLAB using the Cogent toolbox. Participants began each trial holding down two keys on a keypad with their right index and little finger. At the start of the trial, a hand at rest was presented on a computer monitor (Fig 2.1; CRT monitor, 32 x 24 cm, 85 Hz, 21 DPI). The participant's hand was visually occluded, and rotated 90° with respect to the observed hand, such that both index and middle finger movements were at body midline – breaking spatial compatibility between executed and observed actions (Catmur & Heyes, 2011). Participants were instructed to lift either their index or middle finger, producing large rapid single movements. Executed actions were freely selected, and participants were instructed to perform roughly equal numbers of each movement in a random sequence. When participants lifted their finger, the neutral hand image was immediately (within 11.8 ms, given the screen refresh rate) replaced by one depicting the hand performing either an index or middle finger lift. This sequence resulted in apparent motion of the observed finger approximately synchronized with the participant's action. Participants could observe a congruent action outcome (e.g. execute index lift, observe index lift) or an incongruent action outcome (e.g. execute index lift, observe middle finger lift), and both outcome types were presented with equal probability (see below). At a variable time after the observed and executed lift (50 ms or 200 ms), the finger flashed for 100 ms at one of seven intensities (increased brightness by 10-70 %, in 10 % steps) before returning to its original brightness level for a further 300 ms. Following a 1000 ms ISI, a reference square – with equivalent brightness to the target midpoint (40% of the target continuum) - was presented for 100 ms (see Fig. 2.1).

After 400-500 ms participants judged whether the target or reference events were the same or different brightness, responding with a keypress made with their left thumb. They subsequently returned their right lifted finger to the start key, with their finger lifted throughout the trial until this point. The next trial started after 1000 ms.

Participants completed at least 280 trials; 70 at each of the two delays in the congruent and incongruent conditions. The experiment was divided into four blocks. The first three blocks each comprised 70 trials, while the fourth ran until participants had completed 140 trials of each lift. In breaks between blocks participants were given feedback on-screen regarding the distribution of their responses. Responses beyond the 140th trial for each movement were not recorded.

To estimate psychometric functions, responses for each participant were modelled by fitting Gaussians (see Figure 2.2). This procedure was performed separately for congruent and incongruent response data for each delay level. In each condition, bias was inferred from the PSE (the mean of the fitted Gaussian) and precision from the difference threshold. The PSE describes the point where participants judge the target and reference events to have equal brightness, with lower values indicative of brighter target percepts. Judgement precision was inferred from the standard deviation of the fitted Gaussian, with lower thresholds reflected more consistent categorisations, thereby indicating better performance (see Fig. 2.2).

2.2.2. Results

PSE and precision values were analysed via separate repeated measures ANOVAs, with factors of Action Congruency (congruent, incongruent) and Delay (50 ms, 200 ms). No significant effects were found in the precision data (all $p \geq .165$). However, the PSE analysis revealed a significant main effect of Delay, $F(1,25) = 18.911$, $p<.001$, $\eta_p^2 = .431$, alongside a significant interaction between Delay and Action Congruency, $F(1,24) = 7.125$, $p = .013$, $\eta_p^2 = .222$. This interaction reflected the fact that PSEs at the 50 ms delay were lower for congruent outcomes (mean = 34.0 %, SEM = 2.44) than incongruent outcomes (mean = 36.3 %, SEM = 2.62; $t(25) = 2.326$, $p=.028$, $d = .174$), while PSEs at the 200 ms were higher for congruent outcomes (mean = 40.2 %, SEM = 2.37) than incongruent outcomes (mean = 39.2 %, SEM = 2.62) – although this difference did not reach statistical significance ($t(25) = 1.613$, p=.119; see Figure 2.2).

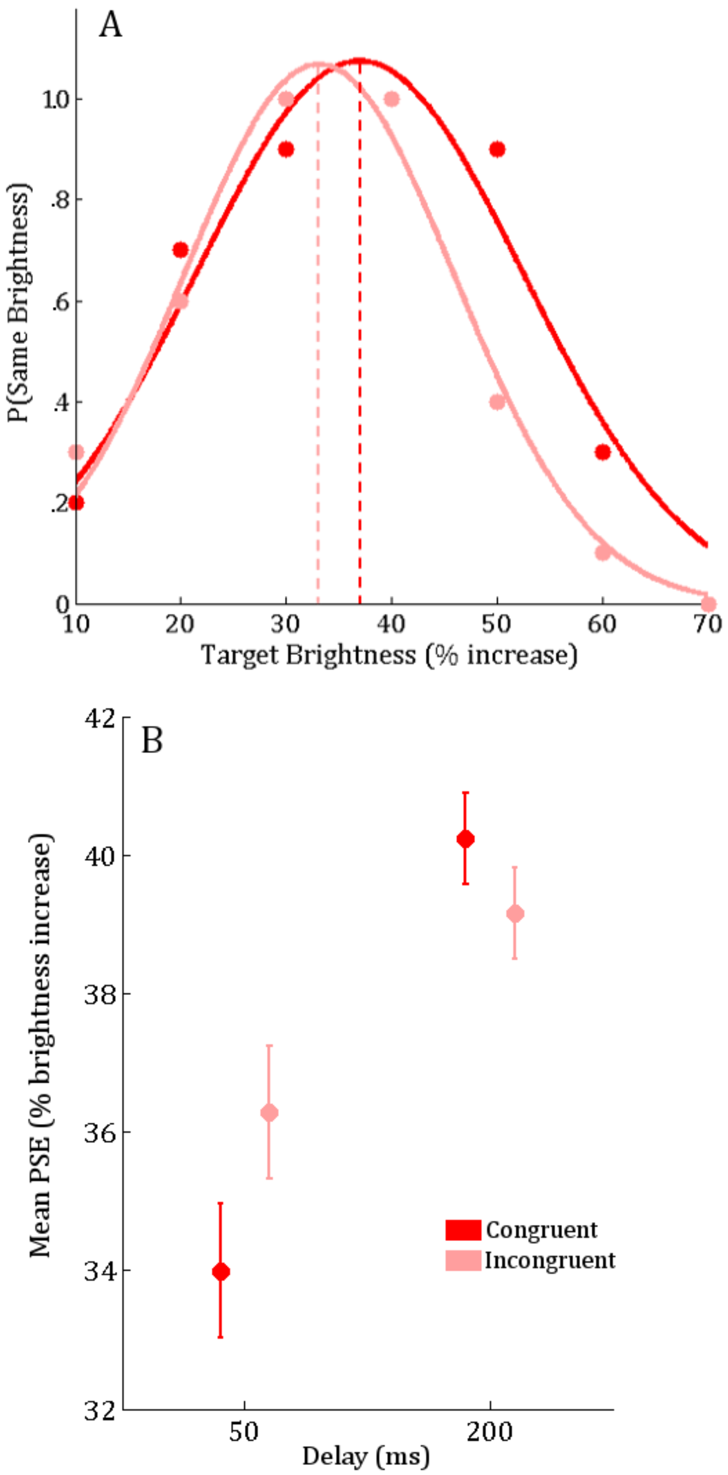

Figure 2.2: A. Demonstration of how the PSE was calculated in Experiment 1 with psychometric functions for an example participant, for stimuli congruent (saturated) and incongruent (faded) with action. B. Mean PSEs across congruency conditions at each delay across all conditions in Experiment 1. Error bars represent 95% within-participant confidence intervals (Cousineau, 2005).

2.2.3. Discussion

This broad pattern of results matches those obtained by Yon and Press (2017; Experiments 1 and 2), with effects of action congruency varying as a function of timecourse, suggesting that previous demonstrations of these effects were unlikely to be driven by response bias. Our experiment replicates the early advantage for expected action outcomes, with significantly lower PSEs (i.e. brighter percepts) for congruent relative to incongruent outcomes. However, we did not observe as convincing evidence for a later cancellation effect. A likely reason for this difference with respect to previous experiments is that equality judgements are more difficult, leading to noisier PSE estimates. Indeed, previous work explicitly comparing judgement types in similar psychophysical tasks suggests that equality judgements have reduced sensitivity to effects on perceived intensity when compared to comparative judgements (Anton-Erxleben, Abrams, & Carrasco, 2010). Nonetheless this experiment finds convincing evidence that at short delays after action expected action outcomes appear as more intense, and that trends toward more intense perception of unexpected action outcomes only emerge at later delays.

The results of Experiment 1 are consistent with the predictions of the OPPOSE model. Under this model, a primary prediction process is hypothesised to facilitate processing of expected action outcomes (increasing the apparent intensity of predicted outcomes) while a secondary prediction-error dependent process is deployed when unexpected events are detected (increasing the apparent intensity of unpredicted outcomes). Importantly, as prediction-errors are necessarily computed after predictions, effects of the former kind should be seen at early delays after action whereas effects of the latter kind should be seen at later delays – as was reflected in the interaction found in Experiment 1, mirroring the results of Yon and Press (2017; Experiments 1 and 2). These results are harder to reconcile with the Cancellation model, which assumes that sensory activity associated with expected outcomes is directly suppressed at an early processing stage (Bays & Wolpert, 2007).

The OPPOSE model suggests that predictions generated during action influence perceptual processing in a comparable way to other kinds of predictive signal – such as those based on sensory context (Summerfield & de Lange, 2014). Indeed, the early effect identified in Experiment 1 closely resembles effects which suggest that grey disks appearing in a predictable shape context appear with greater contrast than equivalent edges presented amongst random lines (i.e. a context that does not support generation of predictions; Han & VanRullen, 2016). Han and VanRullen (2016) tentatively suggest that their effect of shape contexts on apparent contrast reflects excitatory predictive feedback that increases neuronal responses in sensory areas. The results of Experiment 1 are consistent with the possibility that similar predictive mechanisms operate during action.

Through which mechanisms might ‘excitatory predictive feedback’ increase the apparent intensity of congruent action outcomes? Computational models in other areas of psychophysics have suggested that perceived intensity judgements scale with the activation in sensory units (Cutrone, Heeger, & Carrasco, 2014). It has been suggested that effects of expectation in non-action contexts can be understood as operating through an additive mechanism, with predictions biasing activity in expected sensory units (Summerfield & de Lange, 2014; Wyart, Nobre, & Summerfield, 2012) possibly by sharpening the tuning curves of sensory units expected to be stimulated (Summerfield & Tsetsos, 2015). Such predictive enhancements may therefore be expected to produce enhancements in perceived contrast – consistent with the early enhancement in perceived intensity for congruent action outcomes seen in Experiment 1 and Yon and Press (2017).

2.3. Chapter summary

The Cancellation and OPPOSE models make contrasting predictions about how expectations during action should influence the apparent intensity of sensory outcomes. Experiment 1 suggests that at early delays after action execution expected action outcomes appear with greater intensity, with the effect changing at delay – in line with the predictions of the OPPOSE model.

Chapter 3: Influences of action on detection performance

The Cancellation model hypothesises that perceptual processing of expected action outcomes is suppressed – making these events harder to perceive relative to unexpected outcomes. Conversely, the OPPOSE model hypothesises that expectations during action increase the activation of expected sensory units, entailing that observers are more likely to perceive expected outcomes relative to unexpected ones. Experiment 2 was conducted to compare the contrasting predictions of these accounts in a signal detection task. Participants executed finger movements and were required to detect the congruent (i.e. expected) or incongruent (i.e. unexpected) movements observed on an avatar hand, giving confidence ratings alongside their perceptual decisions. Signal detection measures (Green & Swets, 1966) were calculated from both detection judgements and confidence ratings to quantify how participant's objective and subjective sensitivity and bias varied as a function of action-outcome congruency. While no significant effects were found in the metacognitive measures, Experiment 2 did reveal that participants show greater perceptual sensitivity (higher d') to congruent action outcomes, and were also more liberal in reporting their presence (lower c). These results are difficult to reconcile with the Cancellation model, but are consistent with the OPPOSE model and the idea that expectations bias sensory populations towards expected outcomes.

3.1. Introduction

The Cancellation model suggests that sensory predictions generated during movement are used to suppress perception of expected action outcomes. This account has traditionally drawn support from studies which suggest that expected action outcomes are perceived as phenomenally less intense (e.g. you cannot tickle yourself; Blakemore et al, 1999). In Sections 1.2.4 and 2.1, an alternative account of sensory prediction during action was outlined – the OPPOSE model. Under this model prediction during action should have similar facilitatory effects on detection as are seen in passive contexts (e.g. when contextual cues improve accuracy, Palmer, 1975).

Most studies supporting the Cancellation model have used suprathreshold stimuli and have asked participants to make judgements about the intensity of an action outcome – such as the force of a tactile tap (Bays et al., 2005) or the loudness of a tone (Weiss et al., 2011). However, the Cancellation model also makes predictions about how expectations should influence the detection of stimuli presented at-threshold. An advantage of signal detection tasks is that they enable independent characterisation of an observer's objective sensitivity (represented by the statistic *d'*) to a stimulus and their general tendency to report the presence of a stimulus (often represented as the statistic *c*; Green & Swets, 1966; See Section 3.2.1.). Under the Cancellation model, processing of expected action outcomes is suppressed. As the mechanistic details of this suppression are unclear, it is possible that this could manifest as a reduction in sensitivity for outcomes congruent with executed actions. However, the Cancellation model could also predict differences in criterion *c,* e.g., if suppression is achieved by dampening the activity of expected sensory units, generating generally more conservative responding (higher *c*) for congruent action outcomes.

Different predictions are made by the OPPOSE model, which hypothesises equivalent effects of expectation on detection performance in action and non-action contexts. Outside of action, some have reported that valid expectations improve perceptual sensitivity. For example, Stein and Peelen (2015; Experiment 2d) report that when participants detect phase-scrambled target images, *d'* is increased by valid cues about object type (e.g. the written word 'car'). However, it has been queried whether such findings reflect effects of expectation or top-down attention, given that written names could imply the cued item is task-relevant as well as engendering probabilistic expectations (Summerfield & Egner, 2016). A less controversial finding in the signal detection literature is that increasing the conditional probability of a stimulus liberalises responding, generating both more hits and false alarms (Swets, Tanner, & Birdsall, 1961). One reading of these effects suggests that such response liberalisation reflects the fact that observers adopt a decision strategy that reflects stimulus probability, without any changes in perceptual processing. However, a perceptual locus of such criterion shifts was suggested in an experiment by Wyart et al. (2012). Here the authors report a task where participants detected oriented Gabors in noise, but where stimuli could be conditionally more or less probable (i.e. expected and unexpected). While signal detection measures revealed expected stimuli were associated with more liberal response criteria while leaving sensitivity (measured as *d'*) unchanged, a combination of psychophysical reverse correlation analyses (Solomon, 2002) and explicit computational modelling suggested that both hits and false-alarms reflected sensitivity to signal-like noise, and that effects of expectation were best accounted by assuming that expectations induced early sensory biases rather than influences on later decision thresholds. This finding is consistent with the idea that expectations bias sensory populations toward expected outcomes (Summerfield & de Lange, 2014).

The logic of the traditional signal detection task has also been extended by researchers in metacognition to subjective judgements about task performance (Fleming & Lau, 2014). In 'Type II' signal detection tasks participants make the same (Type I) judgements about the presence or absence of a stimulus, alongside (Type II) confidence judgements about the accuracy of their responses (See Figure 3.2). Type II tasks permit the calculation of analogous statistics that capture metacognitive sensitivity (i.e. how much subjective confidence tracks objective accuracy) and metacognitive biases (i.e. tendency to report high or low confidence). While we are not concerned with metacognition specifically in this thesis, it has previously been suggested that metacognitive biases in perceptual tasks may be a useful proxy for the vividness of a perceptual experience– a tendency to be less confident in one condition of a detection task over another may reflect weaker subjective experiences upon which to base a decision (Fleming & Lau, 2014; Peirce & Jastrow, 1884). It is therefore also possible to derive predictions about how expectations during action should influence metacognitive biases in perceptual tasks. Under the Cancellation model, expected action outcomes may have their intensity suppressed, inducing conservative metacognitive responding, whereas under the OPPOSE model participants may show liberal metacognitive responding, as has been seen for expected events outside of action contexts (Sherman, Seth, Barrett, & Kanai, 2015).

The Cancellation and OPPOSE models are contrasted in this chapter by investigating how prediction during action influences Type I and Type II detection performance. In Experiment 2, Participants executed finger tap movements (e.g. index finger movement) while observing congruent (index finger) and incongruent (middle finger) actions of an avatar hand. Observed action stimuli were presented at perceptual threshold and were backward masked, making them difficult to detect. (NB: Movements were altered from finger lifts in Experiment 1 to finger taps in Experiment 2, as piloting found observed lifting movements difficult to backwards mask – perhaps due to the fact that lifting movements lead to slight increases in visual angle - see below). Participants performed both Type I (stimulus presence) and Type II (confidence in Type I) decisions, and signal detection measures were used to compare perceptual performance as a function of action-outcome congruency. For Type I measures, the Cancellation model predicts reduced sensitivity and/or more conservative responding for congruent action outcomes, whereas the OPPOSE model predicts enhanced sensitivity and/or more liberal responding. For Type II measures, the Cancellation model may predict more conservative metacognitive responding for congruent action outcomes, whereas the OPPOSE model may predict more liberal metacognitive responding for these events.

3.2. Experiment 2

3.2.1. Method

Participants

An opportunity sample of seventy participants (58 female, mean age = 27.4 years, SD = 7.2) were recruited from Birkbeck, University of London. Participants were second year psychology students who took part as part of a research methods practical class. It should be noted that a large number of participants were recruited for the present experiment because I was also interested in examining how individual differences in any expectation effects related to individual differences in schizotypy – analyses not reported in the present thesis due to the examination of a theoretically tangential question.

Procedure and Stimuli

Participants were seated approximately 55 cm from a computer monitor inside a dimly lit cubicle, with their right hand positioned over a keypad placed in front of the body midline. Like in Experiment 1, the keypad was positioned perpendicular to the computer monitor to disrupt spatial alignment between the participant's own hand and onscreen hand, and participants were required to position their index and middle fingers above two keys. A black screen occluded view of their right hand. Perceptual and metacognitive judgements were made with the left thumb on a separate keypad. The experiment was run using the Cogent toolbox for MATLAB. Grayscale avatar hand stimuli were created in Poser 10 (Smith Micro Software) and subtended ~15° vertically and~10° of visual field**.** All stimuli were presented against a grey background on a CRT computer monitor (32 x 24 cm, 85 Hz, 21 DPI).

During the experiment participants completed a task that involved the execution of finger taps and judgements about an observed hand presented on screen. The sequence of stimulus events is shown in Figure 3.1. At the beginning of each trial participants were presented with the avatar hand with both index and middle fingers raised. Participants were instructed to perform either a right index or middle finger tap movement (freely selected), recorded by a keypress. Half of the trials were 'signal present', where executed movements triggered an immediate replacement of the initial hand image by one with a single lowered finger, creating the impression of an apparent movement in approximate synchrony with the participant's own tap. The observed movement could either be congruent (i.e. the same finger) or incongruent (i.e. the opposite finger) with the participants' own movements. Congruent and incongruent events were presented with equal probability (see Section 1.2.2). The remaining half of trials were 'signal absent', upon which no movement of the avatar hand occurred. After 17 ms, the region of avatar hand containing the index and middle fingers was backwards masked by an oval stimulus constructed from a texture of overlaid avatar fingers (see Figure 3.1). Such a mask was used given previous work which suggests that biological action stimuli are most effectively masked when masks are composed of the same elements as the target stimulus (Cutting, Moore, & Morrison, 1988). After a further 100 ms, the entire screen was replaced by visual white noise. After a random 300 – 600 ms delay, the screen was replaced by a question about a specific finger on the avatar hand (e.g. Did the INDEX finger move?). Questions probed the congruent or incongruent finger of the avatar with equal probability, though on signal present trials the unmoved finger was never probed. Participants indicated their Type I perceptual judgement with a left thumb press, and were then presented with a prompt to make a Type II judgement (‘High confidence or low confidence?’) with an additional thumb press. Neither response was speeded.

Before the main task participants completed 16 practice trials – an initial eight with executed actions and only perceptual judgements, and a further eight with executed actions and both perceptual and metacognitive judgements. To present stimuli at threshold, practice was followed by an adaptive staircase session in which participants detected masked avatar finger movements as described above, without executing their tapping actions. Half of trials were signal present (observed index or middle movement equiprobable), while the remaining half were signal absent. The magnitude of the observed finger tap on signal present trials (degrees of rotation relative to metacarpophalangeal joint) was adjusted according to a one-up-one-down staircase, with an initial rotation angle of 16° and initial step size of 4°. After each hit the observed rotation decreased by the step size (to a minimum of 1°), whereas each miss increased the observation rotation (to a maximum of 16°). At each reversal point the stepsize halved, to a minimum of 1°. The staircase was terminated after twelve reversals, and the threshold was estimated by averaging the rotation angles presented at the final six reversal points. Observed movements on signal present trials in the main experiment were displayed at each participant’s threshold value.

Participants completed at least 200 trials during the main experiment - 100 each where they executed index and middle finger taps. Responses were recorded for all trials until the experiment this criterion was reached. Breaks were taken every 40 trials, at which point participants received feedback about the distribution of their responses. The experiment was terminated once participants had completed 100 of each movement type. Trial types were randomised across the experiment.

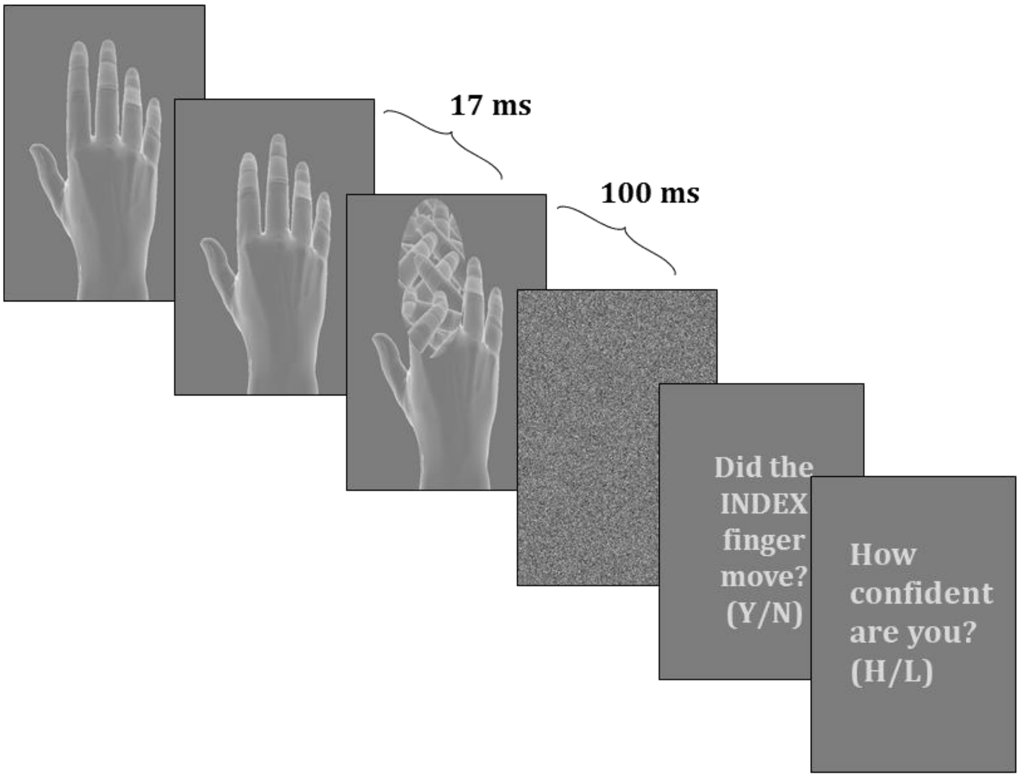

Figure 3.1: Timecourse of presentation and stimuli in Experiment 2.

Measures of Type I and Type II sensitivity and bias

In Type I signal detection tasks stimuli can be objectively present or absent, and participants can respond either ‘present’ or ‘absent’. In Type II signal detection tasks, a Type I response can be objectively correct or incorrect and participants can report ‘high confidence’ or ‘low confidence’. Figure 3.2 illustrates the kinds of responses that are possible in such a task.

Type I Signal Detection

	Signal objectively present	Signal objectively absent
Respond ‘Present’	**Type I Hit**	**Type I False Alarm**
Respond ‘Absent’	**Type I Miss**	**Type I Correct rejection**

Type II Signal Detection

	Type I decision correct	Type I decision incorrect
Confidence ‘High’	**Type II Hit**	**Type II False Alarm**
Confidence ‘Low’	**Type II Miss**	**Type II Correct rejection**

Figure 3.2: Possible response types in a Type I (left) and Type II (right) signal detection task.

For our analyses we used Type I d' as a measure of perceptual sensitivity, calculated as d'= z(Type I hit rate) – z(Type I false alarm rate), and Type I c as a measure of response bias, calculated as c = - 0.5(z(Type I hit rate) + z(Type I false alarm rate)]. High Type I d' values indicate greater perceptual sensitivity, while high c values indicate more conservative responding (i.e. fewer ‘stimulus present’ responses).

It is possible to derive an equivalent measure of Type II d' as an index of metacognitive sensitivity (Kunimoto, Miller, & Pashler, 2001), though it has been argued that such a method depends on the untenable assumption that correct and incorrect judgements generate normal distributions of evidence along an internal decision axis (Evans & Azzopardi, 2007). Maniscalco and Lau (2012) introduced an alternative measure of Type II sensitivity – meta-d – which is estimated on the basis of Type I and Type II responses and does not depend on these assumptions. This measure was therefore used as an index of metacognitive sensitivity, calculated using openly available code (http://www.columbia.edu/~bsm2105/type2sdt/) as an index of metacognitive sensitivity. Metacognitive bias was inferred from Type II *c,* where *c* = - 0.5(z(Type II hit rate) + z(Type II false alarm rate)].

3.2.2. Results

Participants were successful in producing roughly equivalent numbers of index and middle finger actions (mean = 51.54% Index actions, SD = 4.6%). Paired sample t-tests were used to compare Type I sensitivity (*d')* and bias (*c*) values across congruency conditions. *d'* was found to be higher on congruent trials (mean = 1.70, SEM = .108) than on incongruent trials (mean = 1.57, SEM =.097; *t*(69) =2.832, *p*=.006, *d*=.145). Analyses also revealed that observers were more liberal in reporting observed finger movement on congruent trials (mean = .019, SEM = .085) than on incongruent trials (mean = .195, SEM = .070; *t*(69) = 3.574, *p*=.001, *d*=.261). These patterns are illustrated in Figure 3.3.

The same analyses were conducted to compare metacognitive sensitivity (meta-*d*) and bias (Type II *c*) across congruency conditions. No significant differences were found in meta-*d* – *t*(69) = 1.406, *p*=.164 - or Type II *c* - *t*(69) = .863, *p*=.391.

