Causality and Sense of Agency: A Social-Cognitive Approach with Temporal and Spatial Binding

Joshua P I Moreton

A thesis submitted for the degree of Doctor of Philosophy

Department of Psychology

University of Essex

September 2018

Declaration

I declare that this thesis, *Causality and Sense of Agency: A Social-Cognitive Approach with Temporal and Spatial Binding*, represents my own work, except where otherwise stated. None of the work referred to in this thesis has been accepted in any previous application for a higher degree at this or any other University or institution. All quotations have been distinguished by quotation marks and the sources of information specifically acknowledged.

Submitted by: Joshua P. I. Moreton

Signature of Candidate:

Date: 04/01/2019

Table of Contents

List of Figures/Tables

Acknowledgements

Where to begin. I write this section at an early hour during the morning due to my lifelong battle with insomnia. On this particular occasion, the cause (or at least, what I have inferred as the cause), is that within this section, the last I am to write for my thesis, I must now reflect on who to be thankful for – those that endowed me with the drive, motivation, and self-belief to complete this piece of work. In turn, I naturally reflect on every event that has led me to this very specific moment, in what could be either my last step, or perhaps my first step, within the realm of professional academia. My patient examiners may skip this section freely, as I am undoubtedly about to waffle on about nonsensical musings as the emotion, the stress, the elation, and the culmination of my PhD journey is about to end.

I must first and foremost, thank Mr Bradley Snell, who first introduced me to psychology during my A level college program, and gave me my first academic subject that I was genuinely interested in. And yet, it was Mr Snell's enthusiasm and teaching rigour that inspired me, that at a young and suggestible age, had he been as dull and robotic as seemingly every other teacher, I may have not engaged with the subject at all, and seen my life take a drastically different course.

Moving swiftly on, the next person who comes to mind is Dr Netta Weinstein. Netta was my MSc Psychology dissertation supervisor, and her friendly charm and unexpected enthusiasm towards me as a future researcher changed my life forever. Had Netta not given me the opportunity to be her research assistant, to learn the ways of experiments, programming, and writing academic papers, I undoubtedly would not be sitting here at 4am typing these acknowledgements. I cannot thank Netta enough for believing in me and giving me the opportunity to

brag about the various types of social psychology research I have conducted to all my future interviewers.

Next, I would like to thank Professor Mitchell Callan. As with Mr Snell, if Mitch had not injected me with his love and passion for his academic area, I may not be sat here today. Mitch's command as a lecturer is unrivalled in my experience, and as a result I learned to be as enthusiastic and passionate about psychology as him. The day I first sat in a lecture theatre with Mitch Callan at the helm was the day I knew that psychology would most certainly be my pursuit and ambition for the rest of my life. Mitch is as tough as he is passionate, and I still consider myself extremely lucky to have been his PhD student, and for his patience despite my continuous mistakes and errors over the last 3 years.

The natural progression would then be to thank Dr Gethin Hughes, my primary PhD supervisor, whose advice and guidance has been invaluable over the last 3 years. I had not heard of intentional/temporal binding at all until I met Gethin, and to be frank, nor did I care to learn about it. However, Gethin, like Mitch, gave me the enthusiasm over a subject that I learned to thoroughly enjoy, and now take much pleasure in discussing with fellow academics. I must thank him for his patience with my inability to code experiments properly and assist me whenever he could to learn how to program with various software. And finally, for his tireless assistance in proof-reading this thesis, helping me become the best writer I can possibly be.

I must also give special credit to Dr Gerulf Rieger for generously taking me under his wing during the final year of my PhD. During a time of doubt and internal turmoil of my academic and professional abilities, Gerulf helped calm all of the typical mental and emotional woes that would-be high-achievers with

questionable confidence in themselves suffer from, and despite being far from my field of research, became the life-coach every PhD student needs in the fight against imposter syndrome.

One particular lesson I have learned throughout the course of my PhD is that it would not have been as survivable had it not been for the routine support, theoretical discussions, and unhealthy quantities of sarcasm provided by my colleagues in room 4.738 in the Centre for Brain Science. Their advice, equal suffering of imposter syndrome, and reminders of valuable information that had become lost within the chaos of my Essex email account helped me on many occasions.

Of course, one must thank one's parents in such sections, especially if it is one's first ever acknowledgements section. Their pride in me as a son helped fuel my fire to keep going when, as they say, times were tough.

And, last of all, yet most of all, I thank the love of my life, Greta, whose patience, understanding and wisdom not only helped shape my perspective on psychology and academia, but also unendingly supported me throughout the last 3 years. I cannot fathom where I would be right now if it was not for her, but on taking a guess, it would be living a life unfulfilled and without cause for happiness. You are my rock, as I am yours. Thank you, my love.

And of course, if you are reading this, I must thank you, the reader, and to anyone who takes an interest in my work. I hope that you find the next 40,000 words more concise and scientifically written, though I cannot guarantee I did not slip in jovial commentary occasionally to lighten up the coming tough, dense read ahead.

General Abstract

Over a decade of research evidences that the perceived time and space between two events becomes contracted when we believe there is a causal connection between them. The present thesis investigated how such measures of implicit causal beliefs translate to meaningful, social contexts, inspired by research from the areas of sense of agency and immanent justice reasoning. Specifically, how effective were the popularly employed methods of temporal and spatial binding at reflecting or extrapolating beyond simple stimuli or otherwise explicit reports of causal beliefs. Chapter 1 gives an overview of the literature on causality, the sense of agency, the binding phenomenon, and immanent justice reasoning. Chapter 2 focuses how the emotional valence of our outcomes modulates temporal binding. These experiments were conducted to establish a basic effect that in intentional binding, and by extension, our SoA, could be modulated by social outcomes. Chapter 3 focuses on the extent to which unconnected events could be causally linked due to their moralistic congruency, measured via temporal binding. Chapter 4 investigated causal binding with regards to self- and other-actions. These experiments were conducted to explore whether top-down processes (knowledge of a causal relationship) were sufficient to produce the binding phenomenon. Chapter 5 investigated extent to which explicit self-report measures used within immanent justice reasoning research replicates with spatial binding. Finally, Chapter 6 offers a general discussion of the findings from my research, and implications for future research centred on the use of the binding phenomenon in causal and sense of agency research.

Author's Note

Chapters 2 through to 5 of this thesis were written as independent pieces of research, with the aim of being submitted as peer-reviewed manuscripts for publications. As such, there is inevitable overlap between chapters in terms of their background literatures and explanations of terms and methodologies. Chapter 2 has been published in *Consciousness and Cognition*, as of March 2017. Chapter 3 will not be submitted given the nature of the findings. The research within Chapter 4 requires continuation, and thus further research is required in order for the manuscript to be of sufficient quality to submit to a high impact journal. The data presented in Chapter 5 are part of a larger dataset that has been submitted and is currently under review in *Social Psychological and Personality Science*. As no publication attempt will be made for chapters 3 and 4, the manuscripts do not have abstracts. Each of these journals conforms to APA formatting guidelines.

Chapter 1: Inferring Causality and the Sense of Agency

Inferring Causality and the Sense of Agency

> *"The whole sting and excitement of our voluntary life . . . depends on our sense that in it things are really being decided from one moment to another, and that it is not the dull rattling off of a chain that was forged innumerable ages ago"* (William James, 1890)

Understanding causal relations within our environment is vital for social interaction (Bodner, Engelhardt, Minshew & Williams, 2015), learning (Gopnik, Sobel, Schulz & Glymour, 2001), reasoning (Goswami & Brown, 1990) and self-awareness (Duval, Silvia & Lalwani, 2012; Duval & Wicklund, 1973). Without understanding causal relations – if, how, where, when and why one event causes another to occur – we must consider whether we would develop knowledge of the world, and by reflection, ourselves. For instance, to infer causality is the foundation to operant conditioning (e.g., learning that leaving the toilet seat up causes my partner to condemn my lack of bathroom-etiquette), and a demonstration of our brain's finesse to effortlessly compute probabilistic inferences upon external events (e.g., my partner's prior knowledge of my bathroom-etiquette leads her to infer it was myself, and not one of our guests, that left the bathroom seat up). Pivotal to my thesis, understanding causal relations is the bedrock of our understanding of ourselves as causal agents in this universe, that our voluntary actions cause meaningful, intended consequences (e.g., to intentionally irritate my partner).

This sense of our agency (SoA; Moore, Wegner & Haggard, 2009) is typically referred to as the subjective awareness that one is initiating, executing, and controlling one's own volitional actions (for a review, see Moore & Obhi, 2012). To illustrate this, we are often under the impression that we execute our voluntary actions willingly, a feeling that is often confirmed by the predicted or expected outcome of our action. For example, we flick a light-switch, which, to our senses, produces immediate light. In everyday life, the perception that we are causal agents is confirmed routinely by our many, goal-directed actions and the outcomes they produce. In psychological science, the SoA can be measured through explicit self-reports, or attributions, of causality, where the individual is asked whether they believe they, or perhaps another agent, are responsible for an event or outcome that occurred (Dewey & Carr, 2013; Sato & Yasuda, 2005). This is thought to tap into *reflective* agency, as it is measured after the outcome has occurred and allows retrospective inference. However, these explicit self-attributions can introduce bias due to distorted inferential processes and individual differences related to cognitive capacities or personality (Aarts, Custers & Wegner, 2005; Dewey & Knoblich, 2014), and therefore processes of retrospective inference are mainly seen as measures of congruency between current intention and an experienced effect (Wegner, 2003). Therefore, of interest to many researchers is the ability to measure *pre-reflective*, i.e., implicit, processes that indicate SoA.

One such highly popular method of measuring implicit SoA is known as intentional binding, which refers to the compression of the perceived time interval between voluntary actions and their sensory consequences (Engbert, Wohlschläger & Haggard, 2008; Haggard, Clark & Kalogeras, 2002; Moore et al., 2009). More specifically, an outcome (e.g., a tone) is experienced earlier when it is triggered by a voluntary action compared to when it occurs in isolation or is triggered by an involuntary movement. Similarly, actions that trigger an event are experienced later than actions with no discernible outcome (see Moore & Obhi, 2012, for a review). Therefore, rather than just a perceptual bias, intentional binding has been proposed to serve as a function to help construct a coherent conscious experience of our actions in relation to their outcomes (Haggard et al., 2002).

Intentional binding parallels David Hume's (1739/1978) notion of temporal contiguity, where two events occurring temporally close together are more likely to be inferred as causally related. Indeed, the observed temporal contraction between causally-related events seen within the intentional binding literature indicates that not only are temporally contiguous events are inferred to be more causally related, but the belief in the existence of a causal relationship between two events influences our perception to engender the events to become more temporally contiguous (Desantis, Roussel, & Waszak, 2011; Moore & Obhi, 2012). This phenomenon has also been reported directly, where estimated temporal intervals between cause and effect are much shorter relative to non-causally related events (Ebert & Wegner, 2010; Moreton, Callan, & Hughes, 2018).

Such perceptual biases equally transfer to spatial binding, which complements intentional binding; instead of the perceived temporal interval becoming contracted, it is the perceived space between events that contracts with increased conformity to laws of cause and effect; for example, temporal propriety (Buehner & Humphreys, 2010). Although spatial and intentional binding bear a 1:1 mapping of spatial/temporal perceptual distortion, spatial binding has, thus far, been employed as a measure of perceptual biases in general causality (e.g., Buehner & Humphreys, 2010), whereas intentional binding is routinely used in research pertaining to self-caused actions and the SoA (Haggard et al., 2002; Moore & Obhi, 2012).

The binding phenomenon has received extensive attention over the last 16 years since Haggard et al.'s (2002) seminal paper, and despite the question of whether intentional binding truly measures agency (Dewey & Knoblich, 2014; Hughes, Desantis, & Waszak, 2013), intentional binding persists in usage as a proxy measure of SoA. However, the nature of the binding phenomenon, both temporal and spatial, has been heavily debated in terms of whether it reflects SoA or causality in general (Buehner, 2012; Buehner & Humphreys, 2009). The nature of the binding phenomenon and its relation to causality is discussed further within this chapter.