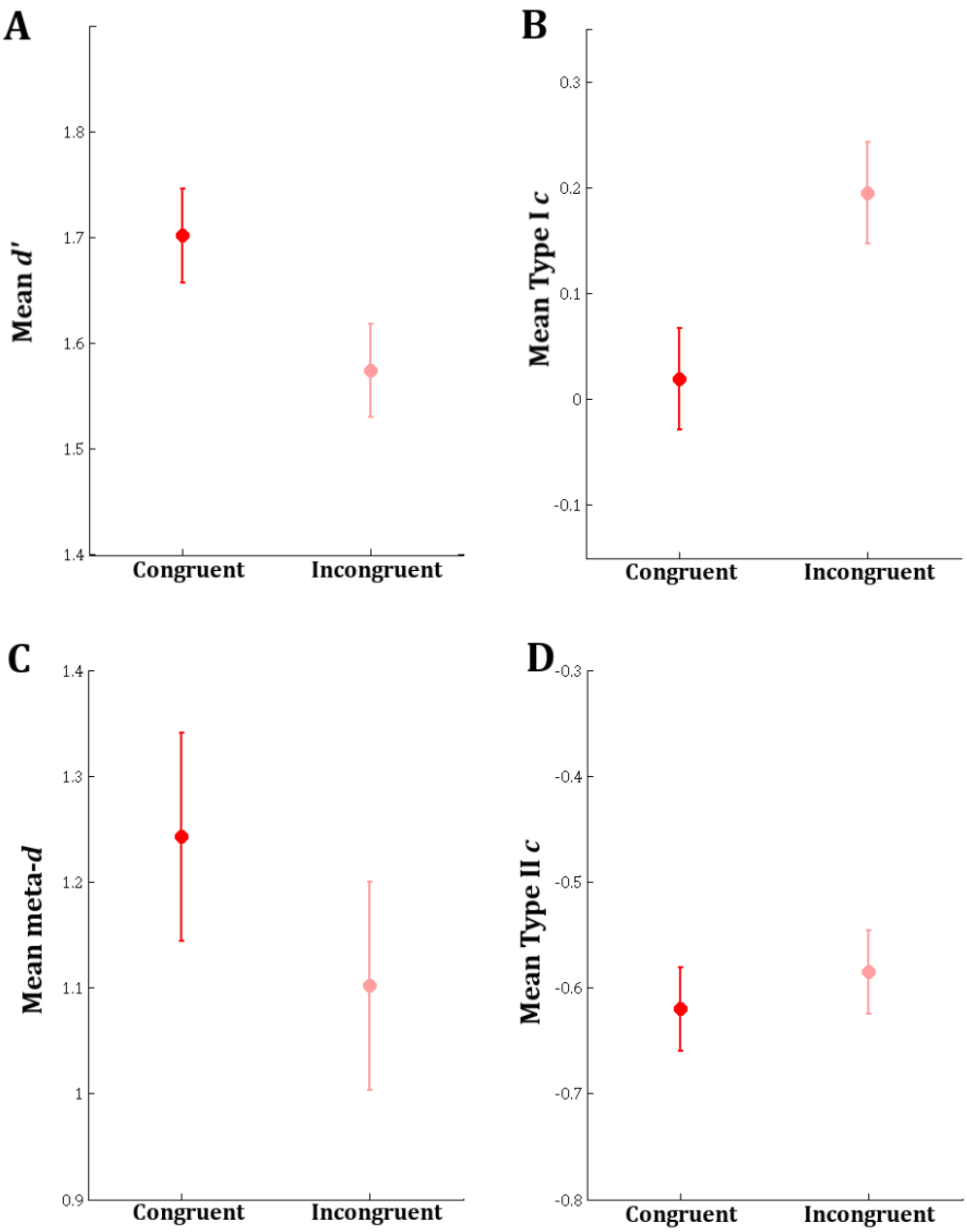

Figure 3.3: Mean *d'*, Type I *c,* meta-*d* and Type II *c* (panels a-d) values across congruency conditions in Experiment 2. Error bars display 95% within-subject confidence intervals (Cousineau, 2005).

3.2.3. Discussion

Experiment 2 found no significant effect of congruency on metacognitive sensitivity or bias. While it is difficult to make inferences from null results, these findings may suggest that predictions during action do not affect perceptual metacognition. Alternatively, influences on these metacognitive measures may exist, but not be detectable given differences in Type I performance. Previous work reporting metacognitive biases induced by instructions about stimulus probabilities ensured Type I performance across conditions was matched (Sherman et al., 2015), and to the extent that Type I performance influences Type II performance (Fleming & Lau, 2014), differences in basic task performance as reported herein could mask potential influences of motor prediction on metacognition.

Experiment 2 found evidence that participants were more sensitive to congruent action outcomes, and less conservative in reporting movement of congruent fingers. These results are inconsistent with the Cancellation model, which hypothesises that predictions act to suppress the activity of expected sensory units during action (Bays & Wolpert, 2007). Such suppression should manifest as either a reduced sensitivity to congruent action outcomes, or more conservative responding about their presence. However, the results reported in this chapter are consistent with the OPPOSE model, which hypothesises the expectations during action operate like predictions in other contexts - where they increase hit rates (Stein & Peelen, 2015) and false alarm rates (Wyart et al., 2012).

Interestingly, our finding that *d'* was higher for congruent action outcomes appears to contrast with findings from the only other experiment that has examined signal sensitivity to action-congruent and –incongruent outcomes. Cardoso-Leite and colleagues (2010) found that low-contrast Gabor patches with orientations congruent with a learned sensorimotor contingency were associated with a reduced *d'*. However, there are differences between the paradigms deployed in these two cases which are likely to explain this discrepancy and which led to speculation that the Cardoso-Leite et al. (2010) findings may not generalise. For instance, the Cardoso-Leite et al. (2010) study utilised a pre-cue (oriented lines) that was similar in nature to the action effects, which could have engendered a form of repetition suppression on congruent trials. Additionally, average *d'* was higher in the experiment reported by Cardoso-Leite et al. (mean = 2.56) than in the experiments reported here (mean = 1.64, SD = .84). It is therefore possible that in Cardoso-Leite et al., (2010) action outcomes were not presented at detection threshold, and that incongruent outcomes benefitted from prediction-error-dependent processes on some trials (see Experiment 1).

Our findings suggest that sensory prediction during action has a similar influence on perceptual processing as other kinds of expectation, and are consistent with the idea that predictions additively bias activity in expected sensory units (Summerfield & de Lange, 2014). However, the specific pattern of these results differs from some previous studies in sensory cognition. For example, Wyart et al. (2012) find that expected outcomes are associated with less conservative responding (lower *c*) but do not find any effects on *d'*. The authors suggest this pattern can be explained if predictions bias sensory populations in an additive fashion – equivalently increasing hits on signal present trials and false alarms on signal absent trials. In contrast, our experiments find the same effect of congruency on criterion (lower *c* values) alongside higher *d'* values for congruent events.

One possible explanation for this difference is that prediction during action increases sensory gain in a multiplicative fashion – increasing the activation in sensory units proportionally with signal strength – in a similar way to that which has been reported for attention (Reynolds & Heeger, 2009; see Figure 3.4). Alternatively, it is possible for an additive mechanism to generate influences on both *c* and *d'* if an observer's threshold for reporting detection is non-stationary across trials (Benjamin, Diaz, & Wee, 2009). A schematic example is illustrated in Figure 3.4.b. If on a given trial a threshold for reporting detection is high, additive biases will have larger effects on hits – increasing the likelihood that present signals cross threshold - but will not increase false alarms on signal absent trials. In contrast, if a threshold for reporting detection is low, additive biases will increase the likelihood that signal-like noise passes decision threshold, while truly present signals will be detected regardless. Importantly, under this scheme the effect of additive biases on *d'* and *c* will depend on the range of detection thresholds an observer adopts across trials. In particular, when observers adopt a generally conservative threshold (as was the case across congruency conditions in Experiment 3, mean *c* = .107), additive biases will have a larger effect on hit rates than false alarm rates – increasing both *d'* and *c.*

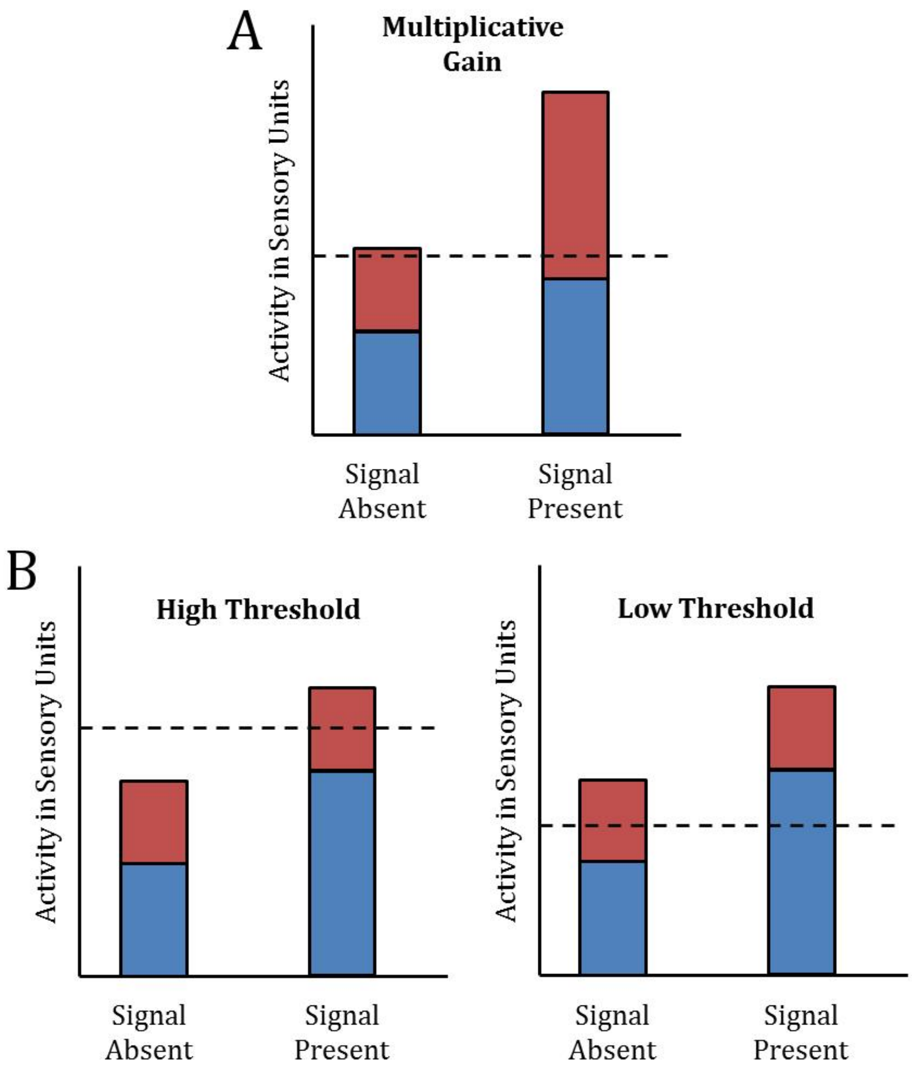

Figure 3.4: Schematics illustrating possible effects of predictive biasing on signal detection performance. Blue bars indicate the signal-like activity (signal + noise on present trials, pure noise on absent trials), while red bars indicate increases in activation resulting from prediction during action. Dashed lines indicate a hypothetical threshold on a given set of trials which activity has to reach for a participant to respond 'signal present'. A: If prediction increases sensory gain in a multiplicative fashion, signal and signal-like noise will be more likely to reach detection threshold, but the effect will be more pronounced for true signals. B: If prediction biases sensory units in an additive fashion, signal and signal-like noise will be equivalently more likely to reach detection threshold – though whether this has a larger effect on signal present or signal absent trials depends on where detection threshold is placed (see Section 3.2.3)

Traditional signal detection theoretic measures are not suited to evaluating these possibilities, as by definition it is not possible to calculate measures of sensitivity and bias separately for signal-present and –absent trials. However, it would be possible to investigate these questions using the psychophysical reverse correlation approach described by Wyart et al. (2012). Here, the authors found that expectations increased the sensitivity of false alarms to signal-like noise, and that effects of expectation were best modelled as an additive increase in sensory activity. If prediction during action also has an additive effect on activity in expected sensory units, comparable effects should be obtained in action tasks. Alternatively, if prediction during action influences gain in a multiplicative fashion (Reynolds & Heeger, 2009), effects of prediction on energy sensitivity should be larger for stronger (i.e. present) signals (Wyart et al., 2012). While external signal-like noise was not systematically varied in this experiment – precluding a reverse correlation analysis – this remains an important question for future work.

3.3. Chapter Summary

The Cancellation and OPPOSE models make contrasting predictions about how prediction during action should influence performance on detection tasks. Experiments 2 suggests that participants are more sensitive to and more liberal in reporting the presence of congruent action outcomes – in line with the predictions of the OPPOSE model.

Chapter 4 – Influences of action on visual brain activity

The Cancellation model suggests that sensory processing of expected action outcomes is suppressed at the input stage. This contrasts with the OPPOSE model, which suggests that expectations generated on the basis of action bias sensory units in line with expected outcomes. This chapter investigates these two contrasting predictions in an fMRI experiment. In Experiment 3 participants completed a task that involved the execution of finger movements (index and little finger abductions) while observing congruent (i.e. expected) and incongruent (i.e. unexpected) actions performed by an on-screen avatar hand. Through an additional orthogonal manipulation of top-down attention, the observed actions could be relevant or irrelevant to the perceptual task. Multivariate pattern classifiers were built to decode the identity of observed stimuli on congruent and incongruent trials, and classifier performance was used a proxy for the quality of underlying sensory representations. This permitted analysis of how visual processing is altered as a function of action-outcome congruency. Results revealed significant decoding of stimulus identity in clusters across the visual cortical hierarchy. Importantly, results revealed superior decoding of congruent action outcomes relative to incongruent outcomes in occipital and occipitotemporal brain regions. Additional analyses of stimulus-specific univariate activity revealed that these decoding enhancements were accompanied in reduced activity on congruent trials in voxels tuned *away from* the presented stimulus. These results are difficult to reconcile with the Cancellation model, but are consistent with the idea that expectations during action bias sensory populations towards predicted outcomes.

4.1. Introduction

The Cancellation model proposes that the expected consequences of action are ‘subtracted’ from the sensory input (Bays & Wolpert, 2007; Blakemore et al., 2000). Support for this hypothesis has been drawn from fMRI studies that have reported that events predictable on the basis of action are associated with reduced BOLD responses in sensory brain areas. In touch, Blakemore et al. (1998) report that activity in the secondary somatosensory cortex (SII) elicited by brushing sensations applied to the hand are reduced when participants operate the brush (i.e. when stimuli are predictable) compared to when the brush is operated by the experimenter. Similarly, a number of studies have reported that force taps which are delivered in synchrony with a participant’s own tapping movement (i.e. in line with temporal predictions) elicit reduced responses in SII compared to force taps delivered at subsecond delays (Shergill et al., 2013, 2014).

Similar results are obtained in visual areas throughout the cortical hierarchy. For example, in an experiment reported by Stanley and Miall (2007) participants viewed an avatar hand performing sequential hand opening and closing movements while producing a congruent action sequence (hand opening and closing) or an incongruent action (wrist rotation). The authors found that congruent visual outcomes were associated with reduced activity in the primary visual cortex (V1). Comparable effects are obtained in superior temporal sulcus (STS), a region implicated in the processing of biological motion and observed actions (Allison, Puce, & McCarthy, 2000). Leube et al. (2003) report an experiment where participants slowly and continuously opened their hand (0.5 Hz) while observing visual feedback of their own movements captured on camera. The authors found that imposing progressively greater delays on the video feedback (0 -200 ms) led to linear increases in BOLD responses in STS – consistent with the idea that the more predictable (i.e. more simultaneous) action outcomes are cancelled. Kontaris et al. (2009) report a similar experiment where participants executed a sequence of hand gestures based on auditory cues while observing video feedback from their current actions, or incongruent feedback recorded in a previous block. This study found that BOLD responses in STS are reduced when actions and outcomes are congruent. In all cases, reduced univariate activity is interpreted as consistent with the assumption that predicted sensory signals have been suppressed, i.e., as consistent with the Cancellation model (Bays & Wolpert, 2007; Blakemore et al., 2000).

Outside the action literature, it is commonly reported that predicted sensory events are also associated with reduced BOLD responses (Summerfield & Egner, 2009). For example, visual dots that are presented along an apparent motion trajectory (Alink et al., 2010), and geometric shapes (den Ouden, Friston, Daw, McIntosh, & Stephan, 2009) and visual objects (Utzerath, St John-Saaltink, Buitelaar, & de Lange, 2017) predictable on the basis of a prior stimulus all elicit reduced activity in early visual brain areas compared to unpredicted stimuli. However, these reductions in univariate signal for expected events are not always thought to imply reduced processing by expected sensory populations. Indeed, as described previously in this thesis (See Sections 1.2.4. 2.1, 3.1) an influential idea in sensory cognition is that expectations bias sensory representations toward predicted outcomes (Summerfield & de Lange, 2014), which could be implemented by sharpening the tuning curves of sensory neurons sensitive to expected features (Summerfield & Tsetsos, 2015). This ‘sharpening’ in turn reduces the overall activity in sensory populations, leading to reduced BOLD responses relative to unexpected events.

A powerful way to investigate this ‘sharpening’ hypothesis is to investigate how expectations influence multivariate measures of BOLD activity (Haynes & Rees, 2006) given that sharpened population responses should be easier to decode using pattern classification techniques. Kok et al. (2012) investigated this possibility in an fMRI study where participants observed visual grating stimuli, the orientation of which was predicted by an auditory cue. In line with previous studies (e.g. Alink et al., 2010), the authors found that BOLD activity in early visual areas (V1-V3) was attenuated when expectations were valid relative to when they were invalid. Kok et al. (2012) then built separate multivariate classifiers that were trained and tested on their ability to decode the orientation of observed gratings from patterns of voxel activity for expected and unexpected trials. This analysis revealed superior decoding in V1 of stimulus identity from trials when expectations were valid relative to invalid, in line with the idea that expectations sharpen the responses of sensory populations.

To date similar multivariate techniques have not been used in fMRI studies investigating sensory prediction during action, and divergent predictions are made by the Cancellation and OPPOSE models about the pattern that should be obtained – which are examined in Experiment 3. Participants performed a task in an MRI scanner where they executed finger abduction movements (index or little finger) and observed congruent and incongruent movements on an avatar hand. Observed movements also revealed a small coloured dot (red, blue). (NB: The effectors and movements used in Experiment 3 were changed relative to Experiments 1 and 2 to maximise the perceptual differences between observed action stimuli, as we had no a priori predictions about which stimulus dimensions are relevant to prediction mechanisms in action). On each trial, participants made a judgement about the stimulus display. On task-relevant blocks these judgements concerned observed actions, while on task-irrelevant blocks judgements concerned the coloured dot. BOLD activity elicited by observed stimuli was modelled and analysed using multivariate pattern classification techniques, permitting an analysis of how information about observed actions varied as a function of congruency and task-relevance. Under the Cancellation account, the suppression of neural activity associated with expected action outcomes should reduce the signal available to a pattern classifier – leading to reduced decoding accuracy for congruent relative to incongruent events. Conversely, the OPPOSE model suggests, in line with models of sensory prediction outside of action (Summerfield & de Lange, 2014), that predictions bias sensory representations toward expected outcomes – leading to sharper population responses for congruent events, which should in turn be more readily decoded from BOLD activity.

4.2. Experiment 3

4.2.1. Method

Participants

20 healthy right-handed individuals (14 female, mean age = 23.1 years, SD = 2.4) were recruited from University College London and paid a small honorarium. An additional three participants were tested, but were excluded due to excessive movement and discomfort during scanning (two) or as a result of a technical error during image acquisition (one).

Experimental Task

Stimuli were displayed against a black background on a rear-projection screen using a JVC DLA-SX21 projector (26 x 19.5 cm, 60 Hz). Observed hand stimuli were generated in Poser 10, and consisted of a gender-neutral right hand viewed from a canonical first-person perspective (height ~13 °, width~9° see Figure 4.1). Participants lay in the scanner with their right and left hands on separate MR-compatible button boxes. The right hand box was positioned across the midline of the participant's body, such that the index finger was above the little finger on the dorsal-ventral axis. Participants depressed two buttons on the right hand box with their index and little finger except when executing movements (see below). The left hand button box was positioned below the right hand box on their left leg, and participants placed their left thumb between two response keys.

Each trial began with the presentation of a white fixation cross, which remained present throughout stimulus presentation. After 750 ms, a neutral hand image was presented behind the fixation cross. On Movement trials, this neutral hand image was also accompanied by an outline shape (square or circle) indicating which action (index or little abduction) the participant was required to perform. On Movement trials, this display remained onscreen until participants executed the appropriate action (releasing the correct key), at which point the neutral hand image was immediately replaced by an image of the avatar hand abducting either its index or little finger. This sequence created apparent motion of the observed finger that could be congruent or incongruent with the participant's own movement, and that was apparently synchronous with their movement. The apparent movement of the avatar hand also revealed a coloured dot (red or blue) in the previous fingertip location (see Figure 4.1). We also included No Movement trials, where an imperative shape cue did not appear and the apparent motion sequence of the onscreen hand occurred after a fixed delay 438 ms (matched to the average action execution RT in a pilot experiment; a fixed delay was implemented such that the onset of movement was approximately comparable in predictability relative to trials on which participants moved themselves, where they could predict that the movement would be present when they moved their finger). These trials were designed to allow us to define regions of interest (ROIs) based on responsivity to the observation of these stimulus events. After 500 ms, the hand image was removed and the screen was blanked for 1000 ms.

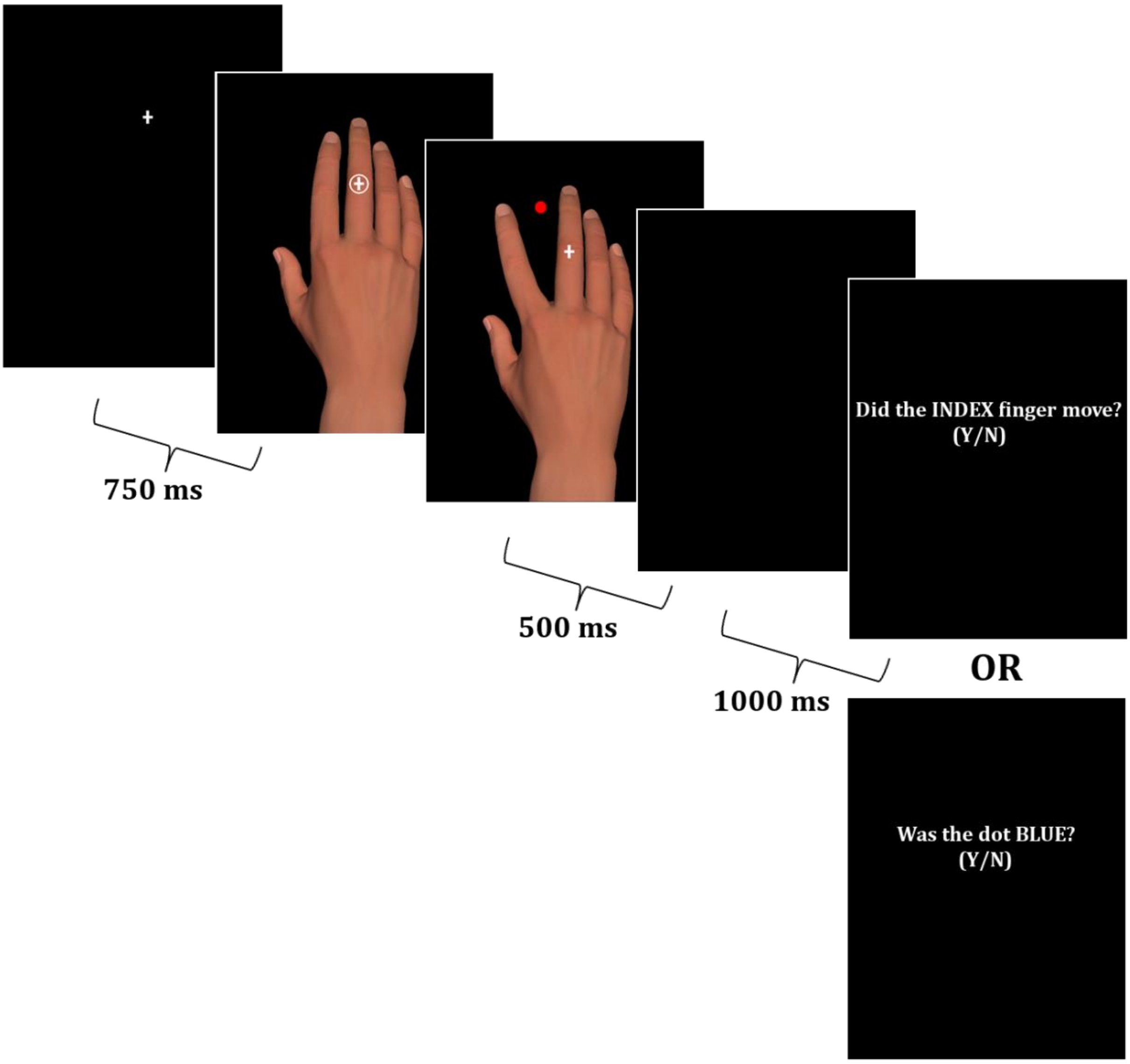

Figure 4.1: Timecourse of presentation and stimuli used in Experiment 3. In task-relevant blocks participants made judgements about observed finger movements, while in task-irrelevant blocks judgements concerned the colour of the dot revealed after the avatar finger abducted.

Participants completed one of two tasks. On task-relevant trials, participants were required to make a judgement about the identity of the observed finger abduction (e.g. did the index finger move?). On task-irrelevant trials, participants were required to make a judgement about the colour of the dot revealed by finger movement (e.g. was the dot red?). On each trial the question was presented for 1500 ms, within which time participants were required to indicate their response via a keypress with their left thumb. The next trial began after a jittered ITI of 2-6 seconds.

The experiment was conducted in eight scanning sessions. Each session comprised 48 trials. On two thirds of these trials participants executed index or little abductions with equal probability, and subsequently observed either Congruent or Incongruent action outcomes with equal probability (16 each). The remaining third of trials were No Movement trials (16), where participants observed index or little abduction movements with equal probability. The task-relevance manipulation was blocked within each scanning session, such that half of the blocks involved judgements where observed actions were task-relevant and half involved judgements where observed actions were task-irrelevant. The task participants performed alternated across blocks, with the order counterbalanced across participants. At the beginning of each task block participants were reminded of the task they were performing, as well as the mapping between imperative shape cues and executed actions. This mapping was counterbalanced across participants, and was also reversed halfway through the experiment (i.e. the beginning of the fifth scanning session) to unconfound cue-outcome congruency and action-outcome congruency over the experiment.

Image Acquisition

Images were acquired using a 3T Trio MRI scanner (Siemens, Erlangen, Germany) using a 32-channel head coil. Functional images were acquired using an EPI sequence (ascending slice acquisition, TR = 3.36 s, TE = 3.29 s, 48 slices, voxel size: 3 x 3 x 3 mm). Structural images were acquired using an MP-RAGE sequence (voxel size: 1 x 1 x 1 mm).

fMRI Data preprocessing

Preprocessing was conducted in SPM12. The first six volumes of each participant's data in each scanning run were discarded to allow for T1 equilibration. All functional images were spatially realigned to the mean image (yielding head movement parameters) and temporally-realigned to the 24th (middle) slice. Functional images were then coregistered to the participant's structural scan and segmented to estimate forward and inverse deformation fields to transform data from participants' native space into Montreal Neurological Institute space, and vice versa.