Cognitive biases can emerge during the attribution of causal agency (Campbell & Sedikides, 1999; Takahata et al., 2012). A prominent example within social psychology of a bias in causal attribution is immanent justice reasoning, where actions are thought to bring about deserved outcomes; rewards for moral or good behaviour, and punishments for immoral or bad behaviour. However, the crucial element of immanent justice reasoning lies within the lack of any plausible connection between the two events. As such, two events, such as someone behaving immorally, and then encountering something bad happening to them later, are often more causally linked due to beliefs that the world is governed by, not only laws of physics, but equally laws of justice, order and predictability (Callan, Ellard & Nicol, 2006; Callan, Sutton, Harvey & Dawtry, 2014). In other words, those who believe more in a just world are more likely to attribute a causal connection between a bad person (e.g., somebody who had committed murder) and something bad happening to them (e.g., they later suffered a fatal traffic accident), simply because they deserved their fate (Callan, Sutton & Dovale, 2010; Harvey & Callan, 2014).

The previous research pertaining to causality, the SoA, and immanent justice reasoning amalgamates towards the aim of this thesis and my doctoral research project: how causal perception can be modulated during social settings. To this end, I combined two complimentary areas of causal perception: The Sense of Agency and Immanent Justice Reasoning, and employed the binding phenomenon as an implicit measure of causality. Both SoA and immanent justice reasoning contain research demonstrating that, despite our capacities for complex causal judgements, this process can be skewed in favour of personal biases and beliefs. These phenomena are particularly well suited to further explore how we infer causality, what biases moderate the causal attribution process, and how this process translates to social contexts, since they provide an indirect measure of the causal relationship between events in the world. Such phenomena equally afford an alternative to the self-report measures typically used in immanent justice reasoning research, where any underlying cognitive mechanisms have yet to be explored. In short, whether cognitive or personal biases modulate the perception of causal relations.

Crucially, preliminary evidence already suggests that the SoA is amenable to top-down modulation (Desantis, Roussel & Waszak, 2011; Hughes, 2015; Kuhn, Nenchev, Haggard, Brass & Gallinat, 2011; Pantelis & Feldman, 2012), and is subject to similar biases (e.g. self-serving bias, Baumeister, 2010) as explicit causal attribution (Yoshie and Haggard, 2013; Takahata et al., 2012). This provides an excellent basis for investigating the extent to which intentional binding is moderated by social factors, such as the emotional consequences of actions, and equally, if binding is a general measurement of causal inference via temporal contiguity (Buehner, 2012), whether it can be applied to immanent justice reasoning.

The rest of this chapter will be devoted to discussing the background literature to my research in greater detail, centring in turn on causality, the SoA, intentional binding, and immanent justice reasoning. These will be discussed within the broader context of what my thesis aims to explore: how attributions of causality, specifically the perceptual error phenomena observed within SoA research, translate to social, meaningful contexts. Additionally, I aim to address the effectiveness and accuracy of binding, both temporal and spatial, at measuring the perceived causality of events.

Causal Inference: A Background

"All knowledge degenerates into probability." (David Hume, 1739)

Causality has been both debated extensively within philosophy, physics and psychology for at least, to our knowledge, 2000 years. Causality's earliest origins within the academic arena appear to hark back to Aristotelian philosophy where 'cause' means to explain or to answer why or how 'effects' occur. Causality, or, specifically, understanding causal relations, appears at the heart of learning, experience and knowledge (Gopnik, Schulz, & Schulz, 2007), where even Aristotle himself advocated "*we do not have knowledge of a thing until we have grasped its why, that is to say, its cause*." (Aristotle, Physics 194 b17–20). Although this appears self-evident, that our knowledge is gained through experience of the world, of the physical mechanisms that interact and give rise to many phenomena, of interest to many researchers of the current century, myself included, are *causal reasoning* and *inference*. In other words, what moderates how we infer causal conclusions based on premises, or using premises to explain conclusions, and what can moderate or influence this process.

Three distinct types of causal reasoning emerge through these millennia of discourse: deductive, inductive, and abductive reasoning (Copi, Cohen & McMahon, 2016; Evans, Newstead & Byrne, 1993; Josephson & Josephson, 1994). Deductive reasoning refers to the logical deduction of a conclusion based upon one or more premises. To continue our toilet-themed examples: if men have bad bathroom etiquette (first premise), and I am man (second premise), I therefore have bad bathroom etiquette (conclusion). Where deductive reasoning is the application of logical premises to arrive at a concrete conclusion (i.e., going from the general to the specific), inductive reasoning utilises observations to form a potentially sound theory and enable generalisations. Inductive reasoning bases the strength of its conclusion upon available evidence, allowing for the possibility that the conclusion may be false (i.e., going from the specific to the general); for example, our discovery of biological life forms upon Earth has shown a dependence upon water to exist. Therefore, all biological life depends, to some extent, on water to exist. Thus, inductive reasoning allows for probabilistic predictions; the conclusion, in other words, is a testable hypothesis.

Lastly, abductive reasoning is a logical inference process that seeks to explain observations with the best or most reasonable explanation based upon all available evidence. For example, witnessing a moving eight ball on a billiard table, one might logically explain the eight-ball's movement by the ball having been recently struck by the white cue ball. Abductive reasoning is, therefore, a type of retroactive inference, making use of prior, albeit often incomplete, knowledge to arrive at the most likely conclusion, of which will be updated in the light of new information. Use of abductive reasoning, however, should be treated with caution. For example, the famous Hawthorne effect (Roethlisberger & Dickson, 1939) was discovered by a series of experiments set within the Hawthorne Works factory, Illinois, where work productivity was incorrectly linked to different lighting settings installed by the researchers. Confused at the temporary nature of the boost in work productivity, they later realised the latter increased as a result of their presence in the factory, declining quickly after their departure.

Thus, utilising abductive reasoning also requires scientific inquiry, given that there are often multiple possible causes for an effect, in order to arrive at the most likely cause from the available evidence. However, when used as a basis to form hypotheses, and conclusions drawn are critically evaluated, abductive reasoning is regarded as the foundation for scientific realism (Harman, 1965; Ladyman, Douven, Horsten, & van Fraassen, 1997; Lipton, 1992; 2009).

Inductive and abductive reasoning are routinely employed for hypothesis-generation and testing. Indeed, when generating and evaluating scientific hypotheses, inductive and abductive reasoning are combined into what is known as the 'inference to the best explanation' (Peirce, 1974; 1992; Raftopoulos, 2016). In other words, we generalise from specific, known truths, which generates specific hypotheses based on the available evidence. For the purposes of my doctoral topic, it is this inferential process that is key; separate from actual causality, inference of causality is a mental process that can occur consciously (Dik & Arts, 2007) or subconsciously (Hassin, Aarts, & Ferguson, 2002). Causal inference forms the foundational basis to which we interpret causal events and where our biases may interfere with the perception of causal relationships between ourselves and the external world.

Some 2000 years after Aristotle, David Hume composed his Treatise of Human Nature (1739/1978), detailing his account of human knowledge acquisition; how knowledge is garnered through inferences from probability. Hume remains a strong influence upon scientific and statistical inquiry, continuing to guide our understanding of causality well into our era of contemporary research (Buehner & Humphreys, 2009; Granger, 1980; Holland, 1985). Hume argued we *infer* causal relations based on these experiences, rather than from direct perception of this relation, and there is strong evidence to support this notion given our sensory organs' inability to detect causal relations, nor does sensory input provide explicit information as to causal relations (Nadel & Hardt, 2011; Shanks, Holyoak, & Medin, 1996; Scholl & Tremoulet, 2000).

However, a primary contribution by Hume relevant to my doctoral thesis is his identification of three criteria upon which we generate causal associations: spatial/temporal contiguity, temporal succession, and constant conjunction. Spatial/temporal contiguity refers to two events occurring close together in space and time. The importance of contiguity, for example, can be illustrated by its necessity to behaviourism: for example, within Pavlovian conditioning the neutral stimulus must occur temporally and spatially contiguous to the unconditioned stimulus to become a conditioned stimulus (Rescorla, 1988; Siegel & Allan, 1996). In operant conditioning, where and when the action and outcome occur are key to generating association/causal inference (Schwartz, 1989). For example, if one presses a lamp's switch, and the lamp turns on, you may reasonably infer that your button press turned the light on. However, if a different lamp turned on (a lamp not spatially contiguous to your switch) or if the lamp turned on 5 minutes later (an occurrence not temporally contiguous to your action), the action's effect becomes ambiguous, and causality would be harder to infer. Temporal succession is largely self-explanatory, where the cause, as is universally advocated, must precede the effect (Rigden, 2005). Lastly, the constant conjunction refers to the events of cause and effect occurring consistently, and that over time the perceived causal connection strengthens with each consistent repetition of the cause-effect circumstance (or is subsequently weakened or abolished if the effect does not follow the cause). Our sensory and perceptual experiences throughout life deliver us constant conjunctions between a myriad of cause and effects, and through this our internal representation constructs what we believe as the causal external world (Hume, 1973; Wegner & Wheatley, 1999).

Thus, Hume set out specific criteria with which we learn causal relations via inductive reasoning, in which we use this knowledge in future instances when judging the same or similar causal relations (i.e., abductive reasoning). Contemporary cognitive science continues to demonstrate the strength of Hume's criteria within the causal reasoning process (Buehner, 2012; Griffiths & Tenenbaum, 2009; Holland et al., 1985). Employing Humean criteria, the power PC (probabilistic contrast) theory of causal inference (Cheng, 1997; Cheng & Novick, 1990; 2005) stands as a widely accepted account of causal inference, where both learning through covariation and *a priori* knowledge are necessary to acquisition of causal relations (Buehner, Cheng and Clifford, 2003; Tenenbaum, Griffiths & Kemp, 2006). In other words, we infer causal relations depending on the frequency of covariation between cause and effect, and a given cause's predictive power to produce the effect. A widely known example of this from medical statistics is that, although several factors can influence the incidence of lung cancer, the most powerful predictor is smoking, accounting for an estimated 72% of lung cancer cases (Cancer Research UK, Lung Cancer Risk, 2018). Hence, we form a causal structure of varying probabilities of causes that could account for the occurrence of a given effect (Cheng, 1997).

A useful, albeit general, maxim of the plethora of discourse within the field of causality (see Perales & Shanks, 2007, for a review) is that a given cause's predicted strength on producing the cause is defined as the probability of the effect occurring when all alternative candidate causes are absent. However, therein also lies an inherent problem; the existence of multiple contributing causes to an effect leaves room for bias during the interpretation of causal events. Different evidences have different emphases placed upon them depending on one's culture (Nemeroff & Rozin, 1989; Raman and Winer, 2004), learning (Legare, Evans, Rosengren & Harris, 2012) and the extent to which the individual conforms to a variety of personal biases, such as the self-serving bias (Campbell & Sedikides, 1999) and the need to believe in a Just World (Callan & Ellard, 2010; Lerner, 1980), particularly when selecting a primary cause as the most likely.

The rest of this chapter focuses on the potential moderators of implicit causal inference with reference to the experience as a causal agent, how we can implicitly measure the extent to which an individual infers a causal relationship, and how such moderators/measures may be applied to events external to the self.

Causal Inference

Humans are famously curious creatures, constantly striving to know and/or explain the reasons behind behaviour and environmental change. Such knowledge empowers us, allows us to grow, and, evolutionarily speaking, increases our chances of survival (Gopnik, 2000; Horner & Whiten, 2005). Indeed, "*If you want to alter outcomes, you need to know what causes them. Thus, knowing why you are unhappy, sick, or your car would not work is crucial if you want to be content, well, or mobile, respectively*" (Callan, Sutton, Harvey & Dawtry, 2014, pp. 107). However, even the causal mechanisms that underlie natural phenomena that occur in everyday life are rarely simple. For example, it might be easy to apply abductive reasoning to explain the movement of a billiard ball: we can observe that ball, being hit and subsequently moved by the white cue ball, fulfils the three causal criteria suggested by Hume. However, even ostensibly simple phenomena such as an infant's cry, the family car suddenly ceasing to start, or, God forbid, the seemingly unexplainable sudden change in your significant other's mood, rarely have a sufficient evidence/an obvious cause. Though we can employ our internal causal model of probabilities to determine a likely cause, even in a perfectly rational and logical mind-set, we still play a game of chance in our causal estimations. Thus, we are always capable of choosing, and acting upon, bad or incorrect explanations due to poor or lack of information (Roethlisberge & Dickson, 1939), if we are not already motivated towards causal explanations that suit our bias, beliefs and ideologies (Callan et al., 2014; Malle, 2006; Ross, 1977).

It is at this junction that two interesting areas of causal inference arise: that of our own sense as a causal agent, in the form of the Sense of Agency (SoA; see Moore & Obhi, 2012, for a review), and the events that we judge causally related in order to uphold a belief that the world is an organised, predictable and just place, in the form of Immanent Justice Reasoning (IJR; see Callan et al., 2014, for a review). Although these two relatively embryonic areas, academically speaking, are not explicitly related in current research, they are complimentary to one another in terms of assessing the degree to which we make causal inferences: While SoA research investigates what factors influence the extent to which we feel our actions produce specific action-effects (i.e., the perceived causal relationship between our own actions and events in the world), immanent justice reasoning research investigates what factors influence the extent to which beliefs about the world influence the perceived causality between different events in the environment.

Our belief in a just world appears to be related to our sense of self-efficacy (Correia, Salvado & Alves, 2016; Riley & Baah-Odoom, 2012). In other words, our sense of personal causal effectiveness is tied to beliefs that our actions bring about deserved outcomes. Both areas are additionally linked from cognitive biases that serve to promote self-esteem and well-being in healthy individuals (Dalbert & Stoeber, 2006; Dzuka & Dalbert, 2006; Linden & Maercker, 2011; Renes & Aarts, 2017). For example, the self-serving bias (Campbell & Sedikides, 1999) refers to the over-attribution of positive outcomes to the self, and negative outcomes away from the self. This has been demonstrated within SoA research where the temporal interval between an action and its action-effect is perceived to be shorter for outcomes of financial gain vs. loss (Takahata et al., 2012) and of positive audio sounds vs negative (Yoshie & Haggard, 2013). The self-serving bias is equally apparent with regards to IJR, where we often feel less vulnerable on the basis that we believe we have not done anything to cause negative outcomes (Furnham, 2003). Indeed, those who lack self-esteem are also less likely to adopt a self-serving bias (Campbell & Sedikides, 1999) and endorse IJR to explain negative events happening to them (Callan, Kay, Davidenko, & Ellard, 2009).

The research contained within this doctoral thesis explores both concepts of the SoA and IJR. Therefore, the next two sections will be dedicated to relevant background literatures of both areas.

Causal Inference and the Sense of Agency

"The sense that I am the one who is causing or generating an action. For example, the sense that I am the one who is causing something to move, or that I am the one who is generating a certain thought in my stream of consciousness" (Gallagher, 2000, p. 15)

The SoA can be defined as the feeling of control over our voluntary, goal-directed actions and their associated outcomes (Moore & Obhi, 2012; Moore, Wegner & Haggard, 2009). In other words, it is recognising our causal influence on the world. Broadly speaking, SoA is largely a constituent of self-consciousness (Metzinger, 2000; Vogeley, May, Ritzl, Falkai, & Zilles et al., 2004) that internally monitors our actions in the context of our willed intentions. Such intentions (e.g., to turn on a light) generate goal-specific motor programs (hand/finger movement and positioning) that direct motor action and coordination, and the ensuing motor feedback (hand/finger movement in real-time) allows us to manage our motor behaviour to adjust our actions and achieve our goals. Post-action, perceptual (i.e., sensorimotor, visual and proprioceptive) feedback (e.g., the light turning on) allows us to compare the expected outcome of our action to the actual outcome in order to gauge whether our intentions have been achieved, and contributes to an overall volitional, agentic experience of action (David, Newen & Vogeley, 2008).

To further illustrate the SoA in terms of action as a causal agent, I am voluntarily moving my fingers to type this very sentence, as you are voluntarily adjusting your eyes to read this sentence. Accompanying this action, I have the implicit sense that I am controlling and executing button presses to produce the desired articulation of this definition. Hence, SoA is separate from simply the intention to move or act in the future, and it is equally different to body ownership and sense of control; it is the online conscious experience of action fulfilling intention (Gallagher, 2000). We are agents of our actions and thoughts, and we are provided with both a sense and confirmation of this as expected outcomes routinely follow our actions. Indeed, such everyday causal relations between motor commands and sensorimotor consequences are so ordinary and familiar that they tend to be overlooked, occurring beneath our active awareness, and the more automatic or learned our actions are, the deeper they fall from this awareness, becoming ephemeral in the process. However, we still retain an agentive experience of our actions, pre-flective – that is, a subliminal experiential level of control over the consequences to our actions (Pacherie, 2001).

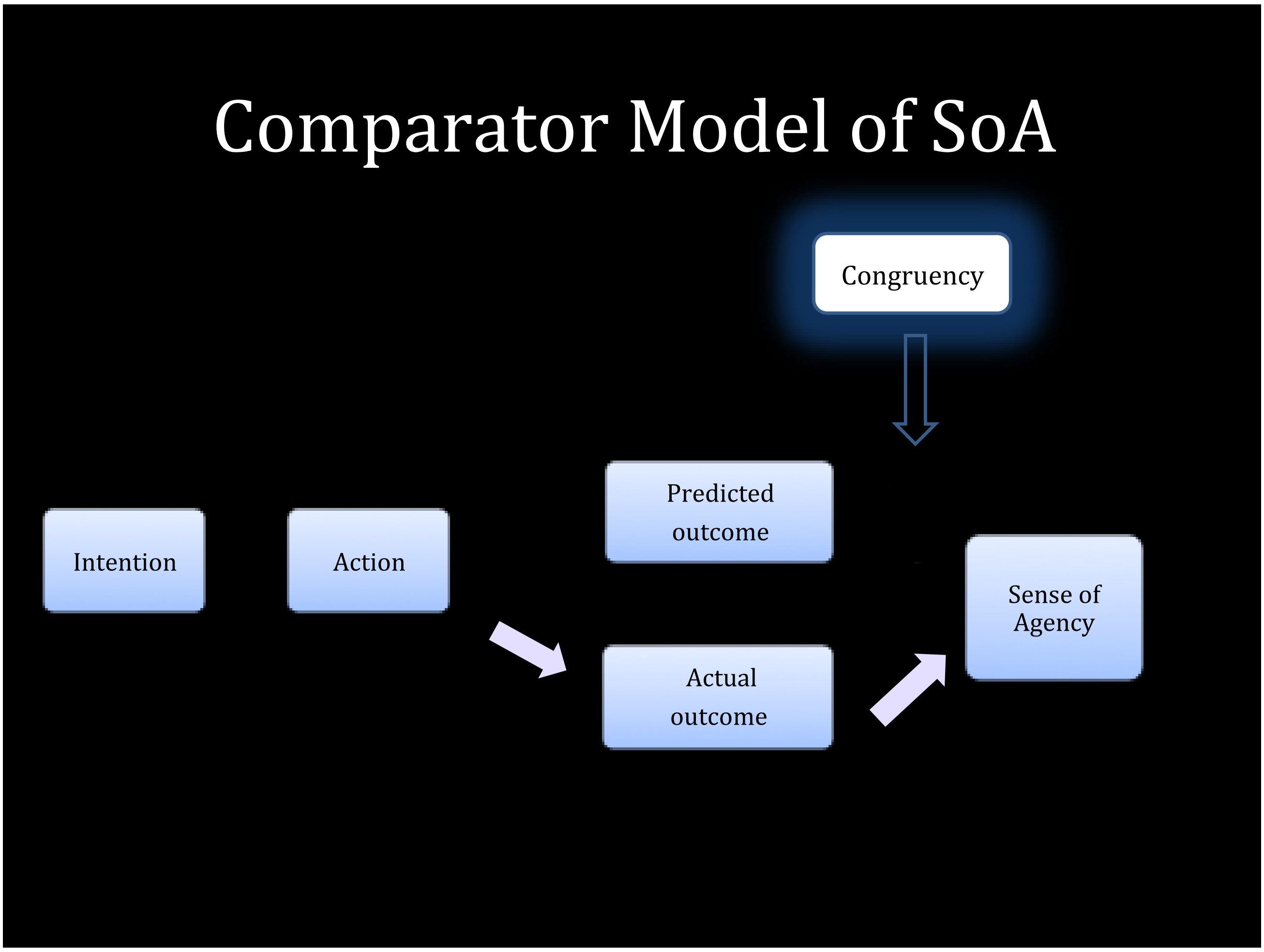
The SoA plays important roles both personally and societally – from distinguishing our actions and their consequences from other peoples' actions, to the notion of responsibility where legal systems are concerned (Gallagher, 2000; Georgieff & Jeannerod, 1998; Jeannerod & Pacherie, 2004). Thus, conceptually, as the reader may have already surmised, the SoA is related to the notion of *free will* (Aarts & Van de Bos, 2011; Davidov & Eisikovits, 2015; Feldman, 2017; Monroe, Dillon & Malle, 2014; Rigoni & Brass, 2014; Rigoni, Sammicheli & Brass, 2011), where, at the psychological level, free will and our SoA are intertwined with feelings of control; that our intentions are causally effective, and appears to operate as a default state of our experiences and perceptions. Free will and SoA are thus very conceptually similar (Feldman, 2017), however the latter distinguishes itself by focusing on the explicit/implicit cognitive processes of intentions, action, and estimations of causal effectiveness and responsibility in the light of sensory feedback – in effect, our SoA is the experience of the exercise of free will, or as some would argue, the illusion of willpower (Wegner, 2003; 2004).

The SoA acts similarly to the principles of causality (Michotte, 1963; De Vignemont & Fournet, 2004; Wegner 2003; Wegner & Wheatley, 1999), whereby our actions can be readily attributed to ourselves if our motor commands produce outcomes that satisfy the criteria of contiguity, temporal succession (also known as priority) and consistency. A further criterion is typically added, known as exclusivity, where no other event can explain the outcome. In the aforementioned light-switch example, the emerging light from the lamp would be less readily attributed to our own actions should there be multiple possible causes (such as another light-switch) or failing to correspond with any other of these criteria. However, voluntariness of action may also be necessary for the SoA to arise: In a seminal study by Haggard, Clark and Kalogeras (2002), intentional binding, a proposed implicit measure of SoA, was shown to only be present in voluntary actions. However, voluntariness of action is hotly debated within SoA research, largely pertaining to whether the methods used to measure SoA, such as intentional binding, require intention as a key component (Buehner, 2012; Hughes et al., 2013). Intention and its contribution to attributing causality are discussed further in the subsection entitled 'Intentional Binding'.

In the previous 16 or so years, much research corroborates the Humean criteria when investigating moderators of the SoA – in other words, when the harmonious flow of intention, action and sensorimotor feedback can be manipulated. For example, we experience less SoA with spatial deviations in visual representations of our movements (Farrer, Franck, Frith, Decety, Georgieff, et al., 2004; Ogawa & Inui, 2007); unexpected outcomes (Sato & Yasuda, 2005); and during ambiguity of outcome authorship (Tsakiris, Haggard, Franck, Mainy & Sirigu, 2005). Other research additionally shows less sensory attenuation, defined as the reduction in the perceived intensity of a stimulus (discussed further later within this chapter), with externally triggered outcomes compared to self-caused outcomes, both auditory (Bäß, Jacobsen, & Schroger, 2008) and visual (Hughes & Waszak, 2011). Interestingly, further evidence stems from patients diagnosed with schizophrenic disorder, i.e., those suffering from loss of the ability to attribute their own thoughts and actions to themselves. Within such patients, self-attributions of actions can become even more impaired with temporal delay (Frank, Farrer, Georgieff, Marie-Cardine & Daléry et al., 2001) and unexpected sensorimotor feedback (Hauser, Knoblich, Repp, Lautenschlager, & Gallinat, et al., 2011).

Delving further, the comparator model, otherwise known as the forward model, is supported by a large body of evidence with regards to illustrating the cognitive computation of SoA. The comparator model proposes an internal, pre-reflective comparison process in which the motor system, governed by intention, dictates a 'desired state'. Subsequently, this desired state is compared to an estimated actual state' based on current motor commands and updates itself continually within a feedback loop in order to perform specific, desired action (Blakemore, Wolpert, & Frith, 2002; Frith, Blakemore, & Wolpert, 2000; Lindner, Haarmeier, Erb, Grodd, & Thier, 2006; Synofzik, Vosgerau, & Newen, 2008;

Wolpert & Flanagan, 2001). However, in terms of SoA, the primary assumption of the comparator model is that SoA arises from congruency between predicted state or 'outcome' and the actual outcome of our actions (see Fig 1.). Intentions result in motor commands that produce action, of which predicts sensorimotor outcomes via an 'efferent copy' of the motor command. These are compared with the actual state or outcome, and congruency between predictions and outcomes increase our sense of sensorimotor control, and thus our feelings of control over our voluntary, goal-directed actions (Kumar & Srinivasan, 2017; Nahab, Kundu, Gallea, Kakareka, & Pursley et al., 2011).