4.2.2. Analysis and Results

4.2.2.1. Multivariate Pattern Analysis (MVPA)

MVPA analyses were conducted for each participant using The Decoding Toolbox (Hebart, Görgen, & Haynes, 2014). All analyses used a whole-brain searchlight approach. In each analysis, a linear support vector machine (SVM) was trained to discriminate which stimulus was observed on a given trial from patterns of BOLD activity across voxels. The initial step in each analysis was the specification of a GLM in SPM12 including a separate regressor for each stimulus type (e.g. Observed Index) in each experimental condition (e.g. Congruent trials) in each scanning run. All regressors were modelled to the onset of the observed stimulus, movement parameters were included as nuisance regressors, and all regressors in the model were convolved with the canonical haemodynamic response function. This GLM generated eight beta images (one for each scanning run) for each stimulus type (Index or Little) in each experimental condition that were used for subsequent decoding analyses.

Separate SVMs were trained and tested on the 16 beta images (eight index and little) in each experimental condition. This was conducted using a leave-one-out cross validation procedure where for each decoding step 14 images from seven scanning runs were used to estimate a linear discriminant function separating Observed Index and Observed Little finger movements, which was then applied to the remaining two beta images to classify them as either Observed Index or Observed Little. Each SVM analysis comprised eight decoding steps, where a different pair of beta images from each scanning run were reserved for classifier testing, and the SVM's accuracy was calculated as the proportion of correctly classified images across all decoding steps.

A whole-brain 'searchlight' approach was used (Kriegeskorte, Goebel, & Bandettini, 2006), which involved conducting a separate SVM for each voxel in the brain using the beta values falling within a searchlight radius of 3 voxels, and assigning the SVM's accuracy to the voxel on which the searchlight was centred. This procedure was used to create decoding maps in participant's native space indicating each voxel's decoding accuracy relative to chance level (50%, i.e., decoding accuracy of 60% is treated as 10%). To allow comparison across participants, these decoding maps were normalised into MNI space using the forward deformation fields estimated in preprocessing and smoothed using a 4mm FWHM Gaussian kernel in SPM12.

4.2.2.2. MVPA Analyses collapsed across task-relevance

To maximise sensitivity, MVPA analyses were initially conducted collapsing across the task-relevance factor. As such whole-brain searchlight analyses were conducted separately for Congruent, Incongruent and No Movement trials, yielding three decoding maps per participant that were normalised and smoothed (described above). Decoding maps from the No Movement condition were used to define regions of interest (ROIs) to compare effects of congruency. This was achieved by subjecting the No Movement decoding maps from each participant to a one-sample t-test in SPM12, and using cluster-wise inference to identify contiguous voxels (no minimum cluster size) where decoding accuracy was significantly above chance at the group level (Loose, Wisniewski, Rusconi, Goschke, & Haynes, 2017). This contrast was restricted to occipital and temporal brain regions using the SPM12 atlas and used a primary cluster-defining threshold of $p<.001$ uncorrected with cluster-wise family-wise error (FWE) of $p<.05$ – a combination which recent authors have suggested appropriately controls false positive rates (Eklund, Nichols, & Knutsson, 2016). This analysis revealed three clusters in bilateral occipital cortex (BO, 1503 voxels), left occipital cortex (LO, 545 voxels) and right occipitotemporal cortex (ROT, 339 voxels) which were carried forward as ROIs (see Figure 4.2).

To investigate the effect of sensory prediction on decoding accuracy, we extracted and averaged the decoding accuracies within each ROI separately for Congruent and Incongruent trials. These mean accuracies were then subjected to a 3 x 2 repeated measures ANOVA, with factors of ROI (BO, LO, ROT) and Congruency (Congruent, Incongruent). This revealed a significant main effect of Congruency – $F(1,19) = 4.781$, $p=.041$, $\eta_p^2 = .201$ and no interaction between congruency and region of interest ($p=.760$); see Figure 4.2).

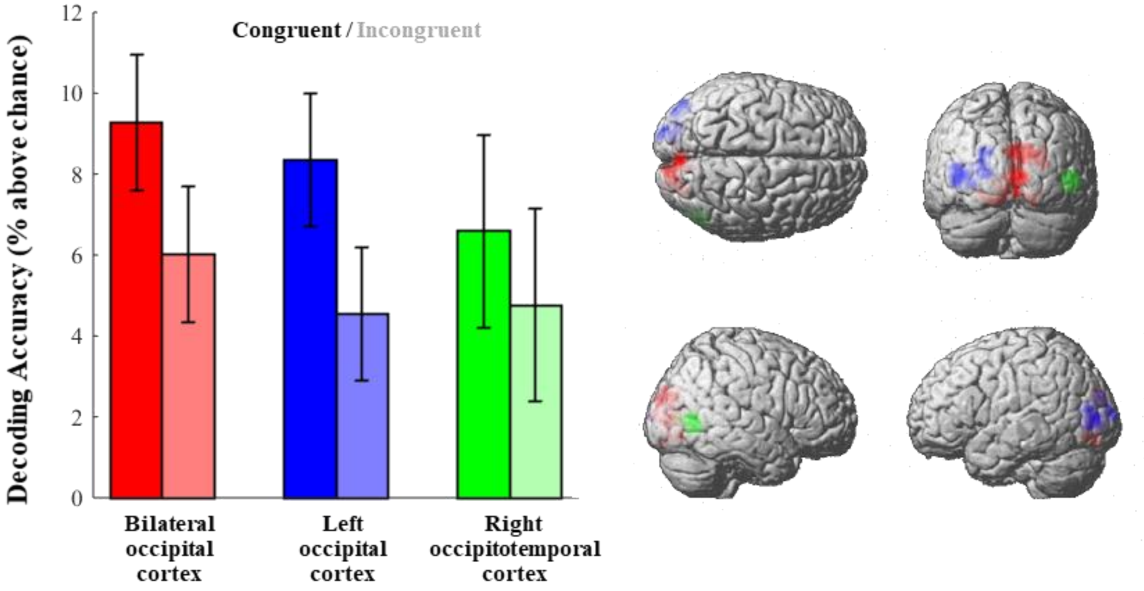

Figure 4.2: Right: Occipital and occipitotemporal clusters containing above-chance information about observed stimulus identity (index vs little) on No Movement trials ($p<.001$ uncorrected, cluster-wise FWE $p<.05$). Left: Mean stimulus decoding accuracies within each ROI on Congruent (saturated) and Incongruent (faded) trials. Error bars represents 95% within-participant confidence intervals.

4.2.2.3. MVPA analyses examining interactions with task-relevance

To investigate whether the effects of Congruency identified above interacted with the task-relevance manipulation, additional whole-brain searchlights were conducted separately for each combination of Congruency and Task-Relevance. This procedure was identical to that described above, though segregating stimulus events by relevance halved the number of stimulus events used to model beta images for decoding. Mean decoding accuracies for each participant were conducted for each condition in each cluster and these values were analysed using a 3 x 2 x 2 repeated measures ANOVA with factors of ROI (BO, LO, ROT), Congruency (Congruent, Incongruent) and Task-Relevance (Relevant, Irrelevant). This analysis revealed a three-way interaction between ROI, Congruency and Task-Relevance ($F(2,38) = 4.725$, $p=.015$, $\eta_p^2=.199$). As shown in Figure 4.3, this three-way interaction reflected the fact that while comparable congruency effects were obtained in the BO and LO ROIs regardless of task-relevance, a significant interaction between Congruency and Task-Relevance was obtained in the ROT cluster ($F(1,19)= 9.678$, $p=.006$, $\eta_p^2=.337$). This two-way interaction reflected a trend toward higher decoding accuracies in the Congruent than Incongruent condition when stimuli were task-irrelevant ($t(19) = 1.930$, $p=.069$) and the opposite pattern when stimuli were task-relevant ($t(19) = 1.401$, $p=.177$).

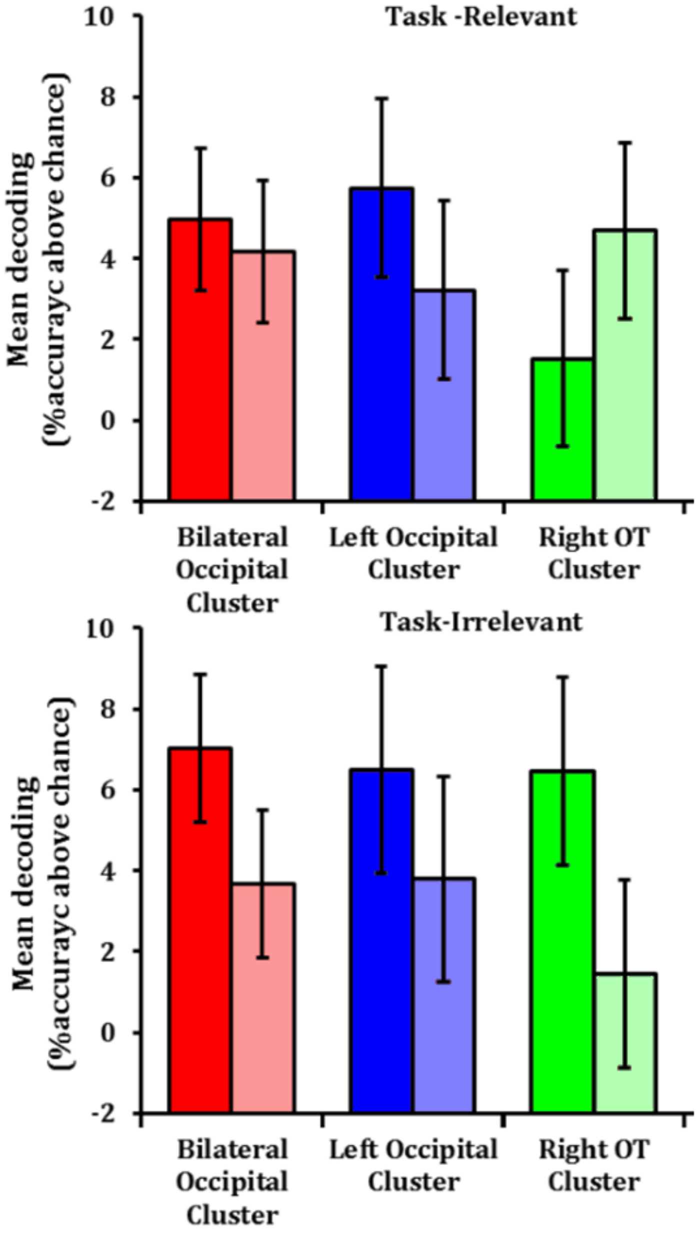

Figure 4.3: Mean stimulus decoding accuracies within (and collapsed across) each ROI on Congruent (saturated) and Incongruent (faded) trials, separately for task-relevant (top panel) and task-irrelevant trials (bottom panel). Error bars represents 95% within-participant confidence intervals.

4.2.2.4. Univariate BOLD amplitude analysis

While Experiment 3 was primarily concerned with the divergent predictions made by the Cancellation and OPPOSE models with respect to how expectations during action should alter multivariate measures of sensory brain activity, additional analyses were conducted to investigate how action-outcome congruency influenced univariate activity. Previous reports of expectation sharpening have suggested that the signal-to-noise ratio of sensory populations is improved by suppressing the activity in sensory units tuned *away from* expected action outcomes (Kok et al., 2012).

We investigated whether a similar effect was obtained in Experiment 3 by assessing how congruency affected stimulus-specific patterns of univariate activity within each of the ROIs used for MVPA analyses. Using the same unnormalised, unsmoothed images used for multivariate decoding, we conducted a t-contrast in SPM12 for each participant comparing activity for observed index finger stimuli and observed little finger stimuli across all conditions. This contrast yielded a t-map for each participant where positive and negative values reflected a voxel's preference for either index or little fingers, respectively.

After assigning a preferred stimulus to each voxel, we extracted univariate BOLD signal (beta values) from each voxel separately for congruent and incongruent trials as a function of whether the stimulus was the preferred or non-preferred stimulus for a given voxel. For example, if a voxel was classified as 'index preferring', the univariate signal on congruent trials where an index finger was presented was congruent-tuned to, whereas signal on the same trials was congruent-tuned away for voxels classified as 'little preferring'. Univariate BOLD signal was extracted from each voxel in each of the clusters used for decoding and analysed with a 3 x 2 x 2 repeated measures ANOVA with factors of ROI (BO, LO, ROT), Congruency (Congruent, Incongruent) and Stimulus-Preference (tuned to stimulus, tuned away from stimulus). This analysis revealed a significant interaction between congruency and stimulus preference ($F(1,19)= 9.306$, $p=.007$, $\eta_p^2=.329$) which did not vary across cluster ($p=.168$). This interaction reflected weaker activity on congruent relative to incongruent trials in voxels tuned away from the current stimulus ($t(19) = 2.214$, $p=.039$), with no congruency effect in voxels tuned towards it ($p =.286$: see Fig 4.4). Therefore, like previous reports of expectation sharpening, this pattern suggests that prediction during action suppresses activity in voxels that are tuned away from currently expected outcomes.

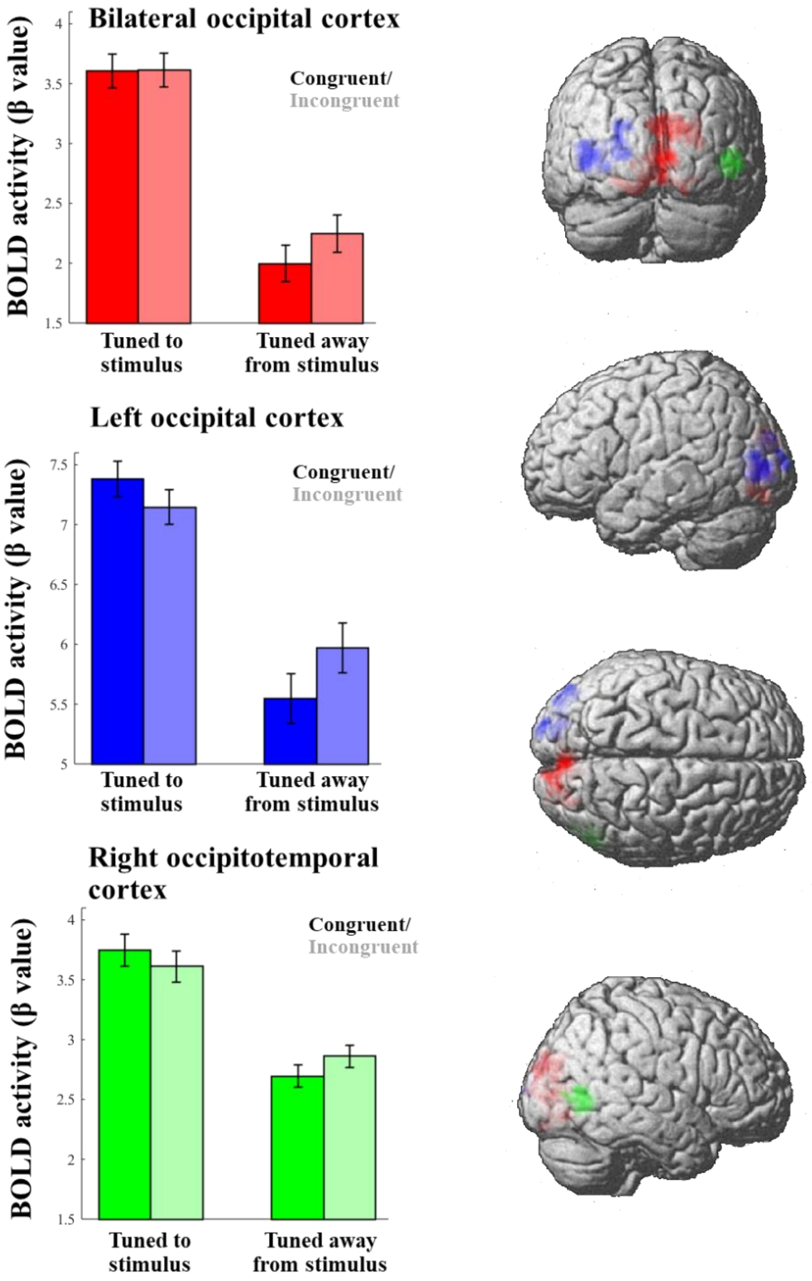

Figure 4.4: BOLD activity across regions of interest in the visual brain. Activity is lower for congruent relative to incongruent trials only in voxels that are tuned away from the observed stimulus (e.g., those voxels tuned towards index fingers on trials where little finger movements are presented). Error bars represent 95% within-participant confidence intervals.

4.2.3. Discussion

Experiment 3 revealed that congruent action outcomes were more readily decoded from BOLD activity than incongruent action outcomes across several visual brain regions in the occipital and occipitotemporal cortices. These results suggest that expectations during action can facilitate visual processing, as improvements in stimulus decoding accuracy are typically taken to reflect improvements in the quality of underlying sensory representation (Jehee, Brady, & Tong, 2011; Kok et al., 2012; Ritchie, Kaplan, & Klein, 2017). Consistent with these enhancements, analyses of stimulus-specific univariate activity revealed lower BOLD signal on congruent trials only in those units tuned *away from* currently expected action outcomes, in line with previous studies showing expectation sharpening (Kok et al., 2012).

This pattern is difficult to reconcile with the Cancellation model, which hypothesises that activity in expected sensory units is suppressed at an early processing stage (Bays & Wolpert, 2007). However, it is consistent with the OPPOSE model which hypothesises that predictions generated during action operate comparably to prediction mechanisms described in non-action contexts, and where they are thought to bias activity in expected sensory units (Kok et al., 2012; Summerfield & de Lange, 2014).

Relative improvements in decoding accuracy for congruent action outcomes were found in clusters that contained above-chance information about stimulus identity on No Movement trials. However, it is important to recognise that the ability to decode a stimulus provides little information about the precise stimulus feature supporting pattern classification (Ritchie, Kaplan & Klein, 2017), and indeed the clusters used in Experiment 3 likely span multiple anatomically-distinct regions that may contribute to different elements of visual processing. While previous studies looking at top-down influences on multivariate measures of visual activity have used more artificial stimuli that differ in a simple feature (e.g. grating stimuli with different orientations – Jehee et al., 2011; Kok et al., 2012), the observed stimuli used in Experiment 3 differ in both lower-level (e.g. retinotopic location of event) and higher-level features (e.g. direction of apparent motion, configural action identity). An interesting consideration therefore concerns the level in the cortical hierarchy at which sensory predictions operate.

It has previously been suggested that expectations predominantly affect processing in V1. In an fMRI study Smith and Muckli (2010) found that when participants viewed visual scenes containing empty regions, it was possible to decode the identity of the visual scene from V1 voxels whose receptive fields centred on the empty region – interpreted as evidence for top-down contextual predictions. Given that such decoding was not possible for comparable voxels in V2, the authors suggested that V1 may be a privileged site for effects of expectation on sensory processing. However, while Kok et al. (2012) similarly find enhanced decoding for expected gratings only in V1 (and not V2-3), they instead suggest this may reflect the fact that influences of expectation are strongest in regions sensitive to the predicted feature. The results of Experiment 3 revealed enhanced decoding for congruent action outcomes across clusters that likely include a number of visual regions, and while it is difficult to delineate these anatomical regions in the absence of (e.g.) retinotopic mapping, this may suggest that predictions influence processing at different hierarchical levels – in line with extant predictive coding models (Friston, 2005; Rao & Ballard, 1999).

An interesting feature of these data are the patterns obtained in the right occipitotemporal ROI. This region closely corresponds to previous reports of the extrastriate body area (EBA), a region which has been implicated by numerous neuroimaging studies in the visual analysis of bodies and actions (Lingnau & Downing, 2015). A handful of studies have reported that executing (Astafiev, Stanley, Shulman, & Corbetta, 2004) or preparing (Kühn, Keizer, Rombouts, & Hommel, 2011) hand actions in the absence of visual stimulation elicits activation in EBA, which has been interpreted as reflecting motor influences on visual processing. However, it has been controversial whether these 'action-related' regions are in fact overlapping with those involved in the perception of body parts (Peelen & Downing, 2005). In fact separate experiments have found that activity related to observed actions in EBA is not modulated by action-outcome congruency (Kontaris et al., 2009), suggesting that these may not be visual representations that are activated during action. However, no previous work has investigated the effect of sensorimotor predictions on multivariate brain activity and the results of Experiment 3 suggest that expectations do act to sharpen stimulus representations in this area, thereby indicating that the representations activated during action at least overlap with those representing the visual properties.

Experiment 3 also found that this right occipitotemporal ROI region exhibited an interaction between action-outcome congruency and task-relevance, with a pattern reflecting trends towards superior decoding of congruent outcomes when these were task-irrelevant and better decoding of incongruent outcomes when these were task-relevant. This result contrasts with previous reports where effects of expectation and task-relevance on stimulus decoding did not interact (Kok et al., 2012), and may be indicative of a mechanistic interaction between top-down prediction and attention (Summerfield & Egner, 2009). This result is predicted by neither the Cancellation model or the OPPOSE model, although it is perhaps worth interpreting cautiously given that neither simple effect contributing to it was significant.

4.3. Chapter summary

Experiment 3 presented evidence for enhanced multivariate decoding of expected, relative to unexpected visual action outcomes in clusters spanning a number of visual brain regions. These results are hard to reconcile with the Cancellation model, but are consistent with the idea that sensory predictions generated during action operate similarly to those in non-action contexts as hypothesised by the OPPOSE model.

Chapter 5: Influences of action on perception of imitative reactions

It is widely assumed that the sensorimotor models we use to predict the direct consequences of our own movements can be used to support perception of imitative reactions in others – leading to smooth, fluent social interactions. However, while the direct outcomes of our movements typically occur at short, predictable delays after action execution, the reactions of others occur at longer, variable delays in the order of seconds. This temporal variability could preclude the prediction of reactions, given that sensorimotor predictions have previously been argued to operate with a subsecond precision. Experiments 4 and 5 were conducted to investigate the temporal profile of visuomotor prediction, assessing whether predictive effects generalise across the suprasecond delays that likely characterise social interaction. Participants completed a task based on the paradigm used in Experiment 1, where congruent or incongruent action outcomes could occur synchronously with action (0 ms) or at a longer delay (1800 ms or 3600 ms) after action execution. Consistent with Experiment 1, Experiment 4 revealed that congruent action outcomes were rated as more intense than incongruent outcomes. Importantly, this facilitatory effect was found irrespective of whether outcomes occurred immediately or at delay. Experiment 5 replicated this finding and demonstrated that it was not the result of response bias. These findings therefore suggest that sensorimotor predictions operate with sufficient temporal-generality to support perception of imitative reactions in others.

5.1. Introduction

Behavioural success for humans and other animals often depends on effectively navigating the social world – on accurately anticipating and interpreting the actions of conspecifics (Happé, Cook, & Bird, 2017). Section 1.3 introduced the influential hypothesis that agents may use sensorimotor predictions derived from models of their own actions to anticipating similar reactions in their interaction partners, given that sensory consequences produced by self and other are perceptually similar (e.g. when we wave to a friend and they return the gesture). This generalised predictive process is argued to increase our sensitivity to matching responses in others; promoting fluent social interactions by facilitating rapid and appropriate responding to our partner's behaviour (see Schütz-Bosbach & Prinz, 2007 for a discussion). However, as noted in Section 1.3.2, existing work looking at the temporal profile of sensorimotor prediction have suggested that the underlying mechanisms operate with a subsecond tuning (Bays et al., 2005). Given that the imitative reactions of other occur in the range of seconds – not milliseconds – such tight temporal tuning may preclude imitative reactions from benefitting from predictive processes.

However, to date no work has considered the temporal precision of visual prediction during action and since we predominantly experience the reactions of others visually it is important to establish whether these mechanisms operate with a temporal profile that could support the perception of imitation. Experiment 4 was conducted to investigate the temporal profile of visual prediction during action, utilising the effect of prediction identified in Experiment 1, where congruent action outcomes were rated to be more intense that incongruent action outcomes when perception was probed at short delays (50 ms) after action execution. The task was adapted such that participants executed finger movements and were presented with congruent or incongruent action outcomes (same or opposite finger) at different delays after execution (0, 1800 or 3600 ms). The suprasecond delays employed mirror the natural delay with which actions are likely imitated, given experimental reports that prosocial effects of being imitated arise with action-outcome delays of 2-4 s (Bailenson & Yee, 2005; see Catmur & Heyes, 2013). If sensorimotor predictions operate across variable action-outcomes in a manner which could support social interaction, signatures of prediction should be found across delays, with participants rating congruent action outcomes as more intense (brighter) than incongruent outcomes. Alternatively, if sensorimotor predictions show tight temporal tuning, congruency should interact with action-outcome delay.

5.2. Experiment 4

5.2.1. Methods

Participants

Twenty-four participants (16 female, mean age = 25.1 years, SD= 6.9) were recruited from Birkbeck, University of London and given a small honorarium for their participation. Two of these were replacements for participants who did not follow task instructions or where PSEs modelled to their responses were beyond the range of presented stimuli.

Procedure and Stimuli

Procedure and stimuli were identical to Experiment 1 with the following changes (see Figure 5.1). The onscreen avatar hand produced observed actions either in synchrony with the participants' own executed finger lift (0 ms delay) or at one of two suprasecond delays (1800 or 3600 ms) with equal probability. Unlike Experiment 1, the increase in brightness on the observed finger always occurred 50 ms after onset of the observed movement (i.e. the 200 ms delay was eliminated). After the presentation of the comparison stimulus participants made a comparative judgement (i.e. which stimulus was brighter?) rather than an equality judgement. While comparative judgements are potentially more prone to response biases (i.e. biases to select one response alternative influence PSE values), previous work has found comparative judgements to be more sensitive than equality judgements given that the latter are typically harder and lead to noisier PSE estimates (Anton-Erxleben et al., 2010). Given previous reports that prediction mechanisms may show subsecond tuning (e.g. Bays et al., 2005) this more sensitive procedure was used in Experiment 4 to minimise the likelihood of a Type II error.

There were at least 420 trials; 70 at each of the three delays where test stimuli were congruent with the lifted finger (i.e. execute index lift, observe index lift) and 70 where test stimuli were incongruent with the lifted finger (i.e. execute index lift, observe middle lift). The experiment was divided into five blocks. The first four blocks comprised 84 trials each, while the fifth ran until participants had completed 210 of each finger lift (i.e. index or middle). In breaks between blocks participants were given on screen feedback regarding the percentage of index and middle finger lifts executed across the experiment. Responses beyond the 210th trial for each movement were not recorded. Trial type was randomized and participants completed eight practice trials. Psychometric functions were modelled separately to congruent and incongruent response data as in Experiment 1, separated by each action-outcome delay condition, though these took the form of cumulative Gaussians rather than standard Gaussians, and associated pDev statistics were calculated to establish the goodness-of- fit for each function (Palamedes toolbox; Kingdom & Prins, 2009; see Figure 5.2.)