Chapter 1, Figure 1. Neurocognitive comparator mechanism underlying the Sense of Agency.

The internal comparison process of SoA has been well documented over a variety of experimental procedures (de Vignemont & Fourneret, 2004; Evans, 1982; Farrer et al, 2003; Farrer, Franck, Paillard & Jeannerod, 2003; Marcel, 2003; Mechsner, Kerzel, Knoblich, & Prinz, 2001; Saito, Mushiake, Sakamoto, Itoyama, & Tanji, 2005), and is equally popular due to its explanation of differentiating sensory events caused by our own actions and those that occur via extrinsic causes (Frith et al., 2001). The comparator model also retains other advantages, such as accounting for attenuated sensorimotor consequences (Blakemore, Frith, & Wolpert, 1999; Haarmeier, Bunjes, Lindner, Berret & Thier, et al., 2001; Lindner et al., 2006) and being intrinsic to action-processing, i.e., via efferent copies of motor commands, without requiring higher-orders of conceptualisation or independent processing from action, and is thus a parsimonious explanation of SoA computation (Synofzik et al., 2008).

Similar to our internal model of causal structures, the comparator model does not operate on a Boolean congruence/incongruence format - we infer sensorimotor consequences, much like causality, from probabilistic contingencies. For example, we still experience SoA despite slight spatial, temporal or perceptual deviations from our predicted outcomes (Daprati & Sirigu, 2002; Farrer et al., 2003, Farrer, Franck, Paillard, & Jeannerod, 2003; Franck et al., 2001; Frith, 2005). Cahill, Silbersweig, and Frith (1996) demonstrated that modifications of our own voices (e.g., altering the pitch) still produce SoA, despite what would have been incongruence between predicted and actual sensorimotor feedback.

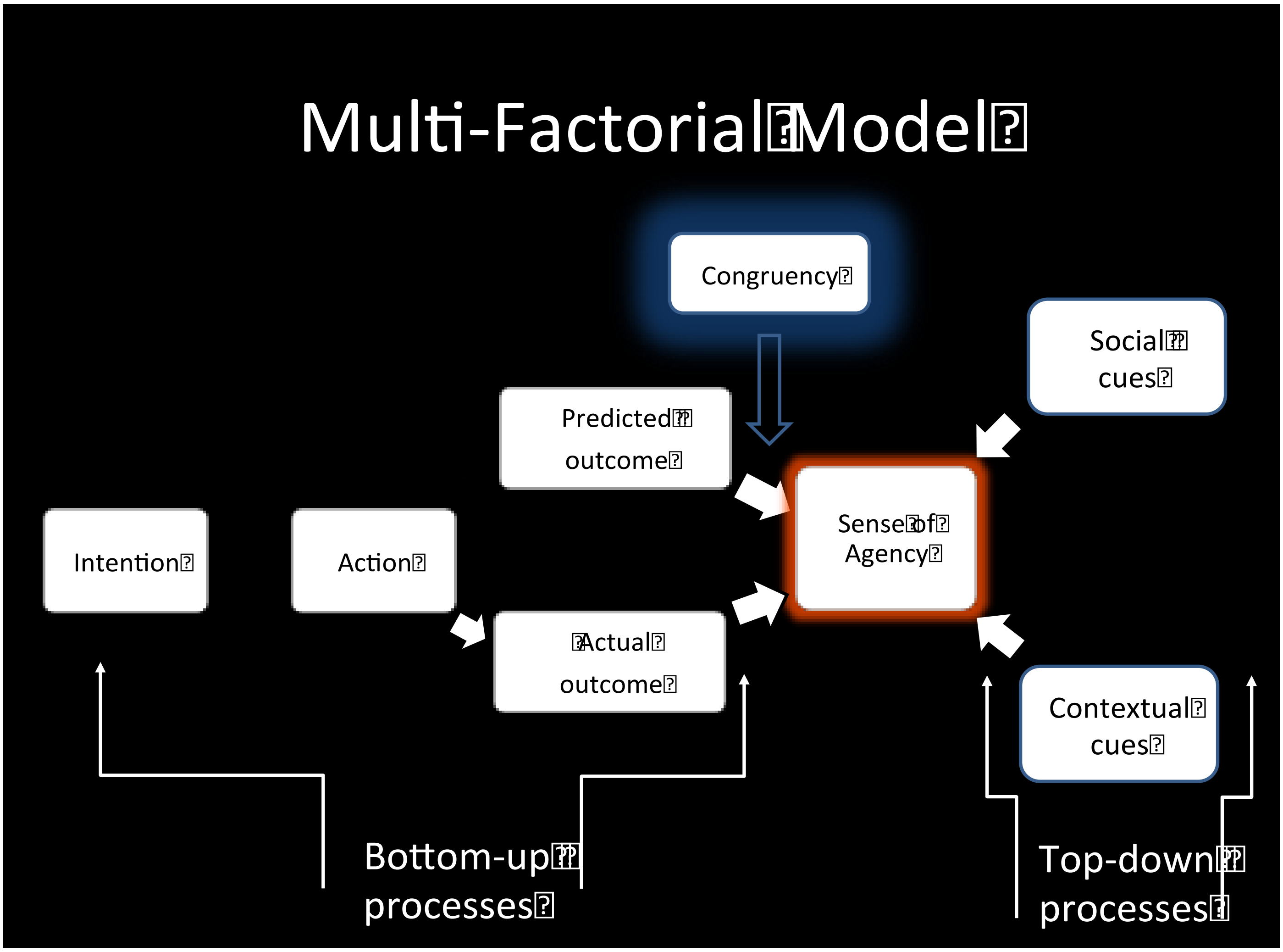
One interpretation is the experience of SoA varies upon a gradient (from no experience to a complete experience) rather than a binary system of whether SoA is present or not, depending on how congruent our prediction is to our action, as restrictive as the window of comparator processing may be (Bays, Wolpert, & Flanagan, 2005; Blakemore et al., 1999). In other words, the extent to which our predictions match the sensorimotor feedback may also predict the amount of agency we feel over those actions, given a statistical relation between events (Jenkins & Ward, 1965; Moore & Haggard, 2008). However, if the comparator model permits a degree of ambiguity in self-attributions, then extrapolating this notion would necessitate additional attribution mechanisms when self- vs. other-causal attributions are unclear (Franck et al., 2001; Farrer et al., 2003). When mismatches between expected and actual outcomes occur, further mechanisms may provide additions to, if not supplant, action-processing in estimating self- versus other- causal attributions (Synofzik et al., 2008). For example, some evidence suggests that proprioceptive and visual cues are equally important for determining agency (Fourneret, Paillard, Lamarre, Cole, & Jeannerod, 2002; Farrer et al., 2003b). Further proposals cite that multiple sensory feedback modalities should be incorporated into the framework of the comparator model, if not indicating a revision of the comparator model to a more general action/sensory feedback comparator system (Sato, 2009; Synofzik et al., 2008; Zopf, Polito, & Moore, 2018). Indeed, recent research demonstrates that visual cues play a vital role in determining agency, and incongruence between our expected and actual visual predictions (e.g., a 3D-rendered hand controlled by our own hand) significantly impairs our SoA (Caspar, Cleeremans, & Haggard, 2015; Zopf et al., 2018). Thus, criticism of the comparator model questions the sufficiency of explaining SoA purely through efferent copies of motor commands (Mechsner, 2004; Prinz, 2003; Saito et al., 2005; Synofzik et al., 2008) and neglecting retrospective elements to attributing agency (for a review, see Wegner, 2003). Regardless of the significance of motor signals in attributing agency, prediction derived from action matching expected sensory feedback is well accepted as a necessary antecedent for SoA to arise.

In terms of perceiving ourselves as a causal agent, herein lies a distinction between *feelings* of agency, and *judgements* of agency: judgements of agency are reflective, following reflection on the congruency between action and expected outcomes. Judgements are typically made through explicit self-reports, or attributions, of causality, where the individual is asked whether they believe they, or another agent, are responsible for the outcome that occurred (Dewey & Carr, 2013; Sato & Yasuda, 2005), the simplest example being "I caused X to occur". Judgements of agency are thus typically a form of abductive reasoning, employing Humean criteria in order to deduce the likeliest of explanations. Conversely, feelings of agency are more elusive, pre-reflective, and constitute a low-level, subconscious feeling of being in control of actions and the events succeeding them. A variety of methods aim to tap into this implicit form of agency, two of which popularly employed are sensory attenuation and intentional binding. Sensory attenuation posits that self-made actions provoke less intense sensory outcomes than outcomes externally generated, a classic example being that we are unable to tickle ourselves (Blakemore, Wolpert, & Frith, 1998a). Intentional binding, on the other hand, is defined as the compression of the perceived time interval between voluntary actions and their sensory consequences (Haggard et al, 2002). More specifically, an outcome (e.g., a tone) is experienced earlier when it is triggered by a voluntary action compared to when it occurs in isolation or is triggered by an involuntary movement. Similarly, actions that trigger an event are experienced later than actions with no discernible outcome (see Moore & Obhi, 2012, for a review). Regarding models of agency, the comparator model, or at least a comparison between actions, intentions, and perceptual (i.e., sensorimotor, visual, proprioceptive, etc.) feedback is concerned, has been argued to tap into this feeling of agency (Kalckert & Ehrsson, 2014; Synofzik et al., 2008).

Whichever type of agency one wishes to measure depends on one's research aims. However, both judgements and feelings of agency are not without their disadvantages: where judgements measure reflective agency, as it is measured after the outcome has occurred, these explicit self-attributions can introduce bias due to distorted inferential processes and individual differences related to cognitive capacities or personality (Aarts, Custers & Wegner, 2005; Dewey & Knoblich, 2014), or prior expectations about the task (Gawronski, LeBel, & Peters, 2007; Synofzik et al., 2008). Moreover, judgements of agency can easily be influenced by contextual or social cues; for example, Wegner (2002) demonstrated that priming participants by subliminally presenting either the word "I" or "Me" prior to target onset prompted more self-attributions than other-attributions of having caused an event when authorship was ambiguous. Conversely feelings of agency are more elusive, and popular implicit measurements of agency are not without their criticisms; for example, that intentional binding may measure temporal predictability and control rather than agency (see Hughes et al., 2013, for a review; see below for further discussion on intentional binding). Of interest to many researchers is the ability to measure pre-reflective, i.e., implicit, processes that contribute to the experience of SoA, void of the potential confounding factors of explicit attributions, and thus a plethora of studies use, and continue to use, intentional binding as a measure of agency whilst strongly proposing a tight link between intentional binding and SoA (Aarts & van de Bos, 2011; Haggard et al., 2002; Haggard & Tsakiris, 2009; Moore & Haggard, 2010; Obhi & Hall, 2011; Pfister, Obhi, Rieger, & Wenke, 2014; Ruess et al., 2017; Takahata et al., 2012).

Sense of Agency: External modulation

Progressing beyond action-prediction mechanisms, external cues constitute a potentially powerful influence over our SoA, particularly when overriding prediction-outcome mismatches. To continue the previous illustration, as I am typing this sentence to further define the SoA, I routinely make typos because I am a clumsy keyboard-user. This knowledge, coupled with the fact I am the only producer of letters upon this document, does not decrease my SoA when I make a typo, despite that my letter-prediction frequently does not match the sensory feedback I desire. As Wegner (2003) points out, external cues, in the form of context and social cues, retrospectively add to our SoA contingency equation to predict the most likely cause of action (see Fig 2). This top-down information, that I am imperfect in my keyboard strokes, maintains my predictions of sensory feedback whilst equally allowing for unpredicted feedback to nevertheless still be attributed to myself.



Chapter 1, Figure 2. Multi-factorial model underlying the Sense of Agency.

However, external cue modulation of SoA is relatively younger in terms of research potential, particularly in terms of social and contextual cues. Indeed, SoA and intentional binding are relatively new constructs within the body of psychology literature. As a result, research has primarily focused on the conditions required for intentional binding and sensory attenuation, implying presence of SoA, and the underlying predictive and reconstructive mechanisms to which we attribute agentive causation (Engbert et al., 2008). Experimental tasks often involve simple actions, such as a button press, producing simple outcomes, such as an auditory tone. These arguably lack the affective or cognitive motivation with which humans perform goal-directed actions to produce meaningful outcomes in everyday life (Moretto et al., 2011). Therefore, despite their relevance, these studies restrict our inference of how SoA translates to (relatively more) meaningful actions and outcomes.

To investigate whether these findings have real-world application, several research paradigms have recently shown that measures of SoA, through intentional binding, can be either be increased or reduced in various social contexts. These effects largely occur when either manipulating the context in which the action performed or the outcome for participants' actions. For example, Desantis and colleagues (2011) found that, when induced to believe that the outcome (a sound) could be produced by either the participant themselves or a confederate, naïve participants showed significantly less intentional binding (i.e., they judged the time interval to be longer) for outcomes when informed the confederate was responsible for the outcome, despite that the participant's actions were the cause of all outcomes incurred.

Existing studies arrive at the head of very recent research that explores how the findings from intentional binding and SoA can be applied to social settings. Findings indicate SoA coinciding with a self-serving bias, a form of attributional bias defined as a need to maintain and enhance self-esteem, or the tendency to perceive oneself in an overly favourable manner (Baumeister, 2010; Bradley, 1978; Miller & Ross, 1975). These findings typically involve greater intentional binding for positive outcomes relative to negative outcomes, such as financial gain vs. loss (Takahata et al., 2012), self- vs. other-generated outcomes (Buehner & Humphreys, 2009; Engbert, Wohlschläger, Thomas & Haggard, 2007), priming with reward-related information (Aarts et al., 2012) and success attributions compared to failure (Mezulis, Abramson, Hyde & Hankin, 2004). Such research highlights the suggestibility of our SoA over socially salient outcomes when specific information is known about the current action-outcome relationship.

Campbell and Sedikides (1999) conducted a meta-analysis of research on the self-serving bias, revealing that a perceived threat to oneself, for example notions that contradict our self-esteem, significantly increases the likelihood of a self-serving bias occurring. Likely a result of evolutionary mechanisms that promote well-being (Taylor & Brown, 1988), the self-serving bias continues to be observed within agency research where the SoA may be modulated to promote self-enhancement or protect pre-existing self-models rather than the pursue accurate, objective self-knowledge (Duval & Silvia, 2002; Gentsch, Weiss, Spengler, Synofzik & Schütz-Bosbach, 2015; Taylor & Brown, 1994). Indeed, other experimental studies have shown that the desire for self-enhancement or verification of pre-existing self-conceptions often overrides motives for obtaining accurate or objective self-knowledge (Gentsch & Synofzik, 2014; Sedikides and Strube, 1995).

As we can gather from the current advances in SoA research, it is reasonable to suggest that, measured via intentional binding, the temporal interval between action and outcome is perceived to be shorter for positive outcomes relative to negative outcomes. In other words, the valence of the predicted outcome for a given action can influence how we perceive the timing of the two events. The prevalence of the self-serving bias indicates that sensory action consequences are subject to the emotional content of action-effects, and thus emphasising the function of social cues interacting with cognitive cues when we infer a causal relationship, such as covariation and temporal contiguity (Amundson & Miller, 2007).

Whilst the influence of social contexts has been recently established, we have only scratched the surface regarding affective components of behaviour. Given the ubiquitous close relation between actions and emotions in our daily activities (Eder, Musseler & Hommel, 2012; Pessoa & Adolphs, 2010; Tamietto & De Gelder, 2010), it would be apt to update current models of SoA with affective components of our actions, either through affective states during action or the emotional value of the outcome. This omission has been recently addressed by a selection of novel studies indicating that SoA may be modulated by emotion: Yoshie and Haggard (2013) asked participants to make a voluntary action (a button press) that produced either neutral, positive or negative valence sounds (Experiment 1), and only either positive or negative sounds (Experiment 2), after an interval of 250ms. The sounds used as outcomes were a tone for neutral valence, applause and laughter for positive, and fear and disgust for negative.

Participants performed their actions in specific blocks dedicated to either valence, such that during one block (consisting of 32 trials), participants only produced outcomes of one valence. Yoshie & Haggard (2013) employed the Libet clock method, such that participants were required to estimate the onset time of either their button press or the ensuring sound. Composite intentional binding scores (combining the total error of onset estimation relative to actual occurrence) showed that, across both experiments, the total amount of error (i.e., perceiving actions closer to outcomes, and outcomes closer to actions) was significantly more for positively valenced sounds compared to negative. On closer inspection, positive sounds tended to produce minor intentional binding effects compared to neutral sounds, with arguably no self-serving bias. However, negative sounds produced significantly less onset estimation error than the other two conditions. Moreover, the majority of perceptual shift was carried largely in favour of outcome binding (i.e., the temporal attraction of outcome towards action).

Yoshie and Haggard (2013) corroborates self-serving bias research and provides insight as to how the self-serving bias is manifested, at least to the extent of auditory stimuli. However, this study suggests that SoA differences in terms of emotional valence are largely due to a reduced SoA over negative outcomes, rather than the assumed self-serving bias notion that positive self-attributions would, at least, equally contribute to these differences.

The findings of Yoshie & Haggard (2013) equally support evidence of emotional distancing (Caspar, Christensen, Cleeremans, & Haggard, 2016), rather than the attribution of positive outcomes to oneself. These findings also mirror observed intentional binding effects when actions produce financial reward or punishment, where the perceived tones indicating punishment are perceived as much later, rather than reward-associated tones being perceived as earlier, compared to neutral tones (Takahashi et al., 2012). Additionally, Hughes (2015) highlights the salience of emotional content of outcomes to our actions using sensory attenuation: Hughes found that stimuli conveying emotion modulated the sensory attenuation effect more so than neutral stimuli, with more neural suppression of expected fearful faces than unexpected fearful faces and both expected/unexpected neutral stimuli.

In the larger context of causality, we may skew temporal contiguity of negative outcomes such that we avoid causally attributing events that may harm our self-esteem or self-image. In other words, we infer less of a causal relationship between our actions and negative outcomes, relative to positive outcomes. Rather than the basic premise that causal inference allows us to learn and understand the world accurately, particularly in terms of ourselves as causal agents, cognitive biases emerge where perceptions are altered in order to perceive the causal relations we want to see, or at least, those that fall within our prior beliefs.

The first set of experiments of my PhD research, contained within chapter 2, aimed to build upon this body of research, and investigate causal inference biases with respect to SoA. This was achieved by investigating socially salient outcomes for voluntary actions in the form of emotional valence. Given my research aims of investigating whether cognitive biases modulate the perception of causal relations, I employed a binding paradigm to measure SoA. One particular goal was to explore if and how affective components modulate SoA experiences (given the aforementioned self-serving bias). Specifically, together with my PhD research supervisors (Dr. Gethin Hughes and Prof. Mitchell Callan) we conducted three studies that implicitly measured the SoA via intentional binding over actions that produce positive or negative emotionally valenced outcomes. Our expectations were weighted upon the preceding evidence that a self-serving bias largely contributes to SoA experiences.

Intentional Binding

> *"An hour sitting with a pretty girl on a park bench passes like a minute, but a minute sitting on a hot stove seems like an hour." (Einstein, Einstein in America: The Scientist's Conscience in the Age of Hitler and Hiroshima, Sayen, 1985, p. 130)*

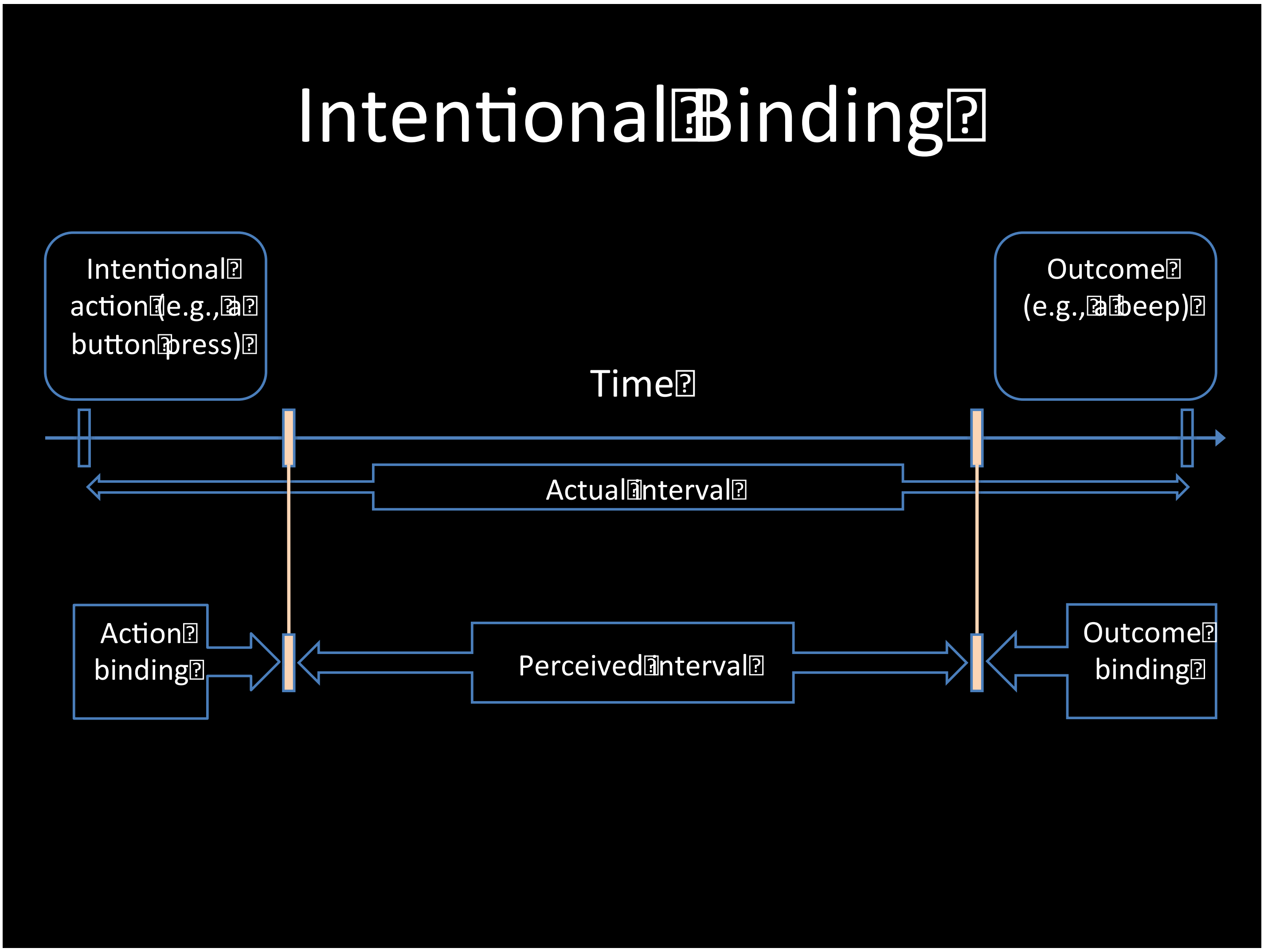
Although intentional binding has been discussed previously within this chapter, the debate surrounding its nature and underlying mechanisms necessitates a more elaborate discussion. Indeed, as my research identifies over the subsequent chapters, the binding phenomenon is not unlike other psychological phenomena that provoke both intrigue and dispute.

To define intentional binding sufficiently, binding refers to the compression of the perceived time interval between voluntary actions and their sensory consequences (Haggard et al., 2002; see Fig. 3). Libet, Gleason, Wright and Pearl (1983), developed a clock-face method (thus dubbed the 'Libet clock') upon which participants would make temporal estimations of the onset of events. The Libet clock consisted of a screen displaying a clock-face with a singular rotating hand that took 2560ms per rotation, and numbered in intervals of 5, displaying 5-60.

Employing the Libet clock some decades later, Haggard et al., (2002), explored how voluntariness of action impacted the perceived timing of events –

an action and its following outcome. In two conditions, participants, at their leisure, pressed a button. This button produced no outcome (condition 1) or produced a tone that occurred 250ms after the button press (condition 2). Each time the participant pressed the button, they were asked when, referring to the clock-face, their action occurred (for example, '17'). In two further conditions, participants heard a tone in isolation (condition 3) or pressed the button that produced the tone, again occurring 250ms later (condition 4). In these conditions, participants were asked when the sound occurred. Relative to the temporal judgements of actions producing no outcomes and sounds heard in isolation (i.e., no prior action or cause), judgements of the onset of actions that produced tones, and of the onset of tones that had been produced by the button-press, were distorted; actions were judged to have occurred much later, and sounds were judged to have occurred much earlier.