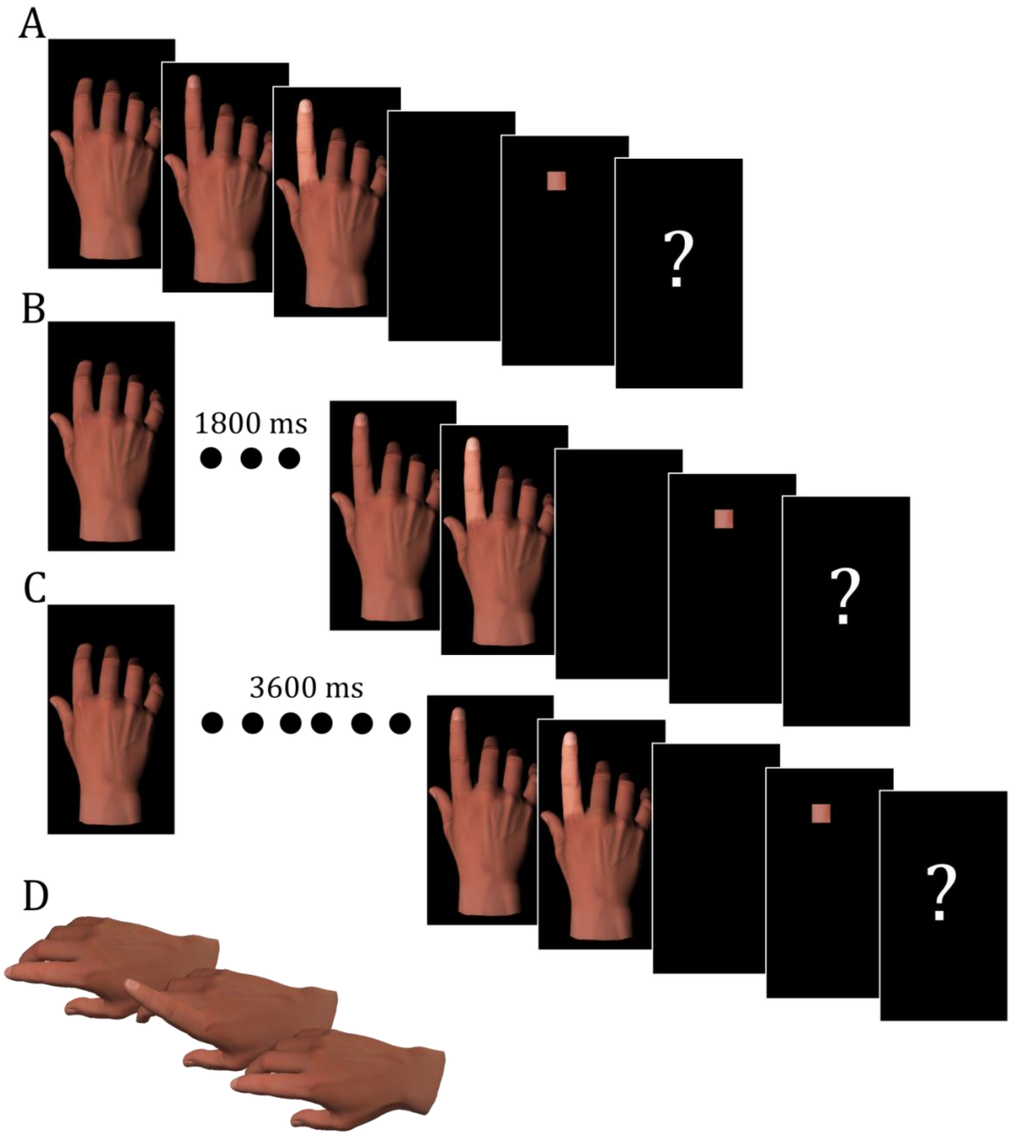

Figure 5.1: The timecourse of stimulus presentation on a single trial in both Experiments 4 and 5. Sensory outcomes were presented at a 0ms (A), 1800 ms (B) or 3600 ms delay (C) relative to the participant's own executed action (D). In the above figure the observed outcome (index lift) is congruent with the executed action. Note that in Experiment 4 participants were required to report which event (target or reference) was brighter, while in Experiment 5 participants were required to report whether the target and reference events were the same brightness or not. Hand stimuli were generated using Poser 7.0 (Smith Micro Software).

5.2.2. Results and Discussion

PSE and precision values were analysed with a 2 x 3 repeated measures ANOVA, with factors of Action Congruency (congruent, incongruent) and Delay (0, 1800, 3600 ms). Where appropriate, Greenhouse-Geisser corrections were applied. The precision analysis found no effect of Action Congruency (p=.264) or Delay (p=.086) and no interaction between these factors (p=.423). However, importantly, the PSE analysis revealed a significant main effect of Action Congruency ($F(1,23) = 7.96$, p=.010, $\eta_p{}^2$ =.257) reflecting the fact that participants had lower PSEs (i.e., brighter target percepts) for congruent (mean = 36.7%, SEM =1.6) compared to incongruent stimuli (mean = 39.4%, SEM =1.9; see Figure 5.2.). No main effect of Delay (p=.067) or importantly, any interaction between Congruency and Delay (p=.807) was observed.

To examine whether the non-significant interaction between Congruency and Delay reflects the absence of an effect or a lack of statistical power, we calculated a Bayes Factor (BF_{10}) which represents the ratio of evidence for the alternative model over evidence for the null model. It is assumed that BF_{10}<.33 provide good evidence to support the null (Jeffreys, 1939; Lee & Wagenmakers, 2014). We conducted a repeated measures Bayesian ANOVA in JASP (Love et al., 2015) with the same factorial structure. This analysis revealed evidence for the null hypothesis over an interaction effect ($BF_{10} = 0.145$).

The effect of Action Congruency, and the absence of its interaction with Delay, provides evidence that sensorimotor predictions exhibit the requisite generality to predict imitative responses across action-outcome delays comparable to those found in natural social settings.

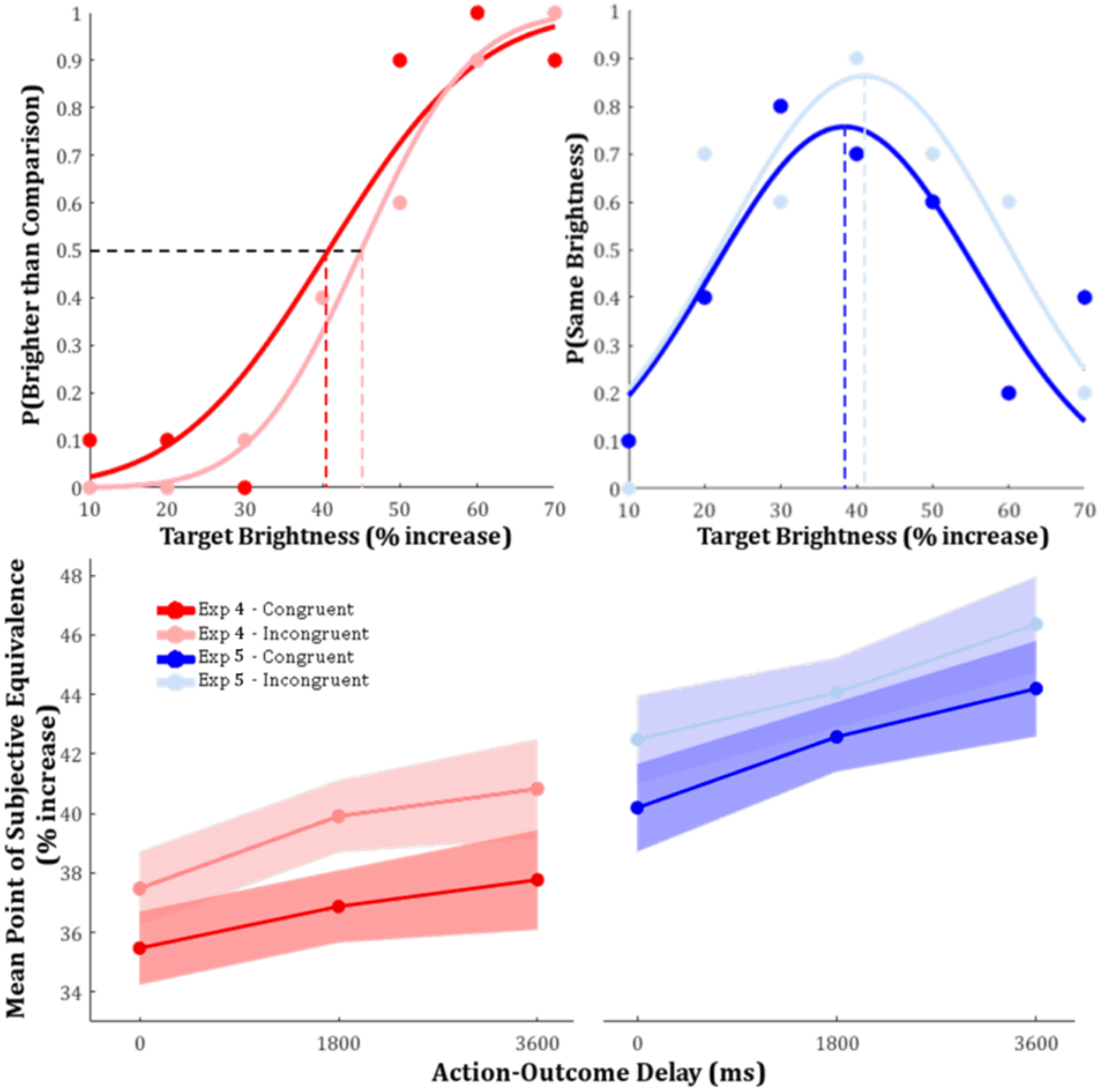

Figure 5.2: Top panel: Demonstration of how the point of subjective equivalence (PSE) was calculated in Experiment 4 (left) and Experiment 5 (right) with psychometric functions from example participants for stimuli congruent (solid) and incongruent (faded) with executed actions. The PSE describes the point where participants judge the target and comparison events to have equal brightness. When target stimuli are perceived as brighter PSEs tend towards *lower* values. Bottom panel: Mean PSEs for stimuli congruent and incongruent with action, across all action-outcome delays across both experiments. Error bars represent 95% within-participant confidence intervals (Cousineau, 2005).

5.3. Experiment 5

The results from Experiment 4 provide evidence that congruent action outcomes are judged to be more intense than incongruent action outcomes regardless of whether these outcomes occur immediately after action execution (0 ms condition: replicating Yon & Press, 2017) or if delays in the order of seconds are imposed between actions and outcomes (1800 and 3600 ms conditions). This result is consistent with the hypothesis that sensorimotor predictions may influence perceptual processing of imitative reactions, increasing their apparent intensity via additive biases in sensory activity that result from prediction (Han & VanRullen, 2016; Summerfield & de Lange, 2014; Yon & Press, 2017).

However, the PSE measure chosen to index perceptual biases induced by sensorimotor predictions could be influenced by response biases (as discussed in Chapter 2). For example, if all action outcomes are perceived with equivalent brightness but participants are biased to select the interval where congruent outcomes occur, PSEs would be lower for congruent than incongruent events. Such biases could occur given previous reports that activity in decision making circuits can be biased towards responding for expected events before a stimulus has been presented (Hanks et al., 2011; Summerfield & de Lange, 2014). A particular concern is that sensorimotor predictions may generate perceptual biases when action outcomes are simultaneous (given that Experiment 1 in the present thesis already rules out response biases for the simultaneous effects), but generate response biases at later delays – giving the impression of a temporally-general perceptual prediction.

As discussed, Experiment 4 used a comparative judgement to minimise the likelihood of a Type II error. However, given that effects were observed in this experiment, Experiment 5 was subsequently conducted to address the concerns associated with the use of a comparative judgement; by utilizing an equality judgement (i.e. were the two events the same or different brightness?). The advantage of the equality judgement is that it no longer requires participants to select a stimulus as more intense on a given trial, and ensures biases to opt for a given response alternative (e.g. same or different) no longer influence the PSE value (Han & VanRullen, 2016; Schneider & Komlos, 2008). If the effects identified in Experiment 4 operate at a perceptual locus they should also be found in Experiment 5, while effects driven by response biases should not remain.

5.3.1. Methods

Participants

Twenty-four new participants (21 female, mean age = 21.4 years, SD= 2.5) were recruited from Birkbeck, University of London and given a small honorarium for their participation. Fourteen of these were replacements for participants who did not follow task instructions or where points of subjective equivalence modelled to their responses were beyond the range of presented stimuli. This relatively high exclusion rate in comparison to Experiment 4 most likely reflects the fact that equality judgements are much more difficult and generate noisier estimates of underlying psychometric functions than do comparative judgements with equivalent stimuli (Anton-Erxleben et al., 2010).

Procedure and Stimuli

The procedure and stimuli used in Experiment 5 were identical to Experiment 4. However, participants were not asked to judge whether the target or reference stimulus was brighter, but whether the two events were the same or different brightness. Psychometric functions were therefore modelled as standard (not cumulative) Gaussians (as in Experiment 1).

5.3.2. Results and Discussion

PSE and precision values were analysed with a 2 x 3 repeated measures ANOVA, with factors of Action Congruency (congruent, incongruent) and Delay (0, 1800, 3600 ms). Where appropriate, Greenhouse-Geisser corrections were applied. The precision analysis revealed a main effect of Delay - $F(2,46) = 4.843$, $p=.023$, $\eta_p^2 =.174$ – which followed a significant quadratic trend, $F(1,23) = 6.139$, $p=.021$, $\eta_p^2 =.211$. This effect reflected the fact that judgement precision was significantly lower for events presented at the 1800 ms action-outcome delay (mean = 29.9 %, SEM =2.08) than events presented at either the 0 ms delay (mean =25.9 %, SEM = 1.88; $t(23) = 2.373$, $p=.026$) or the 3600 ms delay (mean = 26.8 %, SEM = 1.50; $t(23) = 2.366$, $p=.027$; precision at the 0 ms and 3600 ms delays did not differ, $p=.330$). No other significant effects were obtained in the precision data (all $p\geq.530$).

The PSE analysis also revealed a significant effect of Delay - $F(2,46) = 4.152$, $p=.022$, $\eta_p^2 =.153$. This main effect of Delay followed a linear trend – $F(1,23) = 7.22$, $p=.013$, $\eta_p^2 =.239$ – and reflected lower PSEs (i.e. brighter percepts) for events presented at the 0 ms delay (mean = 42.1 %, SEM = 2.31) relative to the

3600 ms delay (mean = 45.7 %, SEM = 1.78; $t(23) = 2.69$, p=.013), with the PSEs lying intermediate at the 1800 ms delay (mean = 43.8%, SEM = 2.19) although not differing significantly from 0 ms (p = .211) or 3600 ms (p =.079) conditions.

Crucially, the PSE analysis again demonstrated a significant effect of Action Congruency - $F(1,23) = 7.313$, p=.013, $\eta_p{}^2$ =.241 – which did not interact with Delay (p=.928). This main effect reflected lower PSEs for congruent action events (mean = 42.9 %, SEM = 2.06) than incongruent events (mean = 44.8 %, SEM = 1.96) – suggesting that predicted action outcomes were perceived as brighter than unpredicted ones, and that this pattern does not reflect a response bias caused by expectation (Hanks et al, 2011). The non-significant Action Congruency x Delay result was again shown through a Bayesian ANOVA in JASP to reflect the absence of an effect ($BF_{10} = .126$).

The effect of Delay on judgement precision was unpredicted. A speculative explanation for this effect is that the timing of the event is perhaps least expected at 1800 ms (see Rohenkohl, Cravo, Wyart & Nobre, 2012), relative to the timepoint at which action outcomes typically occur (0 ms) and the timepoint which is the latest employed in this paradigm (3600 ms; i.e., if an event has not been presented by this point in the epoch then its arrival is certain). One could further speculate that such an effect is more pronounced in more difficult tasks (see Ratcliff, 2014), explaining why it was obtained with the harder equality judgement in Experiment 5 but not the easier comparative judgement in Experiment 4 (Anton-Erxleben et al., 2010). The main effect of Delay on PSEs may similarly reflect greater processing of events occurring at the earliest timepoint, given the typically strong likelihood that outcomes are simultaneous with action.

Most importantly, the main effect of Action Congruency and the convincing absence of an Action Congruency x Delay interaction suggest that the results of Experiment 4 were not due to participants being more likely to select the congruent event through response bias. Instead, they provide further support for the hypothesis that sensorimotor predictions exhibit generality across action-outcome delays.

5.4. General Discussion

It is widely assumed that the sensorimotor predictive processes we use for controlling our own actions are useful for supporting perception of the imitative reactions of others (e.g., Schütz-Bosbach & Prinz, 2007). However, it has been unclear whether sensory predictions generated during action have the requisite generality across time to be useful for perceiving imitation in others. Both Experiments 4 and 5 found evidence consistent with the hypothesis that such generality is indeed present - signatures of predictive processing, i.e., more intense percepts for action-congruent relative to -incongruent sensory events, were found regardless of whether outcomes were observed immediately or at suprasecond delays.

The findings of Experiments 4 and 5 are consistent with the idea that expectations bias activity in expected sensory units, leading to an enhancement in effective signal strength (Wyart, Nobre & Summerfield, 2012; see also

Tsetsos & Summerfield, 2015; Carrasco, Ling & Read, 2004). As outlined in Section 1.2.2, such a mechanism is thought to facilitate perception of expected events in a number of ways e.g., speeding reaction times (Bar, 2004) or accelerating the entry of stimuli into conscious awareness (Pinto et al., 2015). The present findings crucially demonstrate that visual predictions during action operate with a temporal profile that would allow them to influence imitative reactions. The resulting increased sensitivity to these reactions is proposed to promote fluent social interactions by facilitating rapid and appropriate responses to our partner's behaviour (e.g., Schütz-Bosbach & Prinz, 2007). It has also been suggested that it may facilitate pro-social effects that result from perceiving imitation in others, e.g., increased social motivation and rapport (e.g., Bailenson & Yee, 2005; Neufeld & Chakrabarti, 2016; van Baaren, Holland, Steenaert, & van Knippenberg, 2003) given that the signal on which these effects depend is enhanced. Furthermore, one could speculate that – by comparison between predictions and sampled inputs - these predictive processes could also assist the detection of unexpected *absences* of imitation (e.g., when a friend does not return a wave). Sensitivity to such prediction errors could be useful in highlighting the need for further learning (e.g. to infer whether they are upset with are us) or the need for a novel response (e.g. to call their name).

Interestingly, the temporal generalisation found here appears inconsistent with a handful of studies in the tactile domain which suggest that sensorimotor predictions show precise temporal tuning[1]. For example, signatures of sensorimotor prediction on tactile perception are seen when action execution and sensory outcomes are synchronous, but these influences are not seen when delays of ~300 ms are imposed between action and outcome (Bays et al., 2005; Blakemore et al., 1999). These findings have been argued to reflect adaptive processes for action control, given that subsecond tuning will allow the predictive processes to contribute to rapid error correction (Wolpert et al., 1995) and the agentive labelling of self-produced events (Frith, Blakemore, & Wolpert, 2000). However, the present experiments present the first investigation of the temporal properties of *visual* prediction during action and these discrepant findings can therefore be integrated if it is assumed that visual predictions exhibit broader temporal tuning than tactile predictions. Given that imitative reactions are typically perceived visually, such a separation between modalities may be most adaptive in balancing temporal sensitivity to optimise action control with the temporal generality required to optimise perception of others.

The hypothesis tested in Experiments 4 and 5 proposes that predictions we generate about the consequences of our own actions also influence perceptual processing of imitative reactions of others (Schütz-Bosbach & Prinz, 2007; Wolpert et al., 2003). This hypothesis would appear to assume that the learning

---

[1] The experiments supporting this assumption have been designed to examine action control theories. To this end they have compared predictive effects for simultaneous outcomes and those following subsecond delays and consistently find the largest effects in simultaneous conditions. Given that our primary aim was to test theories from social cognition, it was more appropriate to compare simultaneous conditions against those employing suprasecond delays given that imitative reactions occur in the order of seconds, not milliseconds (see Catmur & Heyes, 2013).

we acquire through sensorimotor experience with our own direct action effects at subsecond delays (e.g., observing our own index finger tapping; (de Wit & Dickinson, 2009; Birgit Elsner & Hommel, 2004; Daniel M. Wolpert, Diedrichsen, & Flanagan, 2011) generalises over suprasecond action-outcome delays (Pearce, 1987). Under this assumption, these processes may only operate for the prediction of *imitative* reactions in others. However, an alternative possibility is that predictions at delay may be derived from specific experience of others imitating our actions at delay. This account is consistent with the idea that the socio-cultural environment is responsible for furnishing the mechanisms needed for social interaction (Brass & Heyes, 2005; Heyes, 2012). This latter possibility is easier to reconcile with reports of temporal precision in touch (Bays et al., 2005), as we have little opportunity to learn that tactile outcomes follow our actions at delay. Importantly, this hypothesis assumes that the predictive process is not dependent on the fact that imitative reactions match our actions but only that they are typically the most probable reaction (see Ray & Heyes, 2011). The same process could therefore in principle facilitate perception of non-matching reactions in settings where these are more likely. For instance, expressions of aggression and dominance (e.g., expansive postures) tend to elicit expressions of submission (e.g., constricted postures) rather than mirrored aggression (Tiedens & Fragale, 2003). Under this account, any predicted reaction would benefit from the same facilitated processing, hence further contributing to smooth and appropriate responses to others during social interaction.

5.5. Chapter summary

Experiments 4 and 5 presented evidence that congruent action outcomes appear with greater intensity than incongruent events, regardless of whether they occur simultaneously with action execution or at a suprasecond delay. This pattern is consistent with suggestions that the predictive processes used to anticipate the direct consequences of our own actions may support perceptual processing of others' reactions.

Chapter 6 – Influences of action on duration perception

While there are a number of empirical reports consistent with the idea that expectations generated on the basis of action are incorporated into perception, a notable gap in this literature concerns our perception of time and it remains unclear whether sensorimotor predictions influence how we perceive temporal features. Interestingly, extant models of time perception make divergent predictions about the possibility of such sensorimotor interactions. This chapter investigates this issue in a series of psychophysical experiments examining whether explicit (Experiments 6 & 7) or implicit (Experiments 8 & 9) variation in the duration of executed actions biases duration perception. Across all experiments, participants were found to judge auditory tones as longer when producing movements of longer duration. Results suggested this effect was due to the parameters of executed movements, after controlling for linguistic and perceptual confounds. These findings are consistent with the idea that actors generate expectations about the likely temporal properties of action outcomes, and that these expectations are incorporated into time perception.

6.1. Introduction

Do sensory predictions generated during action influence how temporal properties are perceived? In Section 1.2.2 we outlined how it would be adaptive for observers to incorporate predictions generated on the basis of action into our estimates of the sensory world. There are numerous demonstrations that perception of spatial properties and event identity is influenced by predictions generated during action. As discussed, these biases toward predicted action outcomes may be adaptive, given that expected outcomes are by definition more likely and therefore these biases will, on average, generate more veridical percepts.

Such adaptive arguments apply equivalently to the perception of temporal features, given that strong correlations between the temporal properties of executed movements (e.g. bowing a cello) and the sensory consequences they produce (e.g. the duration of the note) could provide a rich source of expectations aiding perception. However, existing models of time perception make divergent predictions about the possibility of predictive interactions between motor and sensory timing (Ivry & Schlerf, 2008). Under dedicated timing models, temporal information (e.g. the duration of a bowing action, the duration of a note) is assumed to be explicitly represented in a format that allows interactions in timing across modalities (e.g. Ivry, 1996). In contrast, under intrinsic timing models temporal information is only represented implicitly, in sensory-specific cortical mechanisms, and in a fashion that precludes integration (Buonomano, 2000; Karmarkar & Buonomano, 2007; see Section 1.4.2.).

Previous work has not considered the possible influences of concurrent temporal signals on time perception - in action or non-action domains. Therefore to investigate whether observers incorporate top-down expectations generated during action into time perception, I conducted a series of psychophysical experiments where participants judged the duration of an auditory tone while concurrently producing movements of varying durations. Given the typically strong correlations between the temporal features of actions and their perceptual consequences, observers may be expected to incorporate temporal information from their actions into their perceptual judgements, biasing duration percepts towards concurrent actions.

In Experiment 6 participants were explicitly instructed to produce long and short duration finger taps while judging the duration of a concurrently presented tone. Experiment 7 used a similar design to Experiment 6 while controlling for a potential linguistic confound in the movement cuing procedure. In Experiment 8, an implicit movement paradigm was used (participants reached toward Near or Far targets over a motion-tracker) while participants made similar judgements about tone durations. Experiment 9 mirrored the paradigm used in Experiment 8, controlling for a potential perceptual confound in the implicit movement paradigm.

6.2. Experiment 6

6.2.1. Methods

Participants

Sixteen participants (nine female, mean age = 24 years, SD = 4.85) were recruited from Birkbeck, University of London and paid a small honorarium. Four of these acted as replacements for excluded participants where performance was highly variable, yielding psychometric functions that could not be modelled effectively in at least one condition (see below). The number of recruited participants was lower than that in previous studies due to pilot testing which indicated that the likely effect size was large, and therefore that 16 would be sufficient to detect effects.

Procedure and Stimuli

Participants were seated approximately 55 cm away from a computer monitor inside a dimly lit cubicle, with their right hand positioned over a keypad placed in front of the body midline. A black screen prevented participants from viewing their right hand. Perceptual judgements were made with the left hand on a separate keypad that was positioned to left of the screen covering the right hand. The experiment was run using the Cogent toolbox for MATLAB. All visual stimuli were presented against a black background on a CRT computer monitor (32 x 24 cm, 85 Hz, 21 DPI), while all auditory stimuli were presented over Sennheiser HD 201 headphones (81 dB).

The experiment began with a short block to train participants to produce short and long duration movements. Participants depressed the response key with the index finger of their right hand, maintaining this position during the inter-trial interval. On each trial, an ‘S’ or ‘L’ was presented on the screen, indicating the required movement duration. The cue remained visible until participants lifted their finger to initiate the movement. After the lift, participants re-pressed the response key, attempting to match the criterion movement duration (<500 ms for short responses, >950 ms for long responses). The movement was modelled by the experimenter such that the finger moved smoothly to and from the apex of the lift. The training block consisted of 20 trials, ten short and ten long. Participants received auditory feedback (100 Hz, 500 ms long) when the produced duration was outside the criterion for that trial. Each movement was followed by an inter-trial interval of 500 ms, at the end of which the next cue appeared.

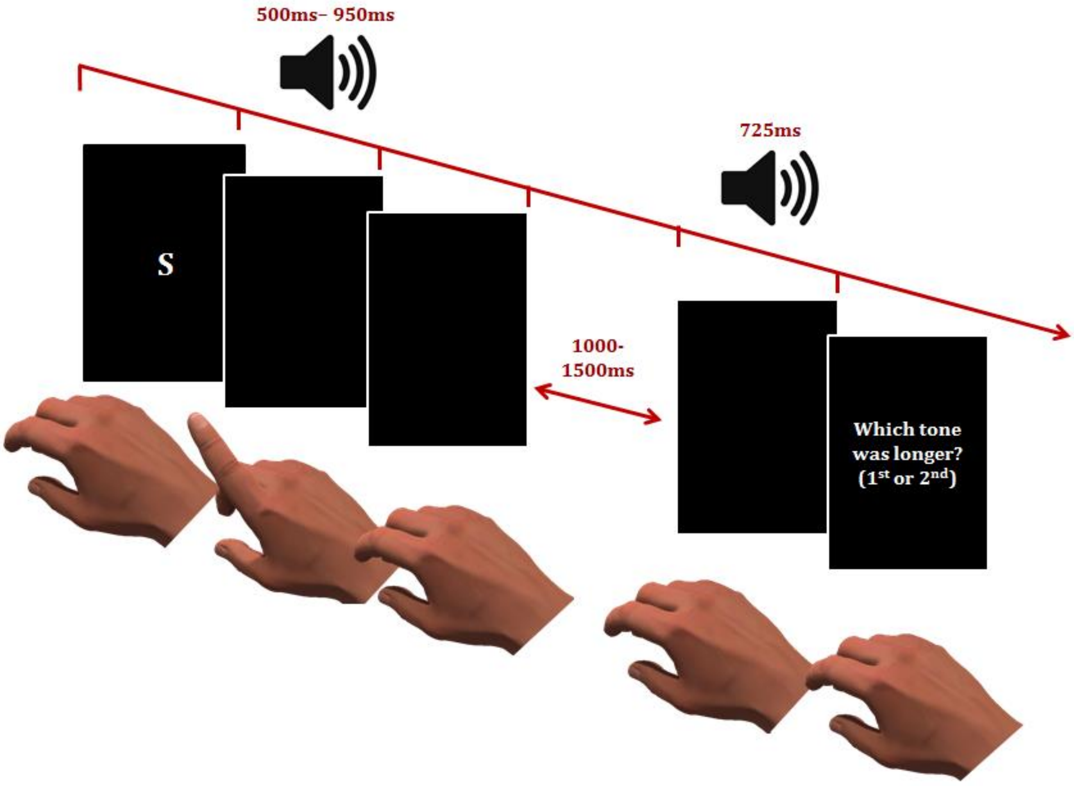

Figure 6.1: Time course of stimulus presentation and participant movements on a single test trial in Experiments 6 and 7.