In the full study, Haggard et al. compared such voluntary actions to two additional conditions: one employing TMS stimulation over the motor cortex to invoke an involuntary action, and sham TMS stimulation. The tone followed voluntary actions, the motor cortical TMS, and sham TMS (indicated by an audible click made by TMS applied to the parietal cortex) by 250ms. Comparing voluntary actions to the TMS pulse, voluntary actions were perceived much closer to the time of the produced tone compared to actions performed in isolation. Equally, sounds produced by voluntary actions were perceived earlier than sounds with no discernible cause. Conversely, a reverse binding effect was observed in the TMS condition, where actions and outcomes were perceived further apart relative to their baseline conditions. The sham TMS condition showed only minor perceptual attraction between actions and outcomes. The authors suggested that binding of intentional actions to their effects represents a mechanism by which the brain integrates intention, motor commands, and sensory feedback into a coherent conscious experience. Moreover, that binding enables the construction of an experience of our own causal agency. This assertion has been further reinforced by other research showing that agency beliefs enhance the binding effect (Desantis, Roussel, & Waszak, 2011; Haering & Kiesel, 2012).



Chapter 1, Figure 3. Illustration of the intentional binding phenomenon.

However, the classic Libet clock paradigm, despite its popular usage (Moore & Haggard, 2008; Haggard et al., 2002; Haggard, Aschersleben, Gehrke, &

Prinz, 2002), has not suffered without criticism. For example, questionable accuracy over event-onset time estimations (Banks & Pockett, 2007) given the subjective nature of time and its fluctuating relation to real time (Eagleman et al., 2005; Hallett, 2007). Additionally, as Humphreys and Buehner (2009) highlight, the Libet clock is, at best, an indirect method of measuring an individual's perceived time between two events given that it captures the perceived onset times of individual events rather than, as Fig. 3 illustrates, the perceived interval.

As such, the Libet clock method enables investigation into the specific contributions of action-binding and outcome-binding separately. However, its criticisms, as well as both the inclination for a direct method of measuring the perceived contraction in time, and to explore whether temporal binding is not simply an artefact of using the Libet clock, led to the application of other methods of measuring time perception. One such method is interval estimation, measured by asking participants to estimate the interval between the two events (such as the action and the outcome), using either a verbal estimate (Engbert et al., 2008; Kumar & Srinivasan, 2017; Moore et al., 2009) or an estimate via a time scale (Moreton, Callan, & Hughes, 2017). Neither a reference interval or feedback is given, as typically the concern is the relational estimation differences between conditions – for example, Engbert et al. (2007) showed that the interval between self-generated actions (a keypress) and somatic stimulation applied to either the participant's or the experimenter's index finger was estimated as significantly shorter than estimations of the same sequence but with the experimenter as the cause of the outcome.

Another alternative of the Libet clock method to measure intentional binding, stimulus anticipation, was developed by Buehner and Humphreys

(2009), where participants are asked to time a button press with the onset of an outcome stimulus (e.g., a visual or auditory cue). Intentional binding occurs when participants anticipate the outcome by pressing a button *before* it appears. Anticipation reflects temporal contiguity: causally related events are perceived closer together in time, where this perceptual bias induces the notion that causally produced outcomes will appear sooner relative to unrelated or merely correlational secondary events (Buehner, 2012). For example, Buehner and Humphreys (2009) showed that auditory tones caused by keypresses were anticipated much earlier than simply timing the keypresses to the onset times.

Following Haggard et al.'s (2002) work, intentional binding has been routinely used to assess the level of agency one experiences over a given action or following outcome predominantly using the Libet clock and interval estimation methods. However, the link between intentional binding and SoA is frequently assumed (Ebert & Wegner, 2010; Engbert et al., 2008; Moore & Fletcher, 2012; Moore et al., 2009; Obhi & Hall, 2011), despite the lack of full understanding of the relationship (Hughes et al., 2013; Moore & Obhi, 2012) and alternate accounts suggesting binding results from knowledge of general causal mechanisms (Buehner & Humphreys, 2009; Stetson, Montague, & Eagleman, 2006). Regardless of the specific mechanism of binding, a plethora of research suggests that, as a causal agent, intentional binding will occur when we voluntarily perform an action that produces an outcome (Cravo, Claessens, & Baldo, 2009; Engbert et al., 2008; Haggard & Clark, 2003; Haggard et al., 2002; Humphreys & Buehner, 2010; Moore & Haggard, 2008). Equally, studies (e.g., Takahata et al., 2012; Yoshie & Haggard, 2013) that sought to modulate SoA with social outcomes employed binding as their measure. In other words, we can infer social factors impact intentional binding (as well as, presumably, SoA), and thus our perception of the outcomes to our actions.

Haggard et al.'s (2002) research first prompted the idea of intentional binding – that intentional actions and their outcomes are drawn together. This emphasises the requirement of voluntary action. However, several studies have shown that intention of action is not necessary to produce the binding effect, but may rather be a contributing factor (Buehner, 2012; Buehner & Humphreys, 2010; Humphreys & Buehner, 2009), particularly given that binding also occurs for observed outcomes (Moore, Teufel, Subramania, Davis, & Fletcher, 2013; Poonian & Cunnington, 2013). Consequently, the term 'intentional binding' has been critiqued to the point that many researchers have adopted the moniker 'temporal binding' instead (Buehner, 2012; Cravo et al., 2009; Engbert & Wohlschläger, 2007; adopted henceforth throughout my thesis), despite continued referral to the binding phenomenon as 'intentional binding' by other researchers (Christensen, Yoshie, Di Costa, & Haggard, 2016; Desantis, Hughes, & Waszak, 2012).

Another possible interpretation posits that a causal relationship between events is sufficient for said events to be drawn together in conscious experience (Buehner, 2012; Suzuki, Lush, Seth, & Roseboom, 2018), highlighted by research showing that binding disappears with actions performed intentionally but without causality (Buehner & Humphreys, 2009). Moreover, Buehner (2012) found comparable temporal binding to self-caused actions by those performed by a machine. In other words, simply knowing of the existence of a causal relationship between two events may sufficiently produce a perceived temporal contraction between them.

Referring to our internal model of causal structures, and to Bayesian principles of ambiguity reduction, one could interpret that, in order to reduce ambiguity during the noisy perception of sensory information, our past experiences of temporally contiguous events indirectly influences our perception of the timings of future events, drawing action-effect pairings closer together. Therefore, intentionality may represent a sub-structure of temporal binding, wherein it supplies an additional cue available to the perceiver that provides information when discerning causal antecedents to outcomes. This temporal contraction between such causally-related events may indicate that not only are temporally contiguous events judged to be more causally related, but the belief of a causal relationship influences our perception of events to become more temporally contiguous.

From the multiple explanations of temporal binding's underlying mechanism, one can conclude that the binding phenomenon is far from cemented in terms of our understanding. Not only are there conflicting explanations, but several studies portray conflicting findings: for example, actions and outcomes appear to be drawn together when movements are congruent to intentions (Ebert & Wegner, 2010; Zopf, Polito, & Moore, 2018). Conversely, other research distinctly found the reverse, where binding is not influenced through manipulating congruence of action-outcomes (Desantis et al., 2012; Hughes, 2018), suggesting alternative mechanisms, in the form of temporal predictability and temporal control, take precedent. One possible explanation of these inconsistencies lies in that action-binding (the perception of our actions occurring temporally closer to an action-effect) and outcome-binding (the perception of outcomes occurring temporally closer to our action) may be driven by different processes, and thus different processes may underlie them (Hughes, 2018; Waszak, Cardoso-Leite, & Hughes, 2012).

One of the most forefront inconsistencies, however, is that binding, as an implicit measure of agency, often fails to correlate with explicit measures of agency (Dewey & Knoblich, 2014). For example, Obhi and Hall (2011) showed that, during an actor-observer paradigm, both participants experienced temporal binding of the actor's actions, but only the actor reported explicit agency. Assessing the degree to which binding is modulated by factors that also modulate explicit agency reports is important to determine the relationship between implicit and explicit agency. Given that explicit (i.e., judgements of) and implicit (i.e., feelings of) agency are proposed to be distinct processes, they will rarely, if ever, show complete convergence, as the former is influenced by inferential processes in addition to the predictive mechanisms outlined in the previous section. However, positive evidence of covariation is important to argue that conscious reports and unconscious biases are indeed measuring the same underlying construct. Recent evidence suggests that neither temporal binding (Dewey & Knoblich, 2014; Saito, Takahata, Murai & Takahashi, 2015) nor sensory attenuation, a prominent alternative to binding in measuring SoA, correlate with explicit reports of agency, or even with each other (Dewey & Knoblich, 2014). Hughes et al.'s (2013) systematic review of temporal binding studies and their methodologies advocates that, although action-effect prediction may contribute to the binding phenomenon, other factors such as temporal predictability (the ability to predict when an outcome will occur) and temporal control (the ability to control the onset of a given outcome) may contribute a greater deal to binding than current research assumes.

However, what we do know currently is that the sensory predictions of our actions are derived from motor and perceptual cues and depend on the context within which they are performed (Dewey & Knoblich, 2014). A belief in a causal relationship between our actions and learned outcomes appears to successfully induce temporal binding (Desantis et al., 2011; Ebert & Wegner, 2010). Moreover, that the self-serving bias has been shown to modulate binding in a few preliminary studies (Takahata et al., 2012; Yoshie & Haggard, 2013) may be interpreted that, as far as binding as a mechanism for forming a coherent conscious experience an agent (Haggard et al., 2002), negative outcomes are less prone to being incorporated into such an experience; or at least, are less prone to forming a *coherent* conscious experience. From the perspective that binding represents a general mechanism of causal perception (Buehner and Humphreys, 2009), less binding for negative outcomes suggests that negative outcomes may be less likely to be causally attributed to oneself. Such theoretical positions lay the foundation for continued research that explores factors that modulate binding, and by extension, contribute towards the discussion on whether intentionality of voluntary action is necessary for binding to occur, as opposed to simply possessing knowledge of or belief in a causal relationship.

Thus, in line with previous research, temporal binding will be used as an implicit measure of the extent to which we self-attribute social outcomes. The findings from my PhD research expected to build upon findings showing support for the self-serving bias modulating temporal binding (e.g., Yoshie & Haggard, 2013). Additionally, given proposed alternative mechanisms (Hughes et al., 2013) and evidence that questions whether temporal binding is indeed a measure of SoA (Dewey & Knoblich, 2014), my PhD research, secondary to my primary research aims, also investigates whether temporal binding holds the validity as a SoA measure throughout current research, and provide more clarity to the discussion. If temporal binding is indeed modulated by social outcomes, as has been demonstrated (e.g. Yoshie & Haggard, 2013), this will both provide evidence of the self-serving bias's impact, consistent with the suggestion that binding reflects implicit SoA. Conversely, if we find, however, that social outcomes do not impact temporal binding, this provokes questions regarding the relationship between the self-serving bias and SoA, and the replicability of findings derived from temporal binding measures. Moreover, given that evidence has shown explicit SoA can be modulated by a self-serving bias (Oishi, Tanaka, & Watanabe, 2018), a lack of effect of social outcomes would contribute towards the discussion that explicit and implicit measures tap into independent constructs (Moore, Middleton, Haggard & Fletcher, 2012).

Measuring Causal Inference

At this juncture of my thesis, and in correspondence to the outcomes of my experiments contained within Chapter 2 (minor spoiler alert), it is necessary to further highlight the lack of solidarity over methods of measuring implicit causal beliefs. Although temporal binding is routinely employed as a measure of SoA, much research questions its proposed validity (Buehner & Humphreys, 2009; Moreton et al., 2018; Stetson et al., 2006), not to mention the struggle to replicate key findings within the literature (Moreton et al., 2017; see Chapter 4). Additionally, although the aim of my thesis was to investigate implicit causal beliefs within the context of SoA first, and then adapt my experiments to more socially relevant settings and known explicit biases in causal reasoning, again, the findings from experiments within Chapter 2 forced a shift in theoretical focus.

David Hume's (1739/1888) notion of contiguity refers to both temporal and spatial parameters. While temporal binding relates to temporal contiguity, spatial contiguity begets spatial binding (Buehner & Humphreys, 2010). The two concepts are virtually identical in theory, save for the one difference being that in spatial binding it is the physical space between two events that becomes perceptually contracted when a causal belief is held or implied. The roots of spatial binding date back to Michottean-era research (Michotte, 1946/1963) where Michotte observed induced perceptual causality with a launching task. In this classic task, one ball would make contact and launch a second ball, and causal beliefs reported by participants. Although a simple task, and not without its limitations (Beasley, 1968; Joynson, 1971), this visual account of perceptual causality guided research to discover a new perceptual bias we humans can make.

The basic Humean notion for contiguity relates to that, if two events occur close together in temporal and spatial proximity, we are more likely to infer that they are causally related (Straube & Chatterjee, 2010; Woods, Lehet, & Chatterjee, 2012). However, the converse is also true: events that are causally linked are perceived to be closer together in both time (Faro, Leclerc, & Hastie, 2005) and space (Buehner & Humphreys, 2010). For instance, actions and their ensuing sensory outcomes are bound towards one another in time (Haggard et al., 2002; Hughes et al., 2013). Similarly, Buehner and Humphreys (2010) demonstrated that the distance between two moving balls is judged to be smaller when the movement of the two balls are causally linked. In their study, Buehner and Humphreys (2010) asked participants to replicate the size of a rectangular bar between two events: one ball hitting the bar on the left side, and another ball launching as soon as the first ball contacts the bar, along the same trajectory. In a second condition, there was a temporal delay of the second ball's launch. The authors found that errors in bar replication showed that participants perceived the bar in the first condition shorter relative to non-causally related events (i.e., spatial binding), whereas the converse was true in the delayed condition. In their second experiment, two further conditions were added: priority violation, where the second ball launches before the first ball makes contact, and upward launch, where the second ball launches upwards instead of along the same trajectory as the first ball. They replicated their findings with the first two conditions, and in the two new conditions also found spatial binding, but at a significantly reduced rate compared to the first condition.

Buehner (2012) followed the previous research by demonstrating that temporal binding is likely causality-induced time compression, as opposed to specifically related to motor-identity planning and intentional action. Thus, overall, both time and space appear to contract when causal beliefs are held about two events. Given that the self-serving bias, an explicit bias, has been reported to interact with temporal binding (Takahashi et al., 2012), other such biases might also be effectively captured using the binding phenomenon.

One such explicit bias, known as Immanent Justice Reasoning (IJR; Callan et al., 2014), is a causal reasoning bias where one morally-valenced event is said to have caused a following, morally-congruent event, despite an absence of plausible physical mechanism between the two events. As stated in a previous sub-section of this thesis, and to keep the shift in theoretical focus thematically related, SoA and IJR are complimentary given the fact both are supposedly linked through cognitive biases that serve to promote self-esteem and well-being in healthy individuals (Dalbert & Stoeber, 2006; Dzuka & Dalbert, 2006; Linden & Maercker, 2011; Renes & Aarts, 2017; Yoshie & Haggard, 2013). Furthermore, exploring IJR with temporal binding allows us to assess the degree to which the binding phenomenon translates to more complex stimuli without apparent physical mechanisms or causal laws, and investigate whether personal biases (for instance, the belief in a just world) impacts not only our attitudes and judgements, but also our cognitive capacities for perceiving causal relationships within our environment. In other words, how do the perceptual error phenomena observed within SoA research translate to social, meaningful contexts?

Chapter 2 questions whether temporal binding can be modulated by emotional valence, where one possibility is that the binding phenomenon is a result of general causality between events (Buehner & Humphreys, 2010; Buehner, 2012; Schlottmann, Ray, Mitchell, & Demetriou, 2006; Scholl & Tremoulet, 2000) rather than forward-motor identity prediction (Haggard et al., 2002). Unlike emotional valence, IJR stems from a wealth of research demonstrating its modulatory power over explicit causal judgements (Callan et al., 2006; Callan et al., 2012). Thus, despite our lack of findings in Chapter 2, we pursued our original research aim in Chapter 3 by exploring whether top-down causal beliefs (such as IJR) may modulate causal binding, similar to how external factors have been shown to modulate SoA (Aarts et al., 2012; Caspar et al., 2016; Engbert et al., 2007).

To be clear, the causal mechanisms of outcomes in temporal binding research are physical by nature; for example, producing a tone or image (e.g., Engbert et al., 2008; Haggard et al., 2002; Moore & Fletcher, 2012) to appear onscreen via a mechanical action performed. As such, the mechanisms of binding appear to arise from a weighted integration of sensory evidence and specific prior belief (Kawabe, Roseboom, & Nishida, 2013; Lush et al., 2018 Synofzik, Vosgerau, & Lindner, 2009). Within IJR research, however, the mechanism is abstract; there is no physical mechanism by which an individual may attribute the moral valence of one event as the cause to another (discounting probabilistic causal sequences of events). Hence, my research additionally explores whether the power of causal beliefs to modulate physical causality (e.g., Desantis et al., 2011) can be applied to abstract causality. If knowledge of a causal relationship is entirely sufficient to produce a temporal or spatial contraction between two events, as suggested by Buehner (2012), then there is reason to suspect that causal events without a physical mechanism may suffer the same modulatory processes. Thus, this work builds upon the concept that binding reflects causal beliefs, rather than reflecting the integration of sensory cues and predicted action-effects during physical causal events. Conversely, if the abstract causal link between two morally congruent events do not produce similar modulated consequences to temporal binding (i.e., contracted when a causal relationship is believed to be present), then one possible interpretation is that temporal binding between two causal events requires weighted integration of sensory evidence from physical events.

Furthermore, IJR is predominantly measured through self-report, and thus, if binding, both temporal and spatial, were truly measures of implicit causal beliefs, it would be of interest to explore how such explicit ratings of causal judgements correlate with implicit measures of causality via temporal (Buehner, 2012) and spatial binding (Buehner & Humphreys, 2010).

Exploring IJR with implicit causal measures such as temporal and spatial binding is instrumental to our original research aim of investigating whether cognitive biases modulate the perception of causal relations, exploring if and how temporal binding observed within SoA research translated to social contexts; specifically, to implicit causal inference between morally skewed events. However, as mentioned previously, the findings from experiments within Chapter 2 question whether temporal binding is modulated by social outcomes in the form of emotional valence. However, given the precedence set by research suggesting temporal binding might reflect an implicit sense of causal relations (Buehner 2012), as well as the wealth of evidence in the form of biased causal reasoning in explicit judgements (see Callan et al., 2014, for a review), in Chapter 3 we decided to pursue our original research aims. Thus, the goal of the Chapter 3 was to extend previous IJR research by employing temporal binding, where participants estimated the perceived temporal interval between events. The next subsection focuses on the background literature of IJR, before culminating in an overview and description of the research presented throughout this thesis. Furthering my research conducted on emotional valence and IJR, Experiments 6 and 7 of my PhD research, contained within Chapter 4, I, along with my PhD supervisors, employed a stimulus anticipation method to test causal binding with regards to self- and other-actions. Experiment 6 conceptually replicated Buehner (2012), who compared self-caused outcomes to those of a machine's to explore whether intention or simply the appearance of a causal mechanism would necessitate temporal binding. These experiments were conducted to extrapolate Buehner's (2012) findings and investigate whether causal knowledge is truly necessary to elicit temporal binding. We used stimulus anticipation again in Experiment 7, this time conceptually replicating Desantis et al.'s (2012) research showing that the belief of another causing an outcome, despite the outcome being self-caused, can reduce temporal binding (compared to believed self-caused outcomes). Given the findings from Experiment 6, we explored how stimulus anticipation, rather than the Libet clock method used in Desantis et al. (2012), could measure the impact of authorship beliefs on temporal binding. If stimuli are anticipated more for self vs. other actions or other prediction signals signifying a following event, and thus there is greater binding, this evidences a self-serving bias and the role of intentionality and causal mechanisms. Equal or less anticipation, however, questions such roles, despite their propagation as necessary.

Causal Inference and Immanent Justice Reasoning

"We're seeking justice, Alfred. How can that ever be a mistake?"

(Batman, Batman Vol. 1, Batman Year 3 - Changes Made, Wolfman & Broderick, 1989)

Implicit measures of causal attributions pertaining to self-actions also apply to those of other-actions (Callan, Moreton & Hughes, *submitted*; Buehner, 2012). Advancing beyond the self-serving bias, another interesting bias that modulates explicit reports of causal inference is immanent justice reasoning, where actions are thought to bring about morally congruent outcomes. Given the promising findings that events can be perceived as causally linked due to inherent desires for justice and predictability within the world (Callan, Ellard & Nicol, 2006; Callan, Sutton, Harvey & Dawtry, 2014), I investigated the perceived causal relation between justice-themed events using implicit measures, offering an alternative to the self-report measures typically used in immanent justice reasoning (IJR) research.

Illustrations of IJR are not uncommon in news cycles, particularly when a person in a position of authority deigns to explain an event via scientifically implausible means, likely to insert their ideological beliefs within their explanation. For example, designated the Great East Japan Earthquake, in May 2011 a magnitude 9.0 earthquake occurred off the eastern coast of Japan. The earthquake and resulting tsunami caused thousands of fatalities and injuries, in addition to destroying over 100,000 households. Globally, experts concerned themselves over the geological explanations of the event. Many who sought to explain the earthquake by other, irrational means, however, accompanied these explanations. Prominently, the governor of Tokyo, Shintaro Ishihara, incited that the catastrophe was deserved and just, proclaiming divine intervention as the cause due to the people of Japan's "selfishness and greed" (McCurry, 2011). This example neatly defines IJR, which is the belief that actions bring about deserved outcomes, be they rewards or punishments, when there is no physically plausible means by which they might have done so (Callan, Sutton, Harvey & Dawtry, 2014). IJR represents a departure from normal types of rewards and punishments – a robber receiving a prison sentence, for example; it is the distinct lack of plausible, or even possible, logical connection between the two events.

IJR is interesting and worthy of study simply because, in contrast to the previous section on causal reasoning, attributing one event as the cause to another, despite the absence of any plausible or rational link, flies in the face of basic logic and scientific understanding of physical laws. Humans are capable of understanding complex causal relationships, but, as it appears, personal bias or beliefs are equally capable of overriding logical causal reasoning. Such thinking has the power to be harmful and cause social unrest, especially when tragic events are attributed to one's religious beliefs or sexual orientation (Burt & DeMello, 2003; Gledhill, 2009).

Causal understanding develops during early childhood (Gopnik et al., 2001), despite disagreements as to whether domain-general causal knowledge is innate (Goodman, Ullman, & Tenenbaum, 2011) or entirely constructed through experience (Carey, 2009). Constrained by lack of knowledge of causal/physical laws, parental sanctions of behaviour, and a tendency to view the complex world as a coherent, intentional system, IJR flourishes where past deeds seemingly have a direct impact on future events (Fein & Stein, 1977; Piaget, 1932/1965). However, the capacity to endorse causal attributions that defy logic persists well into adulthood, and possibly at a greater frequency than children (Callan, Ellard, & Nicol, 2006; Callan et al., 2014; Maes, 1998; Raman & Winer, 2002, 2004; Woolley, Cornelius, & Lacy, 2011). As a classic example, Callan et al. (2006) gave participants a vignette regarding a man ('David') who had suffered a brutal traffic accident. In two conditions, participants were either informed that David recently had an extra-marital affair, or that he had purchased a family vacation. Participants in the former condition attributed, via a Likert rating scale, David's accident significantly more to his prior behaviour than when David had not had an affair.

Multiple studies reaffirm the findings of Callan et al. (2006; Callan, Sutton, & Dovale, 2010; Callan, Harvey, Dawtry, & Sutton, 2013; Harvey & Callan, 2014; Raman & Winer, 2002, 2004; Woolley, Cornelius, & Lacy, 2011), yet there are multiple possibilities as to how adults engage in this particular brand of faulty causal reasoning, and how it comes to exist alongside our internal causal structures and abilities to compute complex causal matrices. Moral intuitions, cultural traditions and emotional connotations, teaching us right from wrong from early age, may persist into adulthood as a bias when it comes to decision-making (Baumard & Chevallier, 2012; Haidt & Joseph, 2007; Rand, Greene, & Nowak, 2012). Equally, absence of knowledge of cause and illiteracy of scientific principles are well known to invoke irrational arguments for causes, a prominent example being the 'God of the gaps' argument (Callan et al., 2014; Coulson, 1955). Additionally, adults develop multi-focused thinking that allows them to entertain multiple arguments for a cause that vary across their grounding in reality. Exposure to cultural messages, be they religious, folklore, etc., increases throughout development, reinforcing notions that our behaviour and outcomes are morally congruent. This process may prompt adults to both maintain and employ both rational and non-rational forms of logic and reasoning, with either or both being exposed given specific circumstances to the individual (Legare, Evans, Rosengren, & Harris, 2012).