In the main experimental blocks, participants performed two, interleaved tasks: the movement task as described above and a duration discrimination task. The time course of the events for each trial is illustrated in Figure 6.1. The trial started with the presentation of the duration cue ('S' or 'L'). When participants initiated their finger lift, the screen was blanked and the test tone was played. The duration of this tone was drawn from a set of seven durations, ranging from 500 to 950 ms (75 ms steps, all 500 Hz). At a variable delay after participants returned their finger to the start position (1000-1500 ms), the reference tone of a fixed duration, 725 ms, was played (again 500 Hz). If the cued movement had been executed correctly, a response screen appeared instructing participants to indicate which tone was longer. The tone judgements were indicated by a keypress with the left thumb, pressing the left key if the first tone was longer and the right key if the second tone was longer. If the movement duration was outside the criterion for that trial, an error message was displayed ('Too fast!' or 'Too slow!'). No perceptual judgements were obtained on these trials and the trial was repeated.

Participants completed eight practice trials of the concurrent movement/duration discrimination task. Test blocks were a minimum of 70 trials each (five of each target duration for each movement condition) and the trial order was randomized. There were two test blocks. Thus, the final data set consisted of 10 trials at each test tone duration in which the required movement duration was correctly produced.

When the experiment had been completed, participants were asked whether they had noticed anything about the tones, and in particular, anything of note about the second tone. None reported being aware that the reference tone was identical in duration on every trial (this same pattern was observed in all experiments).

To estimate psychometric functions, the perceptual judgements for each individual were modelled by fitting cumulative Gaussians, and associated pDev statistics were calculated to establish the goodness-of fit for each function as in previous chapters (Palamedes toolbox, Kingdom & Prins, 2009; See Figure 6.2). This procedure was performed separately for the long and short response conditions.

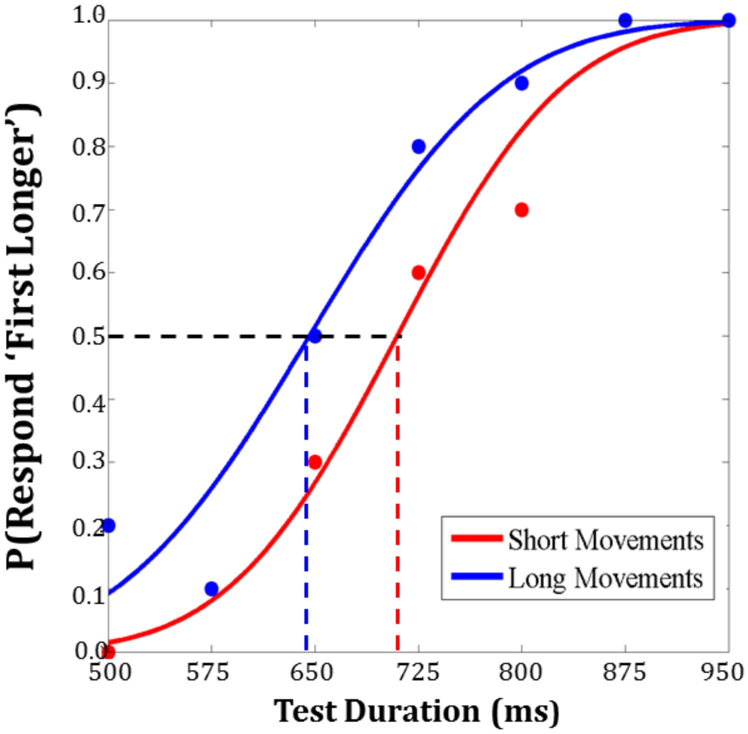

Figure 6.2: Illustration of how PSEs were calculated from cumulative Gaussians modelled to judgements in the short and long movement conditions of Experiments 6-9. The PSE (dotted line) describes the point where participants judge the target and reference events to have equal duration. Please note the data points for the two conditions overlap for stimulus durations of 575 and 875 ms.

6.2.2. Results and Discussion

Participants had little difficulty producing movements of different durations in the two conditions. The mean number of movement errors was 7.7% (SEM = 1.9%). The mean duration of correct movements was 255.4 ms (SD = 105.9 ms) in the short condition, and 1577.5 ms (SD = 521.6 ms) in the long condition (Figure 6.3).

There were no differences between the two movement conditions in terms of the precision of the perceptual judgements ($t(15) = .678$, $p = 0.51$). The mean Weber fractions (precision/duration of the reference stimulus) were .65 (SEM = .09) and .70 (SEM = .09) in the Long and Short movement conditions, respectively. As can be seen in the example psychometric functions (Figure 6.2), participants were more likely to judge the target stimulus as longer on trials requiring a long duration movement. This bias results in lower PSE values for the long duration movement condition compared to the short duration movement condition (Figure 6.3), an effect that was highly significant ($t(15) = 7.82$, $p < .001$, $d = 2.21$). This PSE effect suggests that duration judgements are biased towards the duration of concurrent actions, consistent with a predictive interaction between motor and sensory representations of time.

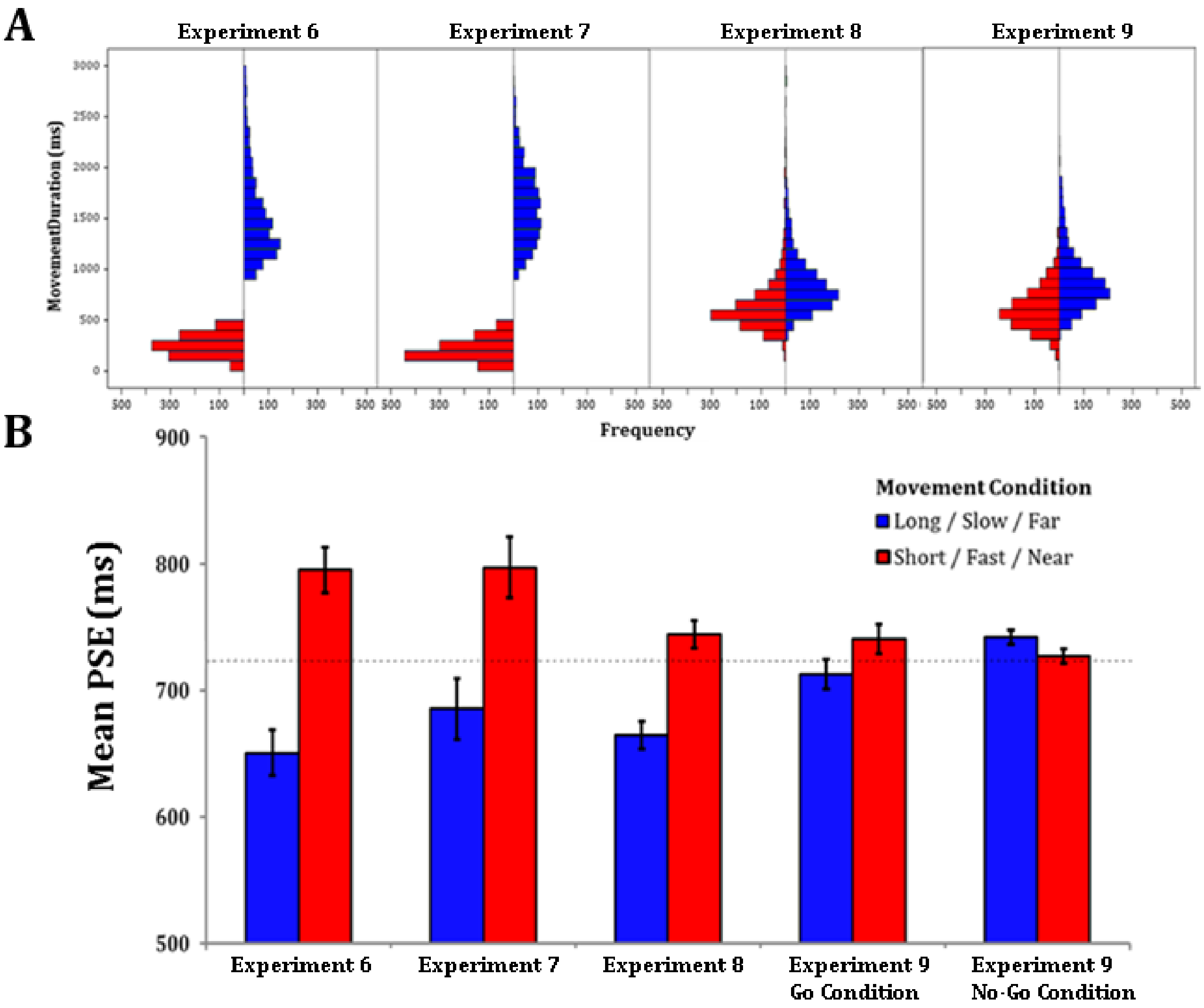

Figure 6.3: A: Distribution of movement times across the conditions of Experiments 6 -9 collapsed across all participants. B: Mean PSEs in the conditions across experiments. Note that lower PSEs indicate a bias to perceive the target stimulus as *longer*. Dotted line indicates the duration of the reference tone (725 ms). Error bars show 95% within-subject confidence intervals (Cousineau, 2005).

6.3. Experiment 7

The results of Experiment 6 are consistent with the hypothesis that predictions are generated during action about the likely duration of subsequent sensory events, and that perception is biased towards these expected durations. However, there are alternative explanations for the pattern of results obtained.

One possibility is that the bias in duration perception described in Experiment 6 arises due to linguistic rather that motor influences. Each movement condition was described in terms of duration, and on each trial, participants saw an explicit linguistic cue, ‘L’ to indicate the production of a long duration movement and ‘S’ to indicate the production of a short duration movement. It is possible that predictions generated on the basis of linguistic representations were responsible for the perceptual bias, given that it is well known that perception of other low-level features (e.g. visual motion direction) can be influenced by linguistic predictions (Francken, Meijs, Hagoort, Gaal, & Lange, 2015).

To evaluate this possibility we repeated the basic procedure of Experiment 6, but modified the manner in which the movement task was described and cued to reduce the linguistic overlap with the perceptual task. In Experiment 7, participants were instead instructed to vary the speed (rather than duration) of their movements, executing ‘fast’ or ‘slow’ movement in response to onscreen cues (‘F’ and ‘S’, respectively). Cuing movement in this manner should reduce the linguistic overlap. If the effect described in Experiment 6 is mediated by linguistic predictions introduced by ‘long’ and ‘short’ codes, we would expected the perceptual biases generated by movement to be abolished in Experiment 7.

6.3.1. Method

Participants

Sixteen new participants (13 female, mean age = 22 years, SD = 3.53) were recruited from Birkbeck, University of London. Seven of these were replacements for excluded participants where performance was highly variable, yielding psychometric functions that could not be modelled effectively in at least one condition.

Procedure and stimuli

The stimuli and procedure for Experiment 7 were identical to Experiment 6 with the following change: participants were instructed to perform 'slow' and 'fast' movements, cued with the letters 'S' and 'F', respectively. The criteria were identical to those of Experiment 6: 'fast' movements were required to be <500 ms in duration and 'slow' movements were required to be >950 ms in duration.

6.3.2. Results and Discussion

As in Experiment 6, participants had little difficulty meeting the demands of the movement task. The mean durations were 213.4 ms (SD = 103.9 ms) in the short condition, and 1644.2 ms (SD = 503.9 ms) in the long condition (Figure 6.3). Movement errors occurred on 7.1 % (SEM = 2.1 %) of the trials.

The comparison of the PSEs on the duration perception task again revealed a highly significant effect ($t(15) = 4.56$, $p < .001$, $d = 1.76$). The PSEs were lower in the slow movement condition relative to the fast condition (see Figure 6.3). As in Experiment 1, there was no difference between the movement conditions in terms of precision ($t(15) = 1.89$, $p = .078$), with mean Weber fractions of .79 (SEM = .15) and .55 (SEM = .04) in the Slow and Fast conditions, respectively.

In sum, the results of Experiment 7 show that tones are perceived to be longer when performing long movements, even when the lexical terms used to describe movement goals do not directly refer to duration.

6.4. Experiment 8

Experiment 7 allows us to reject an explanation of the perceptual bias based on the direct influence of linguistic codes (i.e. 'long' or 'short'). However, it remains possible that effects are mediated via indirect linguistic links even if the conditions were labelled using the terms 'slow' and 'fast'. While adults understand time and speed as distinct dimensions, the lexical terms for time and speed are related. Moreover, young children tend to confuse these dimensions. Interestingly, these confusions reflect the opposite mappings to those defined by Newtonian mechanics: for example, faster moving objects are erroneously judged to move for longer periods of time (Siegler & Richards, 1979). As such, if the effect observed in Experiment 7 was mediating by linguistic representations the opposite pattern may have been expected i.e. longer duration percepts on fast movement trials.

Nonetheless, concerns about linguistic mediation can be more directly addressed by removing the labels entirely from the movement task. To this end, we designed a new task for Experiment 8 in which movements of varying duration were elicited by implicitly manipulating the other dimension that can covary with movement duration, movement distance[2](Fitts, 1954). We used a simple task in which participants reached to a visual target that could appear at

---

[2] Note that movement duration must of course always covary with either distance, speed or both.

different locations on the display. By varying the position of the target, movements could be partitioned into 'short' and 'long, permitting an analysis of whether perceptual judgements differ between the two conditions. This method has the important advantage that the instructions are the same for both condition ('reach to the target'), reducing the likelihood that effects are linguistically mediated. If similar biasing effect to that observed in Experiments 6 and 7, it seems more reasonable to attribute the effect to action duration rather than the explicit framing of the movement task.

6.4.1. Method

Participants

Sixteen new participants (7 female, mean age = 21.4 years, SD = 2.8) were recruited from Birkbeck, University of London and paid a small honorarium for participation. Five of these acted as replacements for excluded participants: four were excluded because their highly variable performance meant that psychometric functions could not be modelled effectively in at least one condition. An additional participant was excluded as movement time analysis (see below) revealed no differences in movement duration between Near and Far target conditions (t = .016, p = .988).

Procedure and stimuli

The visual targets were displayed on a vertically-oriented computer monitor positioned in front of the participant (Figure 6.4). Vision of the arm was occluded by the black screen. Movement position was recorded by a small motion tracker (Leap Motion Controller using the Matleap MATLAB interface, sampling rate = 20 Hz – note that this rate was selected during piloting to eliminate potential interference with the dynamics of stimulus presentation, yet still establish static hand position within 50 ms of arrival at the target location) which recorded the centre of the palm in 3D space.

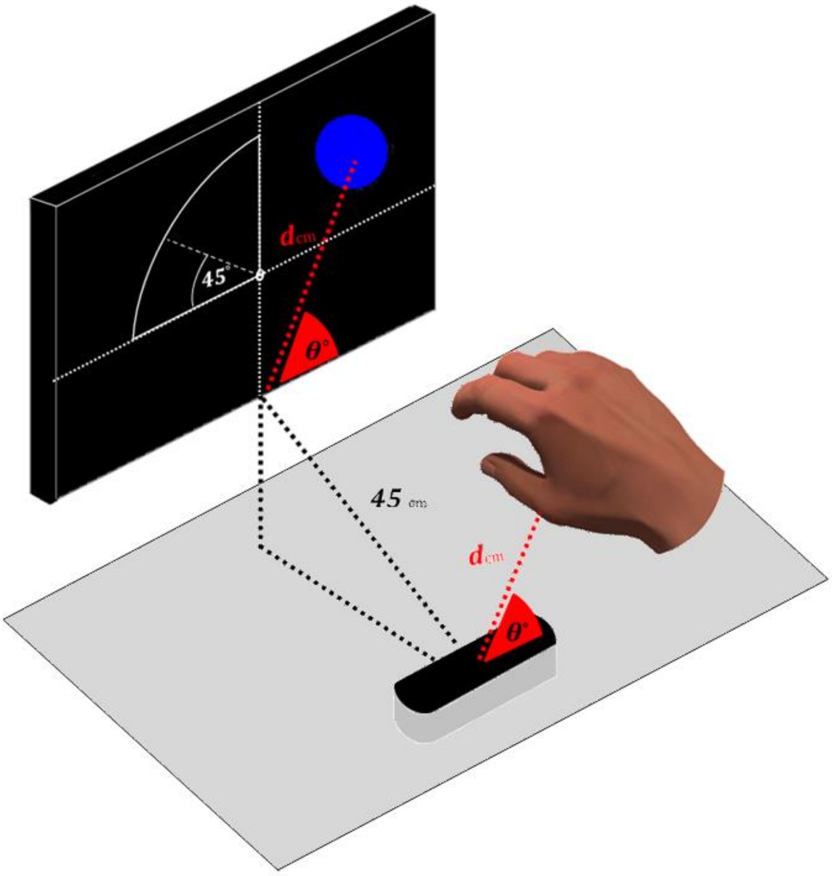

Figure 6.4: Illustration of set up in Experiment 8 and 9. A blue circle, presented on a vertically-oriented monitor, indicated the target location (defined by distance ***d*** and angle ***θ***). Target locations were limited to the upper left and right quadrants of space, presented at a random angle from 0-45° with respect to the horizontal meridian (shown in white).This generated two discontinuous wedges (one left – illustrated, one right) in which the targets could appear. Reaches were made by moving the hand above and along the horizontal surface of a table where the motion tracker was placed. Vision of the hand was occluded by a black screen. On 'No-Go' trials in Experiment 9, the targets were presented, but participants were required to remain at the central starting position.

Each trial began with presentation of cursor feedback (white dot, 0.2° visual angle) corresponding to the participant's palm position. This allowed them to verify that their hand was at a central starting position. After 2000 ms, a blue target circle (3 cm in diameter, ~ 3 ° visual angle) was presented on the display. The location of the target was constrained to fall within one of two 45° wedges about the horizontal meridian, with half of the targets appearing to the left and half to the right. Participants reached to the target by moving their hand above the table surface (without making contact; see Figure 6.4). The instructions emphasized accuracy over speed, with participants told to execute a single smooth movement, without attempting corrections (given there was no movement feedback, corrections would have been infrequent). They were to keep their hand at the final location until the end of the trial. Movement time was recorded as the interval between movement onset (>1 cm change from starting position) and offset (movement velocity<40 mm/s; both criteria were set during piloting to maximize the identification of start-stop movement periods while minimizing false alarms). If participants failed to complete the movement within 3000 ms of the target onset, they received error feedback and the trial was repeated. After 5000 ms, the screen was blanked, signalling the end of the trial. This was followed by a 2000 ms ITI, during which participants were required to move back to the central starting position. The test session began with two practice blocks (10 trials each). In the first practice block, cursor feedback was continuously provided to familiarize participants with the reaching task. This block was followed by a second practice block in which the cursor was only visible for 1000 ms at the start of each trial (also the case in the main task). Error feedback was provided if the final hand positon was outside the target circle. The error feedback was a red dot (0.2°), indicating the position of their palm.

The test block consisted of 168 trials in which participants concurrently performed the reaching and tone duration tasks. The location of the reach targets was manipulated such that 70 reaches were to targets positioned 5 – 8 cm from the start location (Near) and 70 reaches were to targets positioned 15 – 18 cm from the start location (Far). Pilot testing indicated that with these ranges there would be a measureable difference in mean movement time. To prevent participants from becoming aware of the clusters of near and far targets, we also included an additional 28 trials in which the targets were 10 – 13 cm from the start location (intermediate).

A tone was presented as soon as participants initiated the reaching movement. As in Experiments 6 and 7, the duration of this tone varied between 500 and 950 ms (75 ms steps, all 500 Hz). At a random delay after the termination of the test tone (1500 – 2000 ms) a second reference tone was presented (725 ms). Participants were instructed to complete their movements before the onset of the reference tone. Participants were considered to have missed the target if they had not reached the target area within 3000 ms of movement initiation. If the reach had landed in the target zone within this 3000 ms window, participants were required to judge which of the tones was longer, and to move back to the central start position after making their judgement. If the target was missed, participants received error feedback (red cursor indicating hand position) and the trial was repeated at the end of the experiment. Each test tone was presented 24 times; 10 times on trials with Near targets, 10 times on trials with Far targets, and 4 times on trials in which the target was at an intermediate location. Trials were randomized and breaks were taken every 20 trials. Psychometric functions were modelled to responses in Near and Far trials as in Experiments 6 and 7.

6.4.2. Results and Discussion

Despite the absence of visual feedback, participants were accurate in terminating their movements in the target region. The mean number of errors was 11.5 % (SEM = 2.3 %). Participants were also successful in terminating movements before the reference tone began (mean accuracy = 99.5 %, SEM = 0.3 %). As can be seen in Figure 3, the movement time distributions showed considerable overlap. Nonetheless, the means of the two distributions were displaced, indicating that movement time increased with movement amplitude. The mean movement times in Near and Far conditions were 619.0 ms (SD = 267.9 ms) and 859.6 ms (SD = 346.1 ms) respectively. To ensure that our distance manipulation was successful for each individual, we compared the MT distributions for each participant with a t-test. These tests were all highly significant (all ts > 2.87, all ps < .005; bar one excluded participant – see above). Thus, our manipulation of distance was effective in creating different timings for the Near and Far conditions, even though the temporal difference was much smaller than in Experiments 6 and 7.

Turning to the perceptual judgements, a comparison of PSEs on the duration perception task revealed significant differences between the movement conditions - $t(15) = 7.16$, $p < .001$, $d = 1.18$. PSEs were significantly lower in the

Far (long) target condition relative to Near (short) target condition (see Figure 6.3), suggesting that tones are perceived to be longer following reaches to more distant targets. No significant differences were found in judgement precision – $t(15) = 1.78$, $p = .096$ - with mean Weber fractions of .54 (SEM = .06) and .46 (SEM = .04) in the Far and Near conditions, respectively. In line with the findings of Experiments 6 and 7, these results suggest that the duration of executed movements influences the perceived duration of auditory events. The effect is especially striking here given that the duration differences in movement time were an emergent property of target distance, and were not explicitly marked by the instructions.

6.5. Experiment 9

Experiment 8 suggests that biases in perceived duration induced by action duration are not driven by explicit labels attached to the movement. However, by exploiting the coupling between target eccentricity and movement duration, a different confound was introduced into Experiment 8: namely, when participants produced reaches with longer duration, they also observed targets presented at greater eccentricities from the centre of the screen. It is therefore possible that the bias seen in Experiment 8 is driven by the observation of small and large visual eccentricities, rather than by the duration of the executed movements.

To evaluate this possibility we used the task employed in Experiment 8 but introduced a Go/No-Go manipulation. On each trial a target was presented, but movements were only required if the colour of the cursor representing the participant's initial palm position was green (Go trials); No-Go trials were signalled when the cursor was red. If the bias is driven by the perceptual difference between Near and Far targets, then similar biases should be observed on Go and No-Go trials. Conversely, if the bias is driven by the duration of the movements, it should only be observed on Go trials. The task was also adapted such that the reference stimulus was presented before the target and action stimulus, to ensure that any previous effects were not influenced by stimulus order.

6.5.1. Method

Participants

Sixteen new participants (10 female, mean age = 26.1 years, SD =5.3) were recruited from Birkbeck, University of London and paid a small honorarium for participation. Four of these were replacements for participants excluded because their performance yielded psychometric functions that could not be modelled effectively in at least one condition.

Procedure and Stimuli

The procedure and stimuli used in Experiment 9 were identical to Experiment 8 with the following changes. As in Experiment 8, each trial began with the presentation of cursor feedback corresponding to the participant's palm position. In Experiment 9 the colour of this cursor indicated the movement task for the forthcoming trial. On Go trials the cursor was green, indicating that participants would be expected to reach to the target. On No-Go trials, the cursor was red and participants were instructed to keep their hand in the central starting position. After 2000 ms, the 725 ms reference tone was played. At a random delay after the termination of the reference tone (1500 – 2000 ms), a blue target circle was presented. On Go trials, this served as the imperative and movement initiation triggered the presentation of the test tone. On No-Go trials the test tone was presented 500 ms after target onset. This timing was selected to approximate the interval between the target and tone onsets in the Go conditions (anticipated average RT). By using a fixed interval here, the onset of the tone was equally predictable in both conditions. Criteria for successful movements on Go trials were identical to Experiment 8, based on the terminal position of the hand with respect to the target. Movements greater than 5 cm from the start position (in any direction) were considered errors on No-Go trials. We opted to use a liberal criterion here since participants' hands tended to drift from the start position given that the hand was suspended in mid-air, and pilot testing indicated that the criterion was sufficient to detect erroneous reaches.

The test block consisted of 336 trials. Each test tone was presented 48 times: 20 times on trials with Near targets, 20 times on trials with Far targets and 8 times on trials in which the target was at an intermediate location. Half of the trials were Go trials and the other half were No-Go trials. All trial types were randomized. Psychometric functions were modelled separately for the four conditions, Near-Go, Far-Go, Near-No-Go, Far-No-Go. Breaks were provided every 30 trials.

6.5.2. Results and Discussion

Participants were accurate in performing the reaching movements on Go trials, with errors occurring on an average of 9.6 % (SEM = 1.35) of trials. As can be seen from Figure 6.3, the movement time distributions for the Near and Far conditions on Go trials showed considerable overlap. Nonetheless the means of the two distributions were displaced, indicating again that movement time increased with movement amplitude. The mean movement times for Near and Far Go trials were 583.3 ms (SD = 210.0 ms) and 872.9 ms (SD = 298.9 ms). For each participant, movement time distributions were compared with a t test. These tests were all highly significant (all *t*s > 5.98, all *p*s < .001). Thus, the manipulation of distance was effective in creating different timings for the Near and Far Go conditions, similar to that observed in Experiment 8.

Judgement precision was analysed using a 2 x 2 repeated measures ANOVA, with factors of target distance (Near, Far) and task (Go, No-Go). This analysis revealed a main effect of task, $F(1,15) = 16.18$, $p = .001$, $\eta_p{}^2 = .519$, with better discrimination performance in No-Go conditions relative to Go conditions. This effect is expected given that participants are simultaneously performing a motor task when judging the tones in the Go, but not No-Go condition. No other effects were significant (all $p > .058$). Mean Weber fractions of .41 (SEM = .04) and .50 (SEM = .07) were obtained for Near and Far trials in the Go condition, while Weber fractions of .26 (SEM = .02) and .31 (SEM = .03) were obtained for Near and Far trials in the No-Go condition.

The PSEs (now derived from cumulative Gaussians modelled to P[respond ‘second longer’] against test duration) were also analysed using a repeated measures ANOVA with the same factorial structure. This analysis found no effect of task (p = .632) or target distance (p = .385), but a significant interaction between these factors, $F(1,15) = 11.27$, p = .004, η_p^2 = .429. Simple effects analyses revealed a cross-over interaction: In the Go condition, PSEs were lower for Far targets compared to Near targets, $t(15) = 2.29$, p = .037, d = .444, while in the No-Go conditions, PSEs were lower for Near targets than Far targets, $t(15) = 2.69$, p = .017, d = .329. Therefore, as in the previous experiments, the tones were perceived to be longer following reaches to more distant targets. Unexpectedly, the opposite effect was observed in the No-Go condition, with short target locations associated with longer perceived tone durations. This reverse effect may reflect the inhibition of movement on these trials. Previous work has shown that withholding movements can bias perception away from associated spatial features (e.g., when planning a leftward movement, stimuli appearing during the preparatory period are biased to appear rightwards from their actual position; Kirsch & Kunde, 2014). Our reversed effect in the No-Go condition could reflect an analogous effect on temporal features, where inhibiting movements that would have a particular duration biases perception away from intervals associated with those actions. We also note that the bias observed on Go trials in Experiment 9 had a reduced effect size, relative to that observed in Experiments 6-8. Speculatively, this could reflect residual inhibition from No-Go trials.

In line with Experiments 6-8, the results of Experiment 9 therefore indicate that the duration of an executed movement influences the perceived duration of concurrent auditory events. Importantly, the fact that this bias was absent when participants observed targets but were not required to move underscores that the effect is driven by the properties of executed movements rather than spatial properties of the target.

6.6. General discussion

Experiments 6-9 demonstrate that auditory duration judgements are strongly biased by the duration of concurrent executed movements. In Experiments 6 and 7, participants performed two concurrent temporal tasks where they were required to regulate the duration (Experiment 6) or speed (Experiment 7) of a movement. These experiments revealed a strong biasing effect of the executed movement on perceptual duration judgements. Explicit temporal requirements were eliminated from the movement task in Experiments 8 and 9, with a manipulation of movement amplitude instead used to elicit duration differences. The effect of action duration on perceived duration persisted, albeit attenuated in comparison to Experiments 6 and 7. Taken together, the results demonstrate a novel influence of action on duration perception.

The action-induced bias describes in these experiments joins an existing canon of empirical effects where action has been shown to influence the perceived timing of events (Merchant & Yarrow, 2016). For example, it is well established in the intentional binding paradigm that a delayed sound is perceived as closer in time to a keypress when it reliably follows the action (Haggard et al., 2002). The present findings extend this literature by showing that temporal features of action influence the perceived duration of sensory events in an isomorphic fashion – such that sensory durations are perceived to be more similar to executed actions. This isomorphic feature of the effect is more consistent with dedicated timing models which assume that temporal information is represented in explicit metrics in a manner that could support integration across modalities (Ivry, 1996) and are harder to reconcile with the idea that timing across action and perception depends on implicit representations that are encapsulated from each other (Buonomano, 2000).

The results of Experiments 6-9 are reminiscent of predictive effects of action in other perceptual domains, such as reports that perception of visual motion (Wohlschläger, 2000) and action identity (Di Pace & Saracini, 2014) are biased in line with concurrent actions. As such, the temporal bias identified in these experiments is consistent with the idea that actors incorporate expectations generated on the basis of movement into their perceptual estimates, in line with normative accounts of prediction found in sensory cognition (Yuille & Kersten, 2006) and the suggestion that expectations bias sensory populations towards expected outcomes (Summerfield & de Lange, 2014). For example, when participants produce longer actions, predictions may act to increase the gain on units associated with longer durations, biasing the sensory population response in line with expectations. Importantly, while such a mechanism induces perceptual distortions when actions and outcomes are uncorrelated – as in this experimental setting – such predictive biasing may be adaptive in our typical sensory environment where the temporal features of actions and outcomes are highly correlated, and where perceptual shifts toward expected durations generate more veridical percepts.

The idea that an adaptive integration mechanism can sometimes produce biases is reminiscent of findings in the multisensory perception literature where participants integrate signals across different modalities. For example, judgements about the spatial location of an auditory stimulus are often biased towards concurrent visual signals (Alais & Burr, 2004). Interestingly, such integration across modalities has been shown to depend on a 'causal inference' that the two sensory signals have arisen from the same source. For example, Körding et al. (2007) found that increasing the spatial disparity between auditory and visual signals reduced the influence that visual information has on auditory localisation judgements, and this effect could be accounted for by a model that used the spatial disparity between events to infer the likelihood both signals arose from a common cause.

It may be interesting to consider whether the effects reported in this chapter rely on an analogous inference that sensory effects (tones) are indeed caused by the executed actions. A strong version of this account is perhaps unlikely given the patterns across the experiments. Specifically, discrepancies between actions and stimulus durations (which may prevent inferences of a common cause) were largest in Experiments 6 and 7, but these conditions also yielded the largest action-induced biases. Nonetheless, it would be interesting to examine in future work how the predictive integration of information across sensory and motor channels may depend on beliefs about the causal structure of the sensory environment.

6.7. Chapter summary

To date the literature on sensory prediction has largely neglected influences on temporal processing, with dedicated and intrinsic models making divergent predictions about the possibility of predictive interactions. The results of Experiments 6-9 suggest that duration perception is biased toward the duration of executed actions, in line with dedicated timing models and consistent with the idea that expectations generated during action are incorporated into duration percepts.

Chapter 7 – Are predictive influences of action on duration perception biologically-tuned?

Chapter 6 reported a series of experiments that found auditory duration judgements were influenced by the temporal properties of concurrent actions. This chapter presents a preliminary investigation into the nature of the mechanism generating this effect. An influential hypothesis from Schubotz and colleagues proposes that the motor system influences perception of the physical environment using predictive mechanisms that are primarily adapted for processing biological action outcomes. An important prediction of this account is that influences of action on perception should be stronger for biological action outcomes than 'inanimate' nonbiological events. This hypothesis is considered in Experiment 10. Here participants completed a task (adapted from Experiment 7) where participants produced mouth movements of short and long durations, and made judgements about the duration of concurrently presented sounds. In separate blocks, participants either rated natural speech sounds (a biological event) or artificial triangle waves (a nonbiological event). Results replicated the action-induced bias described in Chapter 6, with participants rating sounds to be longer when they produced longer duration movements. However, contrary to the predictions of Schubotz and colleagues, no biological tuning was observed – with participants showing equivalent biases when judging speech and triangle waves. The results are more consistent with the idea that the motor system generates general predictions about features of the sensory environment that influence perceptual processing of both action and non-action events.

7.1. Introduction

Experiments 6-9 reported evidence that auditory duration judgements were strongly influenced by concurrently executed actions, such that tones were perceived to be longer when participants executed longer movements. Such findings are consistent with the idea that observers incorporate top-down expectations generated on the basis of action into their perceptual estimates, which may be adaptive given that expected events are by definition more likely to occur (Yuille & Kersten, 2006). However, while there are indeed strong correlations between temporal features of executed actions and the perceptual consequences which ensue, in Experiments 6-9 participants performed actions with their fingers or hands while judging pure tones. Given that it is unlikely that participants have much experience of producing actions and experiencing pure tones of similar durations, how do these effects arise?

A candidate answer is offered by Schubotz and colleagues (Schubotz, Kalinich & von Cramon, 2004; Schubotz, 2007) who suggest that the motor system is able to generate predictions about a variety of events in the external world using models that are primarily tuned to biological action outcomes. In particular, it is assumed that when we are required to generate expectations about inanimate (non-action) events in the environment we repurpose models that are used to generate predictions about perceptually similar action events e.g. we may use predictive models about the visual consequences of arm motion if we are required to passively predict the trajectory of a crashing wave, given that biological motion somewhat resembles the motion of the wave.

Evidence that the motor system is engaged during the passive prediction of inanimate sensory events is provided by neuroimaging studies. For example, fMRI studies reveal that when participants perform a task that requires predicting spatial, object (e.g. shape) or temporal patterns in a sequence of geometric shapes (in the absence of any requirement to move), specific regions of the premotor cortex are activated as a function of sequence dimension (Schubotz, Kalinich & von Cramon, 2004; Schubotz, 2007). For example, regions of the dorsal premotor cortex are activated for spatial sequences, superior ventral premotor cortex for object sequences and inferior ventral regions for temporal sequences (Schubotz, von Cramon, & Lohmann, 2003; Wolfensteller, Schubotz, & von Cramon, 2007). This pattern of data is interpreted as consistent with Schubotz's model given that motor representations involved in the execution of arm, hand and vocal movements lie in dorsal, superior ventral and inferior ventral premotor regions (Schubotz & von Cramon, 2002), respectively, and the perceptual consequences of our actions principally vary in space, shape and time. The model provided by Schubotz could account for the findings of Experiments 6-9, if it is assumed that predictive models used to anticipate typical biological consequences of our movements are used to generate predictions about the inanimate stimuli used in these paradigms (pure tones).

However, an alternative possibility is that agents possess predictive models of how their actions influence general dimensions of the sensory environment (e.g. event duration) in addition to those models used to anticipate biological action outcomes, and it is these more general predictive mechanisms that are responsible for action-induced biases when perceiving inanimate stimuli. Press and Cook (2015) have noted that the motor system contains representations of actions at different levels of abstraction (e.g. from movement position and duration in the primary motor cortex to more complex representations of effector configuration in the premotor cortex). Given that it is often assumed that sensorimotor predictions about animate action outcomes (e.g. observing hand grasps after executing hand grasps) reflect interactions between motor and sensory representation at similar levels of abstraction (e.g. premotor areas coding for configural features of grasping movements sending predictions to regions of superior temporal sulcus involved in the perception of grasping movements; Kilner et al., 2007), it is therefore possible that both animate and inanimate events are influenced by sensorimotor predictions at these more general levels.

More generic sensorimotor predictions about temporal information could be acquired through statistical learning about the relationships between executed actions and their typical animate and inanimate consequences. As shown in Figure 7.1, early in development, agents may possess separate representations of sensory and motor time that are weakly and non-systemically connected. However, given the natural correlation between the duration of executed actions and the perceptual consequences that are produced, we may acquire predictive associations between particular motor units and sensory units. Importantly, both animate (e.g. observed biological motion) and inanimate action outcomes (e.g. the duration of played musical notes) exhibit correlations with movement duration. It is precisely under these conditions, where the duration of biological and non-biological action outcomes are correlated similarly with action duration, that learning theorists predict the acquisition of predictions that generalise across biological and non-biological inputs (Pearce, 1987). Once such predictions have been acquired, activation of the motor unit through action execution biases its associated sensory units. Such associative activation of expected sensory units through action as described in Figure 7.1 would generate the effect described in Chapter 6 – namely that duration perception will be biased towards the duration of executed actions.

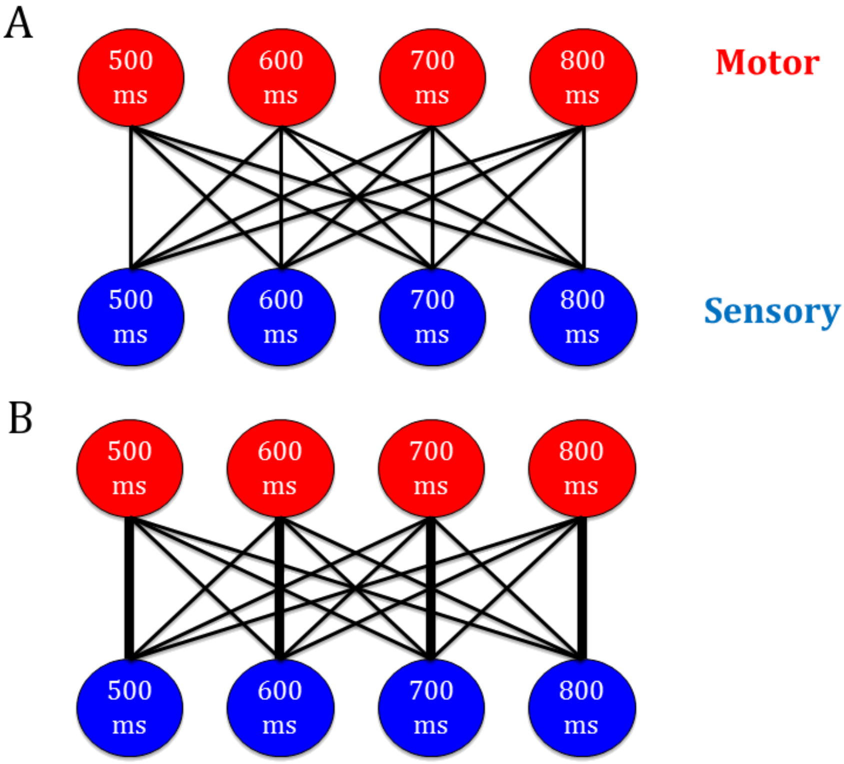

Figure 7.1: A schematic illustration of how sensorimotor associations may be forged through statistical learning. A) Early in life, we may have representations of time in both motor and sensory domains, with no systematic connections between them. B) However after experiencing the natural statistics of the environment – where the duration of an action is typically similar to the duration of the perceptual consequences it produces – we have the opportunity to learn 'matching' associations between motor and sensory duration representations (indicated by thicker connecting lines). These predictive associations would subsequently cause motor activity (e.g. during action execution) to propagate to sensory units, biasing perceptual activity toward similar durations.

A key prediction of Schubotz's 'repurposing' model is that the degree to which a sensory event benefits from motor prediction depends on the degree to which it resembles a biological action outcome (Schubotz, 2007), and that influences on perceptual processing are therefore strongest for biological outcomes relative to inanimate events. However, if agents possess more generic sensorimotor models about stimulus duration as a consequence of statistical learning as outlined above, then similar influences of action should be seen for the perception of biological and inanimate events. Experiment 10 was conducted to explore these contrasting predictions, investigating whether the effects of action execution on duration perception identified in Chapter 6 reflect the operation of a predictive mechanism primarily adapted to biological action outcomes. The paradigm used was modelled on that of Experiment 7 with two principal changes. The first change was that participants produced 'fast' or 'slow' movements with their mouths rather than fingers – as Schubotz and colleagues suggest premotor representations of the articulatory system are most strongly involved in generating temporal predictions (Schubotz & von Cramon, 2002). (NB: While adopting this explicit movement paradigm has some potential methodological concerns that are allayed by the implicit motion tracking task used in Experiments 8 and 9 [e.g. possible linguistic contributions of terms 'fast' and 'slow' to the observed effects, see Section 6.3], practical difficulties associated with tracking mouth movements and manipulating their duration implicitly through a secondary task made incorporating these elements into Experiment 10 infeasible.) The second change was that participants no longer made judgements about pure tones, but instead made judgements about the duration of speech sounds or triangle waves. The selection of these two stimulus types permits assessment of whether predictive effects on perception are larger for biological outcomes (speech) or non-biological outcomes (non-speech), given that triangle waves cannot be produced by biological systems.

7.2. Experiment 10

7.2.1. Method

Participants

Sixteen participants (12 female, mean age = 23.9 years, SD = 4.2) were recruited from Birkbeck, University of London and paid a small honorarium. Five of these acted as replacements for excluded participants where performance was highly variable, yielding psychometric functions that could not be modelled effectively in at least one condition (see below).

Procedure and Stimuli

The procedure of Experiment 10 was identical to Experiment 7 with the following changes (see Fig. 7.2). Participants began each trial with their lips depressing a custom-made response key (Heijo Research Electronics) which was mounted on an adjustable stand at the participant's mouth height. To produce movements of different target durations (<500 ms or >950 ms) participants released and then depressed the key for the target interval – opening their mouths as a result. In the main experiment the key release triggered the presentation of a target duration stimulus that was drawn from a seven-point psychophysical continuum ranging from 500 to 950 ms in duration (75 ms steps).

In Biological blocks, the test stimulus was the open back unrounded vowel (/a/) produced by a male native English speaker (frequency = 117 Hz). This vowel sound is that typically produced by vocalisations with an open mouth, unrounded lips and without the involvement of the tongue i.e. the articulatory action performed by participants when releasing the key. The test continuum was created by taking an original recording of the spoken vowel sound and manipulating the duration contour of this sound in the speech analysis software Praat (Boersma, 2001). This manipulation alters the duration of the sound but leaves its pitch unaffected. In Nonbiological blocks, the test stimulus was a triangle wave produced using the sound-editing software Audacity (frequency = 100 Hz). Following a variable delay ISI (1000-1500 ms) after participant completed their mouth movement, participants were presented with a reference stimulus with a duration of 725 ms (midpoint of the test continuum). Participants then judged whether the first or second sound was longer in duration.

Participants completed 280 trials, split into two blocks with Biological or Nonbiological stimuli (140 each, order counterbalanced across participants). Psychometric functions were modelled to participant responses as in Experiment 7, separately for each combination of movement and stimulus condition. As a manipulation check, at the end of the experiment participants completed computerised scales where they were presented with a test stimulus from the Biological and Nonbiological condition (order counterbalanced) and were asked to clicking on a rating scale from 1 to10 (‘Not at all’ to ‘Very Much’) ‘How much was that sound like human speech?’ and ‘How much was that sound like your speech?’.

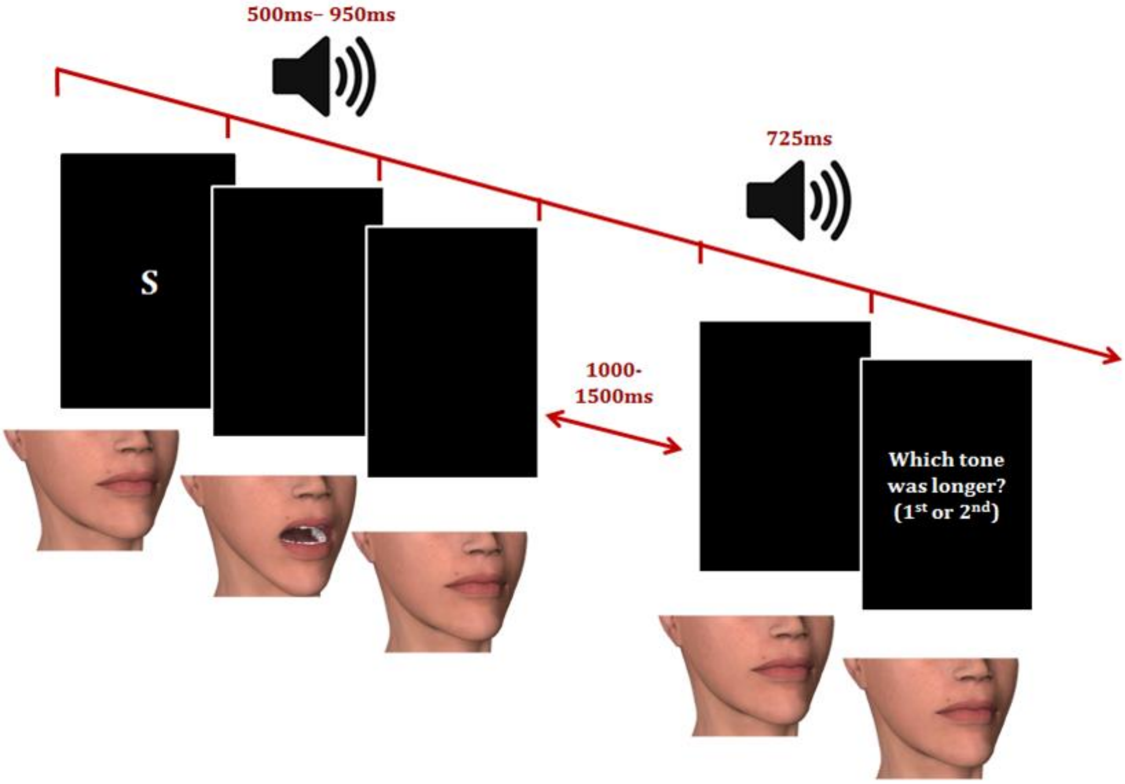

Figure 7.2: Time course of stimulus presentation and participant movements on a single test trial in Experiment 10.

7.2.2. Results

Human-like and self-like ratings for Biological (speech) and Nonbiological (triangle wave) stimuli were compared using repeated measures t-tests. Both tests revealed that participants rated the Biological stimuli as more like human speech (mean rating = 8.74, SEM = .265) than the Nonbiological stimuli (mean rating = .289, SEM = .12424, $t(15) = 28.72$, $p<.001$) and that the Biological speech stimuli also sounded more like their own speech (mean rating = 5.24, SEM = .704) than the Nonbiological stimuli (mean =.234, SEM= .086, $t(15) = 7.11$, $p<.001$). These results suggest that participants perceive robust differences between the identity of the stimuli, and importantly that processing steps used to manipulate Biological stimuli did not cause them to appear unnatural.

Participants were successful in generating movements of different durations in the two conditions. The mean number of movement errors was 13.6 % (SEM = 1.6%). The mean duration of correct movement was 234.9 ms (SD = 109.9 ms) on fast movement trials and 1591.4 ms (SD = 333.9 ms) on slow movement trials. As illustrated in Figure 7.3., comparable differences between fast and slow movement times were obtained in both stimulus blocks.

Precision and PSE values were analysed in separate repeated measures ANOVAs with factors of Movement (Fast, Slow) and Stimulus (Biological, Nonbiological). The precision analyses revealed no significant effects (all $p>.104$). However, the PSE analysis revealed a significant main effect of Movement – $F(1, 15) = 19.201$, $p=.001$, $\eta_p^2= .561$) – which reflected the fact that PSEs were significantly lower (i.e. tones were judged to be longer) in the Slow movement condition (mean = 687.6 ms , SEM = 19.28) relative to the Fast movement condition (mean = 787.6 ms, SEM = 12.6), replicating the results of Chapter 6 . No significant main effect of Stimulus ($p=.261$) or interaction between Stimulus and Movement ($p=.131$) was found.

To investigate whether the lack of interaction reflected support for the null hypothesis we compared the magnitude of the PSE biases in each stimulus condition using a Bayesian t-test in JASP with an identical factorial structure. This test evaluated a one-tailed hypothesis (i.e. biological PSE effect> nonbiological PSE effect) given the directional prediction of Schubotz's model that predictive influences on perception should be more pronounced for biological than nonbiological stimuli. This analysis revealed a $BF_{+0} = 0.113$, suggesting ~8.9 times more evidence for the null hypothesis that PSE effects between stimulus conditions do not differ.

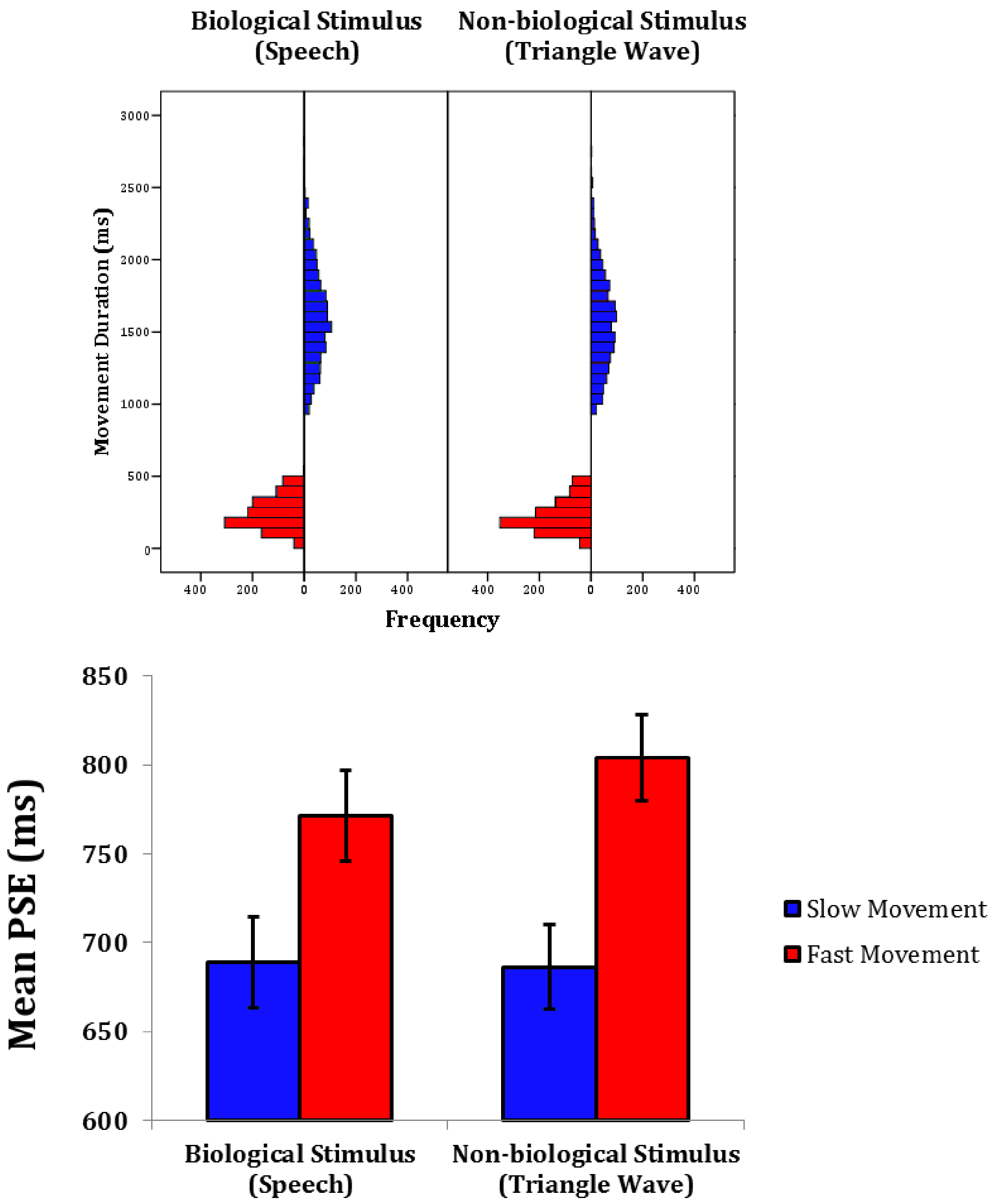

Figure 7.3: Top Panel: Histograms showing movement times in the across all conditions of Experiment 10 collapsed across all subjects. Bottom Panel: Mean PSEs across all conditions of Experiment 10. Lower PSEs indicate a tendency to the rate the test stimulus as *longer* relative to the comparison. Error-bars display 95% within-participant confidence intervals (Cousineau, 2005).

7.2.3. Discussion

Experiment 10 replicates and extend the results of Experiment 6-9, showing that the perceived duration of auditory events is biased towards the duration of executed mouth movements, such that sounds were judged to be longer when movements were longer in duration. Experiment 10 also found convincing evidence against a Movement x Stimulus interaction, suggesting that action-induced biases were equivalent when participants judged biological speech sounds and nonbiological triangle waves. This contrasts with models which suggest that predictive effects of action on perception arise through mechanisms that are primarily tuned to biological action outcomes (Schubotz, 2007), but is consistent with the possibility that agents possess generic sensorimotor models concerning temporal information.