These reasons largely explain how we can preserve non-natural explanations of causal relations into adulthood, but the motivation to engage in IJR stems from the need to believe in a just world (Lerner, 1980). Just World Theory states that we behave and interpret information in a way that allows us to maintain the view that the world is fair, just, and orderly. People get what they deserve, and when making sense of misfortune, especially given a disastrous event with no obvious cause, we can justify faulty causal explanations in order to maintain perceptions of justice and deservingness (Callan & Ellard, 2010; Hafer & Bègue, 2005; Lerner, 1980). Thus, IJR acts as a defence mechanism when we perceive a threat to our belief in a just world. This bears similarity to the frequently observed outcome of victim derogation, where, upon witnessing or learning information about a negative circumstance (for example, an otherwise innocent individual suffering a tragic accident or illness), devaluing or finding fault with the victim helps us 'make sense' of the events transpired (Lerner & Miller, 1978; Ryan, 1971).

Indeed, IJR appears to serve multiple functions from the a Just World perspective (Lerner, 1980): recent research shows that construing events to be consistent with a just world allows us to make long-term goals (Bal & van den Bos, 2012; Callan, Harvey, Dawtry, & Sutton, 2013; Xie, Liu, & Gan, 2011), avoid self-defeating behaviours (Callan, Kay, & Dawtry, 2014), avoid smaller, immediate rewards in order to obtain larger, delayed rewards (Callan, Shead, & Olson, 2009), and maintain a commitment to justice in the face of threat (Hafer & Bègue, 2005).

Rather than a strictly causal-reasoning account of social cognition, IJR represents a form of motivated reasoning when explaining events that seemingly have no causal connection. If no prior cause (other than sheer chance) is available when a specifically positive or negative event occurs, we become motivated to search for morally congruent behaviour to explain the events and shape them accordingly to fit into a Just World narrative. However, an imbalance also exists where negative events appear to threaten the belief in a just world more than positive events (Callan et al., 2006; Percival & Haviland, 1978). Whereas a positive event might incline us to search for good prior behaviour, negative events impact our motivations for justice to the extent that we resolve our justice concerns throughout future events. For example, Callan et al. (2006) gave participants either of two scenarios: one where a woman suffering from AIDS either made no recovery from her treatment, or that she had recovered and was experiencing no further symptoms. Afterwards, participants were given two further vignettes, one detailing a horrific traffic accident involving a bully, and one involving a charitable elderly couple winning the lottery. Participants were asked to what extent did they think the prior behaviour of the bully/elderly couple caused their later misfortune/fortune. Callan et al. found that participants presented with the vignette detailing the continued suffering of the woman with AIDS were significantly more like to attribute the bully's misfortune to his prior conduct, compared to those who read about the woman's successful treatment. However, there was no difference between groups regarding IJR accounts of the elderly couple. Thus, negative events appear to inspire more motivation to resolve threats to justice than positive events. This, as has been said, helps us maintain the view that the world is not chaotic and unpredictable, but rather fair and just, where people get what they deserve if they behave poorly (see Callan et al., 2014, for a lengthier discussion on positive vs. negative differences in IJR).

An underpinning factor related to IJR is the idea of deservingness – that is, when the moral congruency of action and outcome are consistent (Lerner, Miller, & Holmes, 1976). What may directly impact upon the extent we employ IJR, as opposed to systematically judging causal probabilities, is the perceived deservingness of an event given prior unrelated behaviour (Callan et al., 2006; 2010). Deservingness and IJR are separate, yet mutually inclusive, and whether deservingness estimations act as a justification for employing IJR in causal estimations, or rather deservingness is an underlying mechanism and precondition to IJR, is still unclear (Callan et al., 2014). Concerns for deservingness and justice appear to go hand-in-hand when motivations arise to pursue a causal connection that is otherwise physically impossible (Bal & van den Bos, 2012; Callan, Shead, & Olson, 2009, 2011; Hafer, 2000a; Hafer, Bègue, Choma, & Dempsey, 2005) and estimations of causality and deservingness consistently correlate strongly (Callan et al., 2006). However, some evidence suggests that deservingness mediates the relationship between prior behaviour and causal attributions towards future events, such that the more deserved an outcome, the more frequently IJR is used (Preacher, Rucker, & Hayes, 2007). The research contained within chapter 5 of this thesis aims to elucidate the role between IJR and implicit measures of causality.

Moving forward, IJR research often utilises self-report methods to measure the strength of the causal link between someone's prior (mis)deeds and a future event. For example, reading a vignette detailing a fatal accident had occurred to either a good or bad person, followed by the question ''to what extent do you feel that what happened to *xxx* was a result of his/her conduct?" (Callan et al., 2013). To address this dependence on explicit measures of causality, Callan, Ferguson and Bindemann (2012) showed that motivations to perceive justice exist beyond retrospective causal judgements using eye-tracking. Participants listened to a recorded vignette describing either a good or bad person through headphones, and their eye-gaze was recorded whilst two possible outcome images (good or bad) were displayed partway through the vignette. Callan et al. found that participant gaze shifted towards the morally congruent image, such that when listening to the bad (good) behaviour of a person, eye-gaze focused onto the bad (good) outcome. For example, when listening to the description of Allen either shouting and swearing at his overworked wife to make him food (bad) or sympathising with his overworked wife, buying her flowers and making his wife her favourite dinner (good), participants' eye gaze more frequently fell to the morally congruent visual depiction of a following outcome (e.g., a successful business contract for a good outcome/terrible car accident for bad) before the outcome was announced. These findings suggest not only are we motivated to retrospectively assert morally congruent reasons for specific outcomes (Callan et al., 2006), we also anticipate morally congruent outcomes when justice is concerned.

The research within Chapters 3 and 5 aims to extend previous IJR findings by exploring whether valence-laden sentences within a scenario are perceived closer in time, measured via temporal binding (Chapter 3) or felt closer in space, measured via spatial binding (Chapter 5). Thus, in terms of IJR, morally congruent sentences should be more temporally contiguous to one another and felt closer in space. Conversely, morally incongruent sentences, and hence less causally related, should be felt further apart.

Increased spatial proximity has already been shown to increase conceptual similarity between words (Casasanto, 2008; Lakoff & Johnson, 1999). Spatial proximity between stimuli underlies a categorisation function of stimuli, impacting how close we feel two sets of stimuli are to each other. Selecting the spatial proximity between morally laden sentences may then be influenced by how easy we feel they are to categorise together. Reversing this logic, sentences that provoke concerns for justice may equally follow the same categorisation protocol, where morally congruent sentences will feel closer in spatial proximity than incongruent sentences.

Another reason for advancing IJR research beyond explicit causal judgements is that IJR represents prereflective and intuitive needs to defend the belief in a just world (Hafer & Bègue, 2005; Lerner, 2002). Unlike the belief in a just world, which is to say, a consciously evaluated belief, the need to defend the belief is not a rationalised appraisal of reality, but rather a preconscious process (Sutton et al., 2008; Sutton & Winnard, 2007). Thus, I investigated IJR using implicit measures of causality via temporal and spatial binding. If IJR reflects such a prereflective and intuitive form of causal judgement, we would expect implicit measures to reflect the explicit causal bias observed within the self-report methodologies thus far (Callan et al., 2006; Callan et al., 2012), where the temporal interval/physical space between morally congruent events would be perceived closer in time/space.

The goal of the Chapters 3 and 5 was to extend previous IJR using temporal and spatial tasks, where participants estimated the perceived temporal interval between events (Chapter 3) and varied the distance between sentences depending on how close they feel in space (Chapter 5): a ranked choice task (Experiment 8), a free positioning task (Experiment 9) and a sentence-chasing task (Experiment 10). On the basis of previous findings (Buehner et al., 2009; Callan et al., 2006; Casasanto, 2008), we explored whether the power of causal beliefs to modulate physical causality (e.g., Desantis et al., 2011) can be applied to the abstract causality, where binding is reflected in biased judgments about the timing or spatial position of abstract representations of moral acts.

However, in the wider context of my research, attention must be redrawn to the research aims with which my research was conducted: whether cognitive biases modulate the perception of causal relations. Additionally, a secondary aim, as a consequence of my findings from Chapter 2, became to assess the viability of temporal binding as a measure of general causal inference. In other words, to elucidate the current discussion over the specific mechanisms of temporal binding as a measure of SoA and implicit causal attribution in general, specifically in terms of whether the perceptual attraction of two events in time depended upon weighted integration of sensory evidence from physical events could be applied to non-physical causal events. Thus, where Chapter 2 sought to investigate temporal binding with emotional valence, Chapter 3 pursued the research aims by employing temporal binding as an implicit measure of causal inference between two morally charged events via IJR. Chapter 4, as a result of non-significant findings from the prior two chapters, explored whether causal knowledge was sufficient for binding. Finally, where our findings from chapters 2-4 remained inconclusive in regard to the research aims, we decided, as a final set of experiments, to explore IJR with spatial binding in Chapter 5.

Overview of Present Studies

Chapter 2: Experiments 1—4

We attempted to conceptually replicate findings that suggest that the emotional content of an action outcome can modulate the effects of intentional binding. Experiments 1 and 2 utilised an interval estimation measurement of temporal binding. Participants made voluntary keypresses that produced visual outcome stimuli after one of three time intervals (100, 400 or 700ms). Shortly afterwards, a time estimation scale was presented to the participants, where participants were asked to estimate the time between their key press and the resulting image. Contrary to previous findings, we found no evidence that intentional binding was affected by the emotional valence of action outcomes. Experiment 3 was conducted to validate the stimuli for equivalence of perceived emotional valence and arousal using Likert scales. Experiment 4 directly replicated Yoshie and Haggard's (2013) original experiment using sound vocalizations as action outcomes and measuring intentional binding via Libet clock method (Libet, 1980). Our replication attempt failed to detect a significant effect of emotion on temporal binding. Subsequently, these studies suggest that the emotional valence of action outcomes exerts little influence on temporal binding. The potential implications of these findings are discussed.

Chapter 3: Experiment 5

In Experiment 5, we investigated immanent justice reasoning using temporal binding within a 2x2 factor design. As with Experiments 1 and 2, Experiment 5 used an interval estimation procedure to gauge temporal binding, where participants were asked to judge the time interval between a keypress and a previously-shown possible outcome. Participants witnessed two possible outcomes happening to an individual, either fortuitously good or bad, in the form of images to a coming scenario. Participants then read the scenario, which depicted the individual as either a good or bad person. After reading, participants then pressed a button that produced one of the two previously presented outcomes after an interval of either 100, 400, or 700ms. Participants then estimated this temporal interval between scenario and outcome. We expected temporal binding scores, as an implicit measure of causal beliefs, to reflect moral congruency of behaviour and outcomes, with morally congruent behaviour and outcomes inducing smaller temporal estimates. However, temporal interval estimations did not vary, regardless of whether the action-outcome pairing was morally congruent (e.g., bad person/bad outcome) or incongruent.

Chapter 4: Experiment 6-7

Due to the findings of my previous experiments, Experiments 6 and 7 were conducted to assess whether temporal binding is brought about through knowledge or belief of a causal relationship (Buehner, 2012; Desantis et al., 2012). Chapter 4 attempts to extend two previous studies (Experiment 6: Buehner, 2012; Experiment 7: Desantis, Roussel and Waszak, 2011) to further clarify whether knowledge of or belief in a causal mechanism modulates temporal binding. Experiment 6 compared temporal binding of visual outcome (as measured via stimulus anticipation) in a self-caused condition and a time condition, where outcomes appeared within a 2 – 5s range depicted by a visual timer. We found no significant difference in anticipation at 500ms, but greater anticipation at 900ms for the timer condition. Experiment 7, also using stimulus anticipation, found no difference between self-caused outcomes and outcomes said to be caused by an online confederate. These studies question the importance of explicit knowledge of causal in driving temporal binding, as measured by stimulus anticipation. The potential implications of these findings are discussed.

Chapter 5: Experiments 8—10

Chapter 5 contains 3 experiments that explored the idea that immanent justice reasoning influences spatial proximity. Specifically, participants positioned representations of people’s fortuitous bad (vs. good) outcomes within an ordered list format (Experiment 8) and when they were free to move the outcome anywhere on the screen (Experiment 9) relative to a representation of a bad (vs. good) person. In Experiment 10, the positive or negative outcomes were being "chased" across the screen