A critic of this interpretation could query whether this null result reflects an absence of biological tuning, or a failure to effectively generate apparently biological and nonbiological stimuli. Researchers have suggested that explicit beliefs about the animacy of an observed stimulus (i.e. whether it was an action produced by an intentional agent) modulate the degree to which such events recruit motor processes. For example, the same dot motion stimulus interferes less with movement execution when participants believe it to be generated by a computer rather than a human (Stanley, Gowen, & Miall, 2007) and participants show reduced imitation of identical hand actions they believe to be non-human in origin (Liepelt & Brass, 2010; though see Press, Gillmeister, & Heyes, 2006). As such, if speech and non-speech sounds were *both* judged to be 'inanimate' in origin (e.g. given the digital manipulation of speech sounds) any biological tuning effects may not be observed. However, given that participants robustly judged the speech sounds as being of human origin and the non-speech sounds as non-human, this alternative explanation is not compelling. It is also worth noting that a recent meta-analysis has suggested there is little evidence for effects of animacy beliefs on laboratory imitation tasks (Cracco et al., in press), which may bear on the present findings insofar as imitation and sensorimotor prediction reflect related mechanisms (Kilner et al., 2007;Wolpert et al., 2003).

These results are more consistent with the idea that the motor system generates top-down predictions about general sensory features (Press & Cook, 2015) that equivalently influence perception of animate and inanimate events. For example, agents may generate sensory predictions about these more generic perceptual dimensions, with a granularity (e.g. 'long duration') that allows such predictions to apply equally to biological and nonbiological action outcomes. Under this account it is not necessary to assume that one has had specific experience of experiencing (e.g.) tones or triangle waves when producing movements to explain these perceptual effects. Rather, it is assumed that the temporal predictions one acquires through sensorimotor experience with one's own body and physical environment are forged at a level of abstraction that allows them to generalise across input domains (Pearce, 1987).

7.3. Chapter Summary

Experiment 10 found equivalent biases of concurrent action duration on the perceived duration of biological (speech) and nonbiological (triangle wave) stimuli. These results are difficult to reconcile with the idea that predictions about inanimate features of the environment arise by repurposing mechanisms primarily adapted to biological outcomes, but are consistent with the idea that agents may acquire generic sensorimotor predictions as a consequence of statistical learning.

Chapter 8 – Do predictive influences of action on duration perception arise through statistical learning?

Chapter 6 found that auditory duration judgements were biased towards the duration of concurrent executed actions – consistent with a predictive influence of action on perception. Chapter 7 investigated the possibility that these effects are generated by mechanisms primarily tuned to biological action outcomes, but found no evidence for biological tuning. Results were instead consistent with an alternative explanation: that action influences duration perception through domain-general sensorimotor predictions acquired on the basis of statistical learning. Chapter 8 presents a preliminary investigation into this statistical learning by interrogating one of its key predictions:that the influences of action on perception should depend on the experienced contingencies between actions and outcomes, and should be recalibrated when these contingencies change. This prediction is investigated in Experiment 11 by examining whether altering contingencies between executed actions and perceptual outcomes recalibrates action-induced perceptual biases. Participants completed a task modelled on the paradigm used in Experiment 9 over four testing sessions. Over the two middle sessions participants were exposed to a sensorimotor contingency that either reflected the natural statistics of the environment (e.g. where long duration movements were paired with long duration outcomes – the Congruent training group) or where the natural mapping between actions and outcomes was reversed (e.g. where long duration movements were paired with short duration outcomes – the Incongruent training group). Pre- to post-test comparisons revealed that incongruent, but not congruent, sensorimotor experience abolishes the action-induced bias, consistent with the idea that the underlying mechanism tracks the statistical relationships between executed actions and perceptual outcomes.

8.1. Introduction

Thus far this thesis has presented a novel temporal influence of action on perception, where the perceived duration of an auditory event was biased toward the duration of executed actions (Chapter 6). Chapter 7 investigated whether this effect was generated by generalising models tuned to biological action outcomes to the prediction of inanimate events – finding negative results. One alternative possibility entertained was that agents generate 'domain-general' predictions about the relationship between actions and outcomes, and these sensorimotor predictions generalise across biological and nonbiological stimulus domains (Pearce, 1987). This possibility is investigated in this chapter.

A venerable tradition in experimental psychology has emphasised that action control depends on learning the relationship between performed actions and perceived outcomes by tracking the statistics of the environment (James, 1890). For example, research in the ideomotor literature has found that experiencing contingencies between actions (e.g. a particular key press) and outcomes (e.g. a pure tone of a particular frequency) can forge sensorimotor associations - such that actual (Elsner & Hommel, 2001) and anticipated action outcomes (Kunde, Koch, & Hoffmann, 2004) can prime their associated actions in reaction time tasks. It has been proposed that sensorimotor learning generates 'automatic imitation' effects, whereby participants are faster to execute hand actions when observing the same action executed by task-irrelevant hand (Brass, Bekkering, & Prinz, 2001; Heyes, 2011). Brass and Heyes (2005) propose that such effects reflect the fact that we have vast amounts of sensorimotor experience that allows us to develop statistical associations between 'doing' and 'seeing'. For example, when we send the motor command to open a hand there is a high probability that we will observe an opening hand (and amongst all observed actions, the congruent action is conditionally most probable). These associations acquired in the natural environment therefore bias action selection in the same fashion as the 'arbitrary' associations acquired during laboratory training (Elsner & Hommel, 2001).

A powerful way to test statistical learning accounts is to investigate whether sensorimotor interactions are altered when the statistical relationship between actions and outcomes is manipulated. For example, in a series of experiments Heyes and colleagues (Cook, Bird, Catmur, Press, & Heyes, 2014) have shown the automatic imitation effect can be abolished after participants are given incongruent sensorimotor experience of 'counter-imitation' - where participants execute an action (e.g. hand opening) in the presence of an incongruent observed action (e.g. hand closing, (Heyes, Bird, Johnson, & Haggard, 2005). Similar effects are seen with neuroimaging measures, where incongruent sensorimotor experience of producing hand/foot movements when observing foot/hand movements has been shown to reverse the pattern of 'mirror' responses in the premotor cortex evoked during passive action observation (Catmur et al., 2008). Furthermore, similar incongruent experience of observing index/little finger abduction movements while executing little/index movements reverses the pattern of muscle-evoked potentials elicited through TMS during action observation (Catmur, Walsh, & Heyes, 2007).

A similar form of statistical learning could furnish the mechanisms generating the action-induced perceptual biases described here (see Section 7.1.) An important feature of this statistical learning account is the assumption that action biases perception toward *similar* perceived durations solely because actors experience a strong correlation between action durations and perceptual durations that are similar. Nothing, in principle, prevents the acquisition of ‘non-matching’ associations, where actors learn to expect that short actions can be followed by long durations, or vice versa, other than statistics of the environment. As illustrated in Figure 8.1.c, the statistical learning account would predict that an individual who receives this kind of incongruent sensorimotor experience should acquire additional non-matching associations. As a result, one would predict that after this kind of incongruent experience action execution (e.g. production of a short action) would activate competing sensory units (e.g. the first-learned ‘short’ sensory units, alongside the second-learned ‘long’ sensory code). This kind of learning would therefore be expected to abolish action-induced perceptual biases, as increases in sensory activation will be symmetric across the sensory population. In particular, this kind of second-learned association would not be expected to extinguish the first-learned association, given that extinction of learning operates over vast time scales (if it occurs at all; Baeyens, Eelen, & Crombez, 1995). Relatedly, it has been shown that second-learned associations acquired in sensorimotor tasks demonstrate specificity to the contexts in which they are trained (Cook, Dickinson, & Heyes, 2012).

Experiment 11 was conducted to investigate this prediction. Participants completed the simultaneous reaching and duration judgement task described in Experiment 8 in four sessions spread over two consecutive days. During the first and last sessions participants completed an identical task to that described in Experiment 8. In the middle sessions (second and third) participants received sensorimotor training. One group of participants received congruent sensorimotor experience, such that when they reached towards near and far targets (producing shorter / longer movements) they heard shorter and longer tones, respectively. A second group of participants received incongruent sensorimotor experience, such that when they reached towards near and far targets (producing shorter / longer movements) they heard longer and shorter tones, respectively. Under a statistical learning account, participants should show attenuated action-induced perceptual biases after incongruent experience, while biases in participants receiving congruent experience should remain intact.

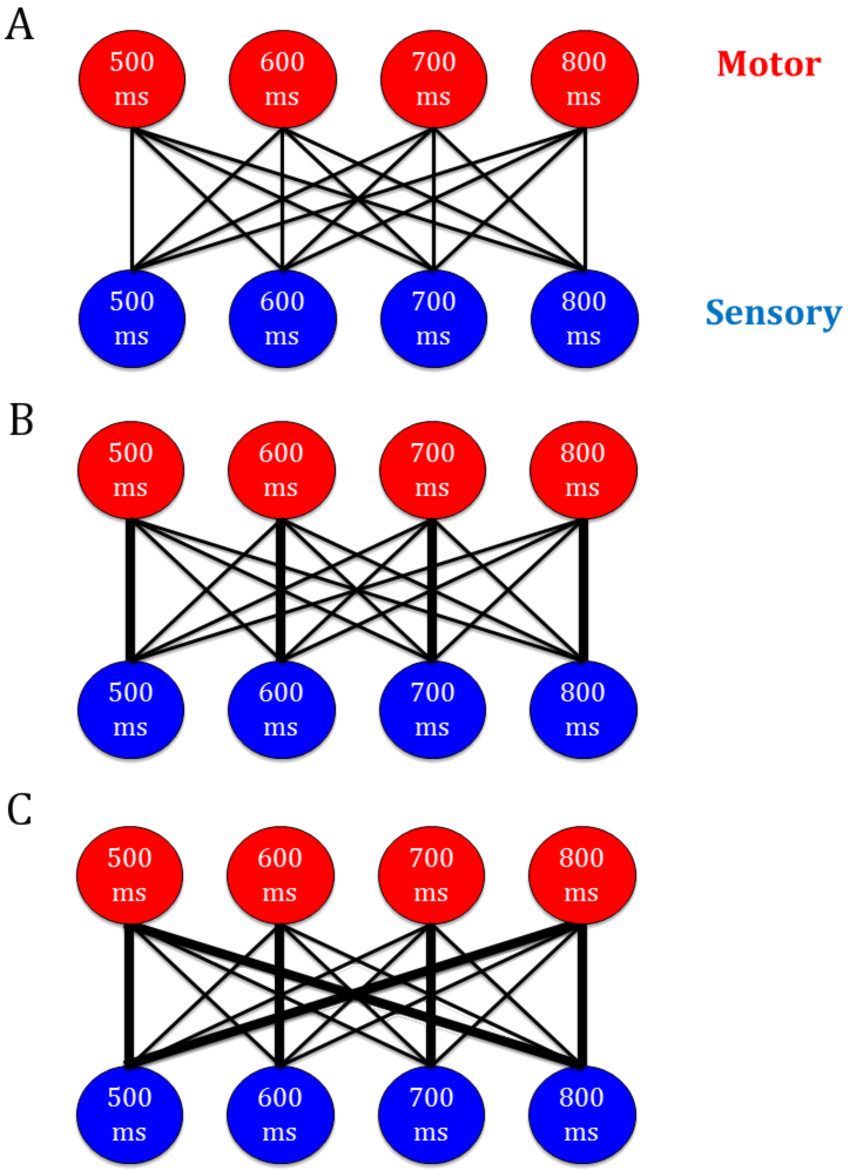

Figure 8.1: A schematic illustration of how sensorimotor associations may be forged through statistical learning. Panels A and B are identical to Figure 7.1, indicating the putative associations acquired by an agent before and after experiment typical sensorimotor contingencies (e.g. longer actions = longer sensory consequences). Panel C illustrates that if an individual subsequently experiences an incongruent sensorimotor contingency (e.g. here where 500 ms actions produce 800 ms effects, and vice versa), a statistical learning account would predict the acquisition of addition non-matching associations that compete with those 'first-acquired' associations. It is this kind of experience that is provided in Experiment 11's Incongruent condition.

8.2. Experiment 11

8.2.1. Method

Participants

Thirty-two participants (24 female, mean age = 24.8 years, SD = 5.5) were recruited from Birkbeck, University of London and paid a small honorarium. Four of these acted as replacements for excluded participants where performance was highly variable, yielding psychometric functions that could not be modelled effectively in at least one condition. Sixteen participants each were randomly allocated to the Congruent and Incongruent training groups.

Procedure and Stimuli

Participants completed four sessions of the paradigm described in Experiment 8, split over two consecutive days. The first and last of these sessions respectively formed the Pre-Test and Post-Test blocks, were identical for both training groups and had an identical procedure to that described in Experiment 8.

The second and third sessions formed the Training blocks. These were identical to Experiment 8 apart from the following change. In the Incongruent training group, Near Target trials were always accompanied by the longest test tone (950 ms) and Far Target trials were always accompanied by the shortest test tone (500 ms). These sensorimotor contingencies are inconsistent with those typically experienced in the natural environment, and provide the opportunity to learn competing sensorimotor predictions. In the Congruent training group, Near Target trials were always accompanied by the shortest test tone (500 ms) and Far Target trials were always accompanied by the longest test tone (950 ms). This sensorimotor contingency roughly corresponds to the natural environment, where short/long actions are associated with short/long sensory outcomes and acted as a control condition. Both conditions also had 16.7 % of Intermediate Target trials as in Experiment 8, and in both groups these targets were accompanied by the midpoint of the tone continuum (725 ms). Participants completed the same task as that undertaken during Pre and Post-Test – judging which of the two presented tones was longer in duration.

Psychometric functions were modelled to participant responses in the Pre-Test and Post-Test sessions as in Experiment 8 to establish whether the effect was replicated in both training groups before training was administered, and to investigate whether the nature of the perceptual bias changed between groups.

8.2.2. Results

Precision and PSE values were analysed using two separate 2 x 2 x 2 mixed ANOVAs with the between-participants factor of Training Group (Congruent, Incongruent) and the within-participants factors of Time (Pre-Test, Post-Test) and Movement Condition (Near Target, Far Target). While no significant effects were found in the precision analysis, the PSE analysis revealed a significant main effect of Movement Condition– $F(1, 30) = 73.6$, $p<.001$, $\eta_p^2 = .710$ – alongside a significant two-way interaction between Movement and Time - $F(1, 30) = 8.522$, $p=.007$, $\eta_p^2 = .221$ – and a significant three-way interaction between

Training Group, Time and Movement Condition - $F(1,30) = 7.85$, $p=.009$, $\eta_p^2 = .207$.

The main effect of Movement Condition reflected the fact that across all conditions PSEs were lower when participants produced movements to Far targets (mean = 664.5 ms, SEM = 9.8) compared to Near targets (mean = 727.2 ms, SEM = 10.1), replicating the finding that participants are biased to judge tones to be longer when they produce movements with longer durations.

To break down the obtained interactions separate 2 x 2 repeated measures ANOVAs with factors of Time and Movement condition were conducted for each Training Group. In the Congruent training group no interaction between Time and Movement Condition was observed ($p=.929$), while in the Incongruent training group this interaction was significant - $F(1,15) = 14.25$, $p=.002$, $\eta_p^2= .487$. This interaction in the Incongruent group reflected the fact that while pre-training PSEs were lower in the Far target condition than the Near target condition – $t(15) = 7.37$, $p<.001$ – this difference was markedly attenuated post-training, with PSEs which no longer differing significantly between conditions ($p=.078$; see Figure 8.2).

This pattern of results is consistent with the idea that incongruent sensorimotor experience can recalibrate predictive relationships between motor and sensory representations of time, such that second-learned associations can abolish action-induced perceptual biases.

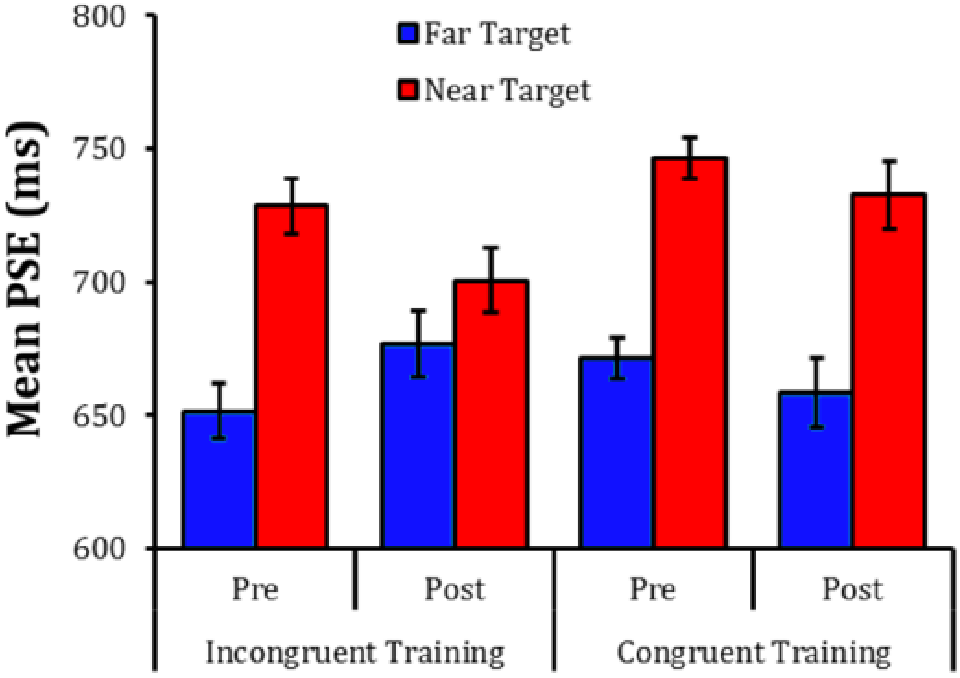

Figure 8. 2: Mean PSEs across all conditions of Experiment 11. Lower PSEs indicate a tendency to rate the test stimulus as *longer* relative to the comparison. Error-bars display 95% within-participant confidence intervals (Cousineau, 2005).

8.2.3. Discussion

Experiment 11 found that a short period of incongruent sensorimotor experience – where participants produced long actions and perceived stimuli with a short duration (and vice versa) – was sufficient to abolish action-induced biases on duration perception. Perceptual biases did not decrease in magnitude after congruent sensorimotor experience; suggesting that this effect is unlikely to reflect generic effects of practice with the task. Instead, these results are consistent with the idea that the mechanism responsible for the perceptual bias tracks the statistical relationship between actions and outcomes (Brass & Heyes, 2005; Hommel et al., 2001). This result is particularly important, as it suggests that the underlying mechanism is *predictive* in the psychological sense – generating influences based on what is conditionally probable (Shea, 2015; see Section 1.1.).

One potential limitation in the design of Experiment 11 is that the training manipulation introduces confounded statistical associations between perceived durations, action durations and target distances. For example, during Incongruent training participants were consistently exposed to Far targets (circles at large eccentricities) and short test tones (500 ms). While this provides the opportunity to learn new statistical relationships between action durations and tone durations (e.g. long movement, short tone), it is also possible that participants learn the relationship between target locations and tone durations (e.g. far target, short tone), and that these purely sensory expectations induce competing perceptual biases at post-test. The possibility that purely sensory learning accounts for the effect of training in the Incongruent group may seem unlikely given the pattern of results seen in the Congruent group. In particular, this form of sensory learning should also occur during congruent training, and should lead to the acquisition of sensory expectations that *exaggerate* (rather than compete with) the perceptual bias in this group. Evidence for this was not obtained given that perceptual biases after congruent training were not significantly increased. Nonetheless, this remains an important issue for future research, and could be resolved in a paradigm where participants undergo incongruent sensorimotor experience but do not produce actions at post-test (cf. No-Go condition, Experiment 9). If acquired sensory predictions drive the training effect it should still be obtained under these conditions, while any effect due to acquired sensorimotor predictions would be absent.

The recalibration of predictive motor influences on perception after incongruent sensorimotor experience is reminiscent of experiments where the perceptual influences on action selection and cortical motor activity are abolished (Heyes et al., 2005) or reversed (Catmur et al., 2007; 2008) after incompatible training. Evidence that sensorimotor experience administered in laboratory experiments is sufficient to recalibrate these kinds of sensorimotor effects has been taken to support the position that sensorimotor experience 'in the wild' plays an instructive role in the initial development of the underlying mechanisms (Cook et al., 2014). The same reasoning can be applied to Experiment 11, suggesting that the predictive effects of action on duration perception depend on an initial stage of statistical learning (see Figure 8.1).

A critic of this line of reasoning may suggest the fact that incongruent sensorimotor experience in the lab alters an effect does not imply that sensorimotor experience was responsible for establishing the mechanism in the first place – for example, matching connections between motor and visual representations could be genetically pre-specified as a result of an inherited adaptation (Gallese, Rochat, Cossu, & Sinigaglia, 2009; or in other words, that the neonate enters the world with a system more like Figure 8.1.b rather than 8.1.a). However, such 'genetic' alternatives may not be compelling given that evolutionary biologists have long stressed the importance of buffering inherited adaptations against environmental changes – if a behavioural trait can be altered by changes in the environment, any advantage conveyed by such a trait is unreliable and the relevant trait is unlikely to be retained through natural selection (Waddington, 1957). Indeed, the fact that the short period of experience administered in Experiment 11 was sufficient to disrupt the action-induced biases suggests that the underlying mechanisms are not buffered against environmental influences as is typical for inherited adaptations.

8.3. Chapter Summary

Experiment 11 presented evidence that incongruent sensorimotor experience can abolish action-induced biases on duration perception described in Chapters 6 and 7. These results are consistent with the idea that these influences arise through a predictive mechanism that tracks the statistical association between actions and outcomes.

Chapter 9 – General Discussion

9.1. Thesis summary

This thesis investigated sensory prediction mechanisms during action, asking three questions about the nature of these processes: First, do predictive processes during action have a functionally dissociable influence on perceptual processing, relative to other prediction mechanisms (functional-specificity)? Second, do motor predictions support our ability to process the actions of other agents (agent-specificity)? Third, do predictions generated during action influence the 'when' – as well as the 'what' - of perceptual estimates (modality-specificity)?

The most influential claims about the functional-specificity of sensory prediction during action are made by the Cancellation model (Wolpert et al., 1995; Bays & Wolpert, 2007) which proposes that perceptual processing of expected action outcomes is suppressed. This contrasts with contemporary thinking about sensory prediction outside of action contexts, where it is typically assume that perception is biased towards more likely outcomes, which will facilitate (not suppress) perception of expected events (Yuille & Kersten, 2006; Summerfield & de Lange, 2014). The novel OPPOSE model was outlined in Section 1.2.4, alongside its principal prediction that sensory predictions during action operate equivalently to those in other contexts. In line with this model, Chapters 2-4 presented evidence that participants rate expected action outcomes as more intense at early timepoints after action (Experiment 1) and that participants are more sensitive to and biased towards reporting the presence of congruent action outcomes when these are presented at threshold (Experiment 2). In addition, Experiment 3 found using multivariate analyses of fMRI data that congruent action outcomes are associated with higher fidelity representations in clusters across the visual cortical hierarchy. All of these results are difficult to reconcile with the Cancellation model, but are consistent with the predictions of the OPPOSE model. As such, these results undermine popular claims about the functional-specificity of sensory prediction during action, and instead suggest that expectations generated on the basis of movement influence perceptual processing in a similar way to expectations established on the basis of other kinds of probabilistic information (Summerfield & de Lange, 2014).

Claims concerning agent-specificity are found in social cognition theories that assume predictions generated to support perception of the direct consequences of our actions also support perception of imitative reactions performed by other agents (Schütz-Bosbach & Prinz, 2007). An important assumption of these theories is that sensorimotor predictions operate across suprasecond delays between action and outcome. Chapter 5 reported two experiments which revealed that congruent action outcomes were perceived to be more intense after substantial suprasecond delays. These results suggest that sensorimotor predictions do indeed operate with a temporal profile that could allow them to influence perceptual processing of others' reactions, leading to smoother and more fluent social interactions.

Modality-specificity was considered in Chapters 6-8 by investigating whether predictions about the temporal features of action outcomes bias duration perception. These experiments found that duration judgements were strongly influenced by the duration of concurrent actions, even when these differences in movement duration were generated implicitly. Subsequent chapters presented preliminary investigations into the nature of the underlying mechanism. Chapter 7 found evidence against biological tuning; suggesting it is unlikely that the perceptual biases arise due to a generalisation of sensorimotor models used for predicting naturalistic action outcomes (Schubotz, 2007). Chapter 8 reported evidence that the action-induced bias can be abolished after a short period of incongruent sensorimotor experience. This finding is more consistent with the idea that participants acquire general predictive models about the relationship between temporal features of their actions and the ensuing perceptual consequences through a process of statistical learning.

In summary, the results of all the experiments reported in this thesis highlight the domain-generality of sensory prediction mechanisms in action across a range of dimensions, and suggest that prediction operates similarly in action and in other contexts.

9.2. Disentangling perceptual and response biasing

The majority of the experiments reported in this thesis used psychophysical techniques to characterise perceptual biases induced by action. A perennial concern in the psychophysical literature is that influences of experimental manipulations on perceptual report may reflect response biases rather than changes in perceptual appearance (e.g. Carrasco et al., 2004; Schneider & Komlos, 2008). Indeed, in a recent paper Firestone and Scholl (2016) suggested that there was weak evidence for *any* top-down influences on visual perception, and that previously reported effects likely reflect changes in later decisional processes rather than changes in perception per se.

In describing the pitfalls of previous work Firestone and Scholl (2016) identify conditions where top-down influences are likely to induce response (rather than perceptual) biases. For example, they suggest responses biases are likely when participants have prior beliefs about how different response options relate to different experimental conditions. A number of experiments reported in this thesis (e.g. those reporting effects of expectation on perceived brightness – Chapters 2 and 5) are not vulnerable to this criticism given that participants are unlikely to have preconceived ideas about the relationship between the independent and dependent variables, and a number of experiments incorporated controls which made biases to select one response alternative orthogonal to the measure of perceptual bias.

However, the issue of decisional bias is less well-controlled in experiments using signal detection measures (Experiment 2). This is because while measures of sensitivity in signal detection tasks (e.g. *d'*) are bias-free, top-down influences at the perceptual or decisional locus can both manifest as changes in measures of response criterion (e.g. *c*; Witt, Taylor, Sugovic, & Wixted, (2015). For example, Wyart et al., (2012) suggest that expectations generate an additive increase in activation at the sensory level – which should inflate hits and false alarm rates, leading to a liberal bias– but similar results would be obtained if expectations instead alter later decision thresholds in the absence of any perceptual effects (Swets et al., 1961). Wyart and colleagues (2012) address this issue by building a computational model of the sensory and decision stages of a putative signal detection process and showing that their results are better accommodated by a model which assumes expectations affect sensory activity rather than decision thresholds. Similar explicit attempts to model predictive effects in the context of action are therefore an important issue to address in future work.

Decisional biases could also in principle be a concern in the experiments which reported that duration judgements were biased toward the duration of executed actions (Chapters 6-8). However, such an account may be unlikely given that effects persist when movement duration is manipulated implicitly, and the effect is abolished after incongruent sensorimotor experience (Chapter 8). If producing (e.g.) longer actions biases participants to respond that a concurrent stimulus is (e.g.) longer in the absence of any changes at the perceptual level, such a mechanism would continue to generate effects after incongruent sensorimotor experience.

9.3. Are effects of action-outcome congruency effects of expectation?

The principal aim of this thesis was to investigate the nature of sensory prediction mechanisms during action, and to consider whether these effects are more consistent with the Cancellation model or with predictive models of perception described in the wider sensory cognition literature. This question was investigated by comparing perceptual performance and visual brain activity as a function of action-outcome congruency, where participants produced a particular action (e.g. an index finger movement) and observed either a congruent (e.g. index finger movement) or incongruent outcome (e.g. middle finger movement), with both kinds of events occurring with equal probability during the experimental task. The logic of these paradigms assumes that differences between congruency conditions reflect influences of expectation, given that participants will have strong expectations that ‘congruent’ sensory outcomes will occur after experiencing the statistics of the natural environment during development (e.g. where observed index finger movements are conditionally more probable after sending the motor command to lift an index finger). As noted in Section 1.2.2., congruency manipulations of this kind are the predominant way in which expectations are manipulated in the action control literature.

However, this congruency manipulation differs from approaches typically used to examine expectations in perceptual tasks by sensory cognition researchers. Here, it is more common for observers to be given sensory cues (e.g. tones with different frequencies) that provide information about the conditional probability of a particular stimulus across the duration of the experiment (e.g. gratings with particular orientations, or different directions of dot motion; Kok et al., 2012; 2013). It should be noted that this manipulation of within-experiment probabilities is perhaps the most logical in the sensory cognition literature given that it often examines the perception of arbitrary events (e.g., gratings) where there would be fewer predictive events prior to any training. However, given that probabilistic relationships between actions and outcomes are not manipulated in most of the paradigms used here, it is important to consider the validity of describing these effects of congruency as ‘predictive’.

The ideomotor tradition in action control research has emphasised that predictive knowledge about the consequences of our actions is acquired through statistical learning (James, 1890), and researchers in this tradition have drawn a theoretical equivalence between those action-outcome associations that are putatively acquired through typical development (e.g. that lifting a finger predicts the observation of a finger lift; Brass, Bekkering & Prinz, 2001) and those which are acquired through arbitrary learning procedures conceptually similar to those described in contemporary sensory cognition research (e.g. where moving a finger generates a high pitched tone; Elsner & Hommel, 2001). As such 'natural' and artificial congruency effects in reaction time tasks are treated as identical phenomena. Indeed, it has been shown that laboratory learning procedures can reconfigure 'natural' congruency effects observed in automatic imitation tasks (see Section 8.1) and that the learning mechanisms responsible for this reconfiguration are sensitive to the statistical contingency between actions and outcomes (Cook, Press, Dickinson, & Heyes, 2010) in much the same way that arbitrarily acquired action-outcome associations are contingency-sensitive (Elsner & Hommel, 2004). Additionally, in the present thesis Experiment 11 revealed that action-induced perceptual effects can be reconfigured by sensorimotor experience, suggesting that the underlying mechanisms track statistical relationships between actions and outcomes, a cardinal feature of prediction (Shea, 2015).

As such, it seems reasonable to assume that effects of action-outcome congruency do indeed reflect influences of a predictive mechanism. However, it may be advantageous in future work to investigate whether similar effects on perceptual processing are obtained in tasks where participants experience novel associations between actions and arbitrary outcomes in the lab, as these would provide strong controls against other forms of top-down influence that could influence perceptual processing (e.g. conceptual relationships between actions and outcomes; Casasanto & Boroditsky, 2008) and would permit even more informed comparison with tasks used in the sensory cognition literature (Summerfield & de Lange, 2014).

9.4. Relationships between attention and expectation

Summerfield and Egner (2009; 2016) have proposed that top-down effects of expectation (driven by stimulus probability) should be carefully distinguished from top-down effects of attention (driven by the relevance of stimulus to a participant's current task). Such dissociations are important given that both kinds of information may be redundant in many settings (e.g. things that are likely are also often relevant). Indeed, this conflation is found in classic 'attentional' paradigms – such as the Posner task (Posner, Snyder, & Davidson, 1980) –where participants are given probabilistic cues about target locations and manipulations are thought to reflect attentional processes. Similarly, Summerfield and Egner (2016) have proposed that experiments reporting detection benefits after valid probabilistic cues to stimulus identity (e.g. Stein & Peelen, 2015) could be explained by attentional – rather than predictive – weighting of cued features.

The experiments reported in this thesis chiefly manipulated expectations via action-outcome congruency (see Section 9.3 above), but largely did not explicitly manipulate task-relevance (with the exception of Experiment 3). However, across these paradigms congruent and incongruent stimuli were equally task-relevant, and similar effects of congruency on visual brain activity were obtained irrespective of task-relevance (Experiment 3). Nonetheless, orthogonally manipulating task-relevance alongside stimulus probability (e.g. Wyart et al., 2012) would provide the most powerful way of determining the contribution (if any) of top-down attentional mechanisms to these effects, and this presents an important avenue for future work.

9.5. Relationship to effects of action on tactile perception

Predictive influences of action on perception were investigated in this thesis using visual and auditory paradigms. The results obtained across these experiments suggested that predictions generated during action have a generally facilitatory effect on perceptual processing, in contrast with the predictions of the Cancellation model. While researchers interested in the Cancellation account have proposed that the predictions should operate equivalently across all sensory modalities (Brown, Adams, Parees, Edwards, & Friston, 2013; Wolpert et al., 2003), some of the most high-profile evidence for this account has been drawn from the tactile domain. For example, considerable empirical and theoretical attention has been paid to the force-tapping task, where participants perceive forces applied to the palm of their hand after tapping a lever as less intense than equivalent forces applied in the absence of movement (Bays et al., 2005). The experiments reported in this thesis have not used tactile tasks and evidence of action-induced facilitation in visual or auditory domains does not imply identical effects will be obtained in touch.

However, it is important to recognise that investigating predictive influences of action on tactile perception is likely complicated by the occurrence of sensory suppression (Chapman, 1994). A number of studies have revealed general decrements in tactile sensitivity during action execution (Chapman & Beauchamp, 2006), which is likely mediated by mechanisms operating at the level of spinal cord (Seki & Fetz, 2012). The mechanisms underlying sensory suppression effects are not sensitive to predictive relationships between actions and outcomes – for example, Juravle and Spence (2011) report that participants are poorer at detecting gaps in ongoing vibration applied to the wrist when they are juggling, but gaps in vibration are not a probable consequence of juggling.

The existence of sensory suppression mechanisms complicates the interpretation of cancellation effects in tasks which compare perception of tactile stimuli during active movement and passive stimulation, even when the stimuli used resemble naturalistic action outcomes (Bays et al., 2005). In these studies, the underlying logic is that participants will not engage predictive mechanisms during passive stimulation and therefore comparison of this condition to active movement reveals influences of prediction. However, these two conditions also differ in the degree to which they recruit non-predictive sensory suppression mechanisms (Chapman & Beauchamp, 2006) and it is therefore plausible that cancellation effects reported in these tasks are not effects of prediction.

This confound could be addressed by adopting conceptually similar paradigms to those reported in this thesis where participants typically produce an action on every trial and either experience a congruent or incongruent tactile outcome. While congruency conditions differ in how they will be influenced by an identity-specific prediction process, these conditions would be influenced equally by non-specific sensory suppression.

9.6. Conclusion

This thesis has presented evidence that sensory predictions generated during action shape perceptual processing. The results are consistent with a model wherein expectations generated on the basis of action influence perception in an equivalent way to other forms of predictive signal.

References

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology: CB*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(8), 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

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

Anton-Erxleben, K., Abrams, J., & Carrasco, M. (2010). Evaluating comparative and equality judgments in contrast perception: attention alters appearance. *Journal of Vision*, *10*(11), 6. https://doi.org/10.1167/10.11.6

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience*, *7*(5), 542–548. https://doi.org/10.1038/nn1241

Baeyens, F., Eelen, P., & Crombez, G. (1995). Pavlovian associations are forever: On classical conditioning and extinction. Journal of Psychophysiology, *9*(2), 127.

Bailenson, J. N., & Yee, N. (2005). Digital Chameleons: Automatic Assimilation of Nonverbal Gestures in Immersive Virtual Environments. *Psychological Science*, *16*(10), 814–819.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629. https://doi.org/10.1038/nrn1476

Bays, P. M., & Wolpert, D. M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. *The Journal of Physiology*, *578*(2), 387–396. https://doi.org/10.1113/jphysiol.2006.120121

Bays, P. M., Wolpert, D. M., & Flanagan, J. R. (2005). Perception of the consequences of self-action is temporally tuned and event driven. *Current Biology: CB*, *15*(12), 1125–1128. https://doi.org/10.1016/j.cub.2005.05.023

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review*, *116*(1), 84–115. https://doi.org/10.1037/a0014351

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

Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can’t you tickle yourself? *Neuroreport*, *11*(11), R11-16.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640. https://doi.org/10.1038/2870

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5,*341-345.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22.

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495. https://doi.org/10.1016/j.tics.2005.08.007

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, *14*(4), 411–427. https://doi.org/10.1007/s10339-013-0571-3

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

Buonomano, D. V. (2000). Decoding temporal information: A model based on short-term synaptic plasticity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *20*(3), 1129–1141.

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., & Waszak, F. (2010). A new look at sensory attenuation. Action-effect anticipation affects sensitivity, not response bias. *Psychological Science*, *21*(12), 1740–1745. https://doi.org/10.1177/0956797610389187

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313. https://doi.org/10.1038/nn1194

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition*, *106*(2), 579–593. https://doi.org/10.1016/j.cognition.2007.03.004

Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *The European Journal of Neuroscience*, *28*(6), 1208–1215. https://doi.org/10.1111/j.1460-9568.2008.06419.x

Catmur, C., & Heyes, C. (2011). Time Course Analyses Confirm Independence of Imitative and Spatial Compatibility. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(2), 409–21. https://doi.org/10.1037/a0019325

Catmur, C., & Heyes, C. (2013). Is it what you do, or when you do it? The roles of contingency and similarity in pro-social effects of imitation. *Cognitive Science*, *37*(8), 1541–1552. https://doi.org/10.1111/cogs.12071

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

Chapman, C. E. (1994). Active versus passive touch: factors influencing the transmission of somatosensory signals to primary somatosensory cortex. *Canadian Journal of Physiology and Pharmacology*, *72*(5), 558–570.

Chapman, C. E., & Beauchamp, E. (2006). Differential controls over tactile detection in humans by motor commands and peripheral reafference. *Journal of Neurophysiology*, *96*(3), 1664–1675. https://doi.org/10.1152/jn.00214.2006

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: from origin to function. *The Behavioral and Brain Sciences*, *37*(2), 177–192. https://doi.org/10.1017/S0140525X13000903

Cook, R., Dickinson, A., & Heyes, C. (2012). Contextual modulation of mirror and countermirror sensorimotor associations. *Journal of Experimental Psychology. General*, *141*(4), 774–787. https://doi.org/10.1037/a0027561

Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(4), 840–852. https://doi.org/10.1037/a0019256

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Cracco, E., Bardi, L., Desmet. C., Genschow, O., Rigoni, D., De Coster, L., Radkova, I., Deschrijver, E., & Brass, M. (in press). Automatic imitation: a meta-analysis. *Psychonomic Bulletin.*

Cutrone, E. K., Heeger, D. J., & Carrasco, M. (2014). Attention enhances contrast appearance via increased input baseline of neural responses. *Journal of Vision*, *14*(14). https://doi.org/10.1167/14.14.16

Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception & Psychophysics*, *44*(4), 339–347. https://doi.org/10.3758/BF03210415

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychological Research*, *73*(4), 463–476. https://doi.org/10.1007/s00426-009-0230-6

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(5), 1175–1185. https://doi.org/10.1093/cercor/bhn161

Dewey, J. A., & Carr, T. H. (2013). Predictable and self-initiated visual motion is judged to be slower than computer generated motion. *Consciousness and Cognition*, *22*(3), 987–995. https://doi.org/10.1016/j.concog.2013.06.007

Di Pace, E., & Saracini, C. (2014). Action imitation changes perceptual alternations in binocular rivalry. *PloS One*, *9*(5), e98305. https://doi.org/10.1371/journal.pone.0098305

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the*

National Academy of Sciences of the United States of America, *113*(28), 7900–7905. https://doi.org/10.1073/pnas.1602413113

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology. Human Perception and Performance*, *27*(1), 229–240.

Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research*, *68*(2–3), 138–154. https://doi.org/10.1007/s00426-003-0151-8

Evans, S., & Azzopardi, P. (2007). Evaluation of a 'bias-free' measure of awareness. *Spatial Vision*, *20*(1–2), 61–77.

Eysenck, M. W. (1998). *Psychology: An Integrated Approach*. Longman.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *The Behavioral and Brain Sciences*, *39*, e229. https://doi.org/10.1017/S0140525X15000965

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381–391. https://doi.org/10.1037/h0055392

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00443

Francken, J. C., Meijs, E. L., Hagoort, P., Gaal, S. van, & Lange, F. P. de. (2015). Exploring the automaticity of language-perception interactions: Effects of attention and awareness. *Scientific Reports*, *5*, 17725. https://doi.org/10.1038/srep17725

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *355*(1404), 1771–1788. https://doi.org/10.1098/rstb.2000.0734

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain: A Journal of Neurology*, *119 (Pt 2)*, 593–609.

Gallese, V., Rochat, M., Cossu, G., & Sinigaglia, C. (2009). Motor cognition and its role in the phylogeny and ontogeny of action understanding. *Developmental Psychology*, *45*(1), 103–113. https://doi.org/10.1037/a0014436

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley, New York, USA.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological Review*, *77*(2), 73–99.

Gregory, R. L. (1970). *The Intelligent Eye.* Littlehampton Book Services Ltd, Worthing, UK.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385. https://doi.org/10.1038/nn827

Han, B., & VanRullen, R. (2016). Shape perception enhances perceived contrast: evidence for excitatory predictive feedback? *Scientific Reports*, *6*, 22944. https://doi.org/10.1038/srep22944

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed Decision Time Affects the Weighting of Prior Probability in a Perceptual Decision Task. *Journal of Neuroscience*, *31*(17), 6339–6352. https://doi.org/10.1523/JNEUROSCI.5613-10.2011

Happé, F., Cook, J. L., & Bird, G. (2017). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annual Review of Psychology*, *68*(1), 243–267. https://doi.org/10.1146/annurev-psych-010416-044046

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *Neuroimage*, *62*(2), 852–855. https://doi.org/10.1016/j.neuroimage.2012.03.016

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience*, *7*(7), 523–534. https://doi.org/10.1038/nrn1931

Hebart, M. N., Görgen, K., & Haynes, J.-D. (2014). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics*, *8*, 88. https://doi.org/10.3389/fninf.2014.00088

Helmholtz, H. von. (1860). *Treatise on Physiological Optics*, *3*, 1.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*(3), 463–483. https://doi.org/10.1037/a0022288

Heyes, C. (2012). Grist and mills: on the cultural origins of cultural learning. *Phil. Trans. R. Soc. B*, *367*(1599), 2181–2191. https://doi.org/10.1098/rstb.2012.0120

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, *22*(2), 233–240. https://doi.org/10.1016/j.cogbrainres.2004.09.009

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(05), 849–878. https://doi.org/10.1017/S0140525X01000103

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology*, *6*(6), 851–857.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *12*(7), 273–280. https://doi.org/10.1016/j.tics.2008.04.002

James, W. (1890). *The Principles of Psychology* . Dover Publications.

Jeffreys, H. (1939). *Theory of Probability*. The Clarendon Press, Oxford, UK.

Jehee, J. F. M., Brady, D. K., & Tong, F. (2011). Attention improves encoding of task-relevant features in the human visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(22), 8210–8219. https://doi.org/10.1523/JNEUROSCI.6153-09.2011

Juravle, G., & Spence, C. (2011). Juggling reveals a decisional component to tactile suppression. *Experimental Brain Research*, *213*(1), 87–97. https://doi.org/10.1007/s00221-011-2780-2

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron*, *53*(3), 427–438. https://doi.org/10.1016/j.neuron.2007.01.006

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2

Kingdom, F. A. A., & Prins, N. (2009). *Psychophysics: A Practical Introduction*. Academic Press.

Kirsch, W., & Kunde, W. (2014). Impact of planned movement direction on judgments of visual locations. *Psychological Research*, 78(5), 705-720.

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

Neuroscience: The Official Journal of the Society for Neuroscience, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kontaris, I., Wiggett, A. J., & Downing, P. E. (2009). Dissociation of extrastriate body and biological-motion selective areas by manipulation of visual-motor congruency. *Neuropsychologia*, *47*(14), 3118–3124. https://doi.org/10.1016/j.neuropsychologia.2009.07.012

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal Inference in Multisensory Perception. *PLOS ONE*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868. https://doi.org/10.1073/pnas.0600244103

Kühn, S., Keizer, A., Rombouts, S. A. R. B., & Hommel, B. (2011). The functional and neural mechanism of action preparation: roles of EBA and FFA in voluntary action control. *Journal of Cognitive Neuroscience*, *23*(1), 214–220. https://doi.org/10.1162/jocn.2010.21418

Kunde, W., Koch, I., & Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *The Quarterly Journal of*

Experimental Psychology. A, Human Experimental Psychology, *57*(1), 87–106. https://doi.org/10.1080/02724980343000143

Kunde, W., Weller, L., & Pfister, R. (2017). Sociomotor action control. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1316-6

Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition*, *10*(3), 294–340. https://doi.org/10.1006/ccog.2000.0494

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press, Cambridge, UK.

Leube, D. T., Knoblich, G., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. J. (2003). The neural correlates of perceiving one's own movements. *NeuroImage*, *20*(4), 2084–2090. https://doi.org/10.1016/j.neuroimage.2003.07.033

Liepelt, R., & Brass, M. (2010). Top-down modulation of motor priming by belief about animacy. *Experimental Psychology*, *57*(3), 221–227. https://doi.org/10.1027/1618-3169/a000028

Lingnau, A., & Downing, P. E. (2015). The lateral occipitotemporal cortex in action. *Trends in Cognitive Sciences*, *19*(5), 268–277. https://doi.org/10.1016/j.tics.2015.03.006

Loose, L. S., Wisniewski, D., Rusconi, M., Goschke, T., & Haynes, J.-D. (2017). Switch-Independent Task Representations in Frontal and Parietal Cortex. *The*

Journal of Neuroscience: The Official Journal of the Society for Neuroscience, *37*(33), 8033–8042. https://doi.org/10.1523/JNEUROSCI.3656-16.2017

Love J., Selker, R., Marsman, M., Jamil, T., Verhagen, A. J., Ly, A., et al. (2015). JASP (Version 0.7.1.12) [Computer software]

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Maruya, K., Yang, E., & Blake, R. (2007). Voluntary action influences visual competition. *Psychological Science*, *18*(12), 1090–1098. https://doi.org/10.1111/j.1467-9280.2007.02030.x

Mechelli, A., Price, C. J., Noppeney, U., & Friston, K. J. (2003). A dynamic causal modeling study on category effects: bottom-up or top-down mediation? *Journal of Cognitive Neuroscience, 15*(7), 925–934.

Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences*, *8*, 22–27. https://doi.org/10.1016/j.cobeha.2016.01.006

Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, *14*(17), 2135–2137. https://doi.org/10.1097/01.wnr.0000098751.87269.77

Müsseler, J., & Hommel, B. (1997a). Blindness to response-compatible stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(3), 861–872.

Müsseler, J., & Hommel, B. (1997b). Detecting and identifying response-compatible stimuli. *Psychonomic Bulletin & Review*, *4*(1), 125–129. https://doi.org/10.3758/BF03210785

Neisser, U. (1967). *Cognitive psychology*. Appleton-Century-Crofts, New York, USA.

Neufeld, J., & Chakrabarti, B. (2016). Empathy Modulates the Rewarding Effect of Mimicry. *Scientific Reports*, *6*, srep27751. https://doi.org/10.1038/srep27751

Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2013). Crossmodal and action-specific: neuroimaging the human mirror neuron system. *Trends in Cognitive Sciences*, *17*(7), 311–318. https://doi.org/10.1016/j.tics.2013.04.012

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. https://doi.org/10.3758/BF03197524

Pazzaglia, M., Smania, N., Corato, E., & Aglioti, S. M. (2008). Neural underpinnings of gesture discrimination in patients with limb apraxia. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(12), 3030–3041. https://doi.org/10.1523/JNEUROSCI.5748-07.2008

Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, *94*(1), 61–73.

Peelen, M. V., & Downing, P. E. (2005). Is the extrastriate body area involved in motor actions? *Nature Neuroscience*, *8*(2), 125; author reply 125-126. https://doi.org/10.1038/nn0205-125a

Peirce, C. S., & Jastrow, J. (1884). On Small Differences in Sensation. *Memoirs of the National Academy of Sciences*, *3*, 75–83.

Pinto, Y., van Gaal, S., de Lange, F. P., Lamme, V. A. F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *Journal of Vision*, *15*(8), 13. https://doi.org/10.1167/15.8.13

Plotkin, H. C. (1997). *Darwin Machines and the Nature of Knowledge*. Harvard University Press, Cambridge, MA, USA.

Pobric, G., & Hamilton, A. F. de C. (2006). Action understanding requires the left inferior frontal cortex. *Current Biology: CB*, *16*(5), 524–529. https://doi.org/10.1016/j.cub.2006.01.033

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, *109*(2), 160–174.

Press, C., & Cook, R. (2015). Beyond action-specific simulation: domain-general motor contributions to perception. *Trends in Cognitive Sciences*, *19*(4), 176–178. https://doi.org/10.1016/j.tics.2015.01.006

Press, C., Gillmeister, H., & Heyes, C. (2006). Bottom-up, not top-down, modulation of imitation by human and robotic models. *The European Journal of Neuroscience*, *24*(8), 2415–2419. https://doi.org/10.1111/j.1460-9568.2006.05115.x

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(2), 870–888. https://doi.org/10.1037/a0034954

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental Science*, *14*(1), 92–105. https://doi.org/10.1111/j.1467-7687.2010.00961.x

Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, *18*(1), 6–7. https://doi.org/10.1111/j.1467-9280.2007.01839.x

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*(2), 168–185. https://doi.org/10.1016/j.neuron.2009.01.002

Ritchie, J. B., Kaplan, D. M., & Klein, C. (2017). Decoding the Brain: Neural Representation and the Limits of Multivariate Pattern Analysis in Cognitive Neuroscience. *The British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axx023

Rochat, P. (1998). Self-perception and action in infancy. *Experimental Brain Research*, *123*(1–2), 102–109.

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *Journal of Neuroscience*, *32*(24), 8424–8428. https://doi.org/10.1523/JNEUROSCI.0804-12.2012

Schneider, K. A., & Komlos, M. (2008). Attention biases decisions but does not alter appearance. *Journal of Vision*, *8*(15), 3.1-10. https://doi.org/10.1167/8.15.3

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218. https://doi.org/10.1016/j.tics.2007.02.006

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *NeuroImage*, *15*(4), 787–796. https://doi.org/10.1006/nimg.2001.1043

Schubotz, R. I., von Cramon, D. Y., & Lohmann, G. (2003). Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *NeuroImage*, *20*(1), 173–185.

Schubotz, R. I., Kalinich, C., & von Cramon, D. Y. (2004). How anticipation recruits our motor system: the habitual pragmatic event map revisited. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.). *Sensorimotor Foundations of Higher Cognition (Attention and Performance)* (pp. 141-162*).* Oxford University Press.

Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: action-induced modulation of perception. *Trends in Cognitive Sciences*, *11*(8), 349–355. https://doi.org/10.1016/j.tics.2007.06.005

Seki, K., & Fetz, E. E. (2012). Gating of sensory input at spinal and cortical levels during preparation and execution of voluntary movement. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(3), 890–902. https://doi.org/10.1523/JNEUROSCI.4958-11.2012

Sekuler, R., & Ball, K. (1977). Mental set alters visibility of moving targets. *Science (New York, N.Y.)*, *198*(4312), 60–62.

Shea, N. (2015). Distinguishing Top-Down from Bottom-Up Effects. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and Its Modalities* (pp. 73–92). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199832798.003.0004

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2013). Modulation of somatosensory processing by action. *NeuroImage*, *70*, 356–362. https://doi.org/10.1016/j.neuroimage.2012.12.043

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry*, *71*(1), 28–35. https://doi.org/10.1001/jamapsychiatry.2013.2974

Sherman, M. T., Seth, A. K., Barrett, A. B., & Kanai, R. (2015). Prior expectations facilitate metacognition for perceptual decision. *Consciousness and Cognition*, *35*, 53–65. https://doi.org/10.1016/j.concog.2015.04.015

Siegler, R. S., & Richards, D. D. (1979). Development of time, speed, and distance concepts. *Developmental Psychology*, *15*(3), 288–298.

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(46), 20099–20103. https://doi.org/10.1073/pnas.1000233107

Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision*, *2*(1), 7–7. https://doi.org/10.1167/2.1.7

Stanley, J., Gowen, E., & Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(4), 915–926. https://doi.org/10.1037/0096-1523.33.4.915

Stanley, J., & Miall, R. C. (2007). Functional activation in parieto-premotor and visual areas dependent on congruency between hand movement and visual stimuli during motor-visual priming. *NeuroImage*, *34*(1), 290–299. https://doi.org/10.1016/j.neuroimage.2006.08.043

Stein, T., & Peelen, M. V. (2015). Content-specific expectations enhance stimulus detectability by increasing perceptual sensitivity. *Journal of Experimental Psychology. General*, *144*(6), 1089–1104. https://doi.org/10.1037/xge0000109

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*, 745-756. https://doi.org/10.1038/nrn3863

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003

Summerfield, C., & Egner, T. (2016). Feature-Based Attention and Feature-Based Expectation. *Trends in Cognitive Sciences*, *20*(6), 401–404. https://doi.org/10.1016/j.tics.2016.03.008

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends in Cognitive Sciences*, *19*(1), 27–34. https://doi.org/10.1016/j.tics.2014.11.005

Swets, J., Tanner, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68*, 301–340.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*(3), 558–568.

Utzerath, C., St John-Saaltink, E., Buitelaar, J., & de Lange, F. P. (2017). Repetition suppression to objects is modulated by stimulus-specific expectations. *Scientific Reports*, *7*(1), 8781. https://doi.org/10.1038/s41598-017-09374-z

van Baaren, R. B., Holland, R. W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of*

Experimental Social Psychology, *39*(4), 393–398. https://doi.org/10.1016/S0022-1031(03)00014-3

Waddington, C. H. (1957). *The strategy of the genes.* Allen & Unwin, London, UK.

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary Observations on Tickling Oneself. *Nature*, *230*(5296), 598–599. https://doi.org/10.1038/230598a0

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011). The self in action effects: selective attenuation of self-generated sounds. *Cognition*, *121*(2), 207–218. https://doi.org/10.1016/j.cognition.2011.06.011

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal Detection Measures Cannot Distinguish Perceptual Biases from Response Biases. *Perception*, *44*(3), 289–300. https://doi.org/10.1068/p7908

Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research*, *40*(8), 925–930.

Wolfensteller, U., Schubotz, R. I., & von Cramon, D. Y. (2007). Understanding non-biological dynamics with your own premotor system. *NeuroImage*, *36*(2), 33–43. https://doi.org/10.1016/j.neuroimage.2007.03.040

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, *12*(12), 739–751. https://doi.org/10.1038/nrn3112

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1431), 593–602. https://doi.org/10.1098/rstb.2002.1238

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science (New York, N.Y.)*, *269*(5232), 1880–1882.

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(9), 3593–3598. https://doi.org/10.1073/pnas.1120118109

Yon, D., & Press, C. (2017). Predicted action consequences are perceptually facilitated before cancellation. *Journal of Experimental Psychology. Human Perception and Performance*, *43*(6), 1073–1083. https://doi.org/10.1037/xhp0000385

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301–308. https://doi.org/10.1016/j.tics.2006.05.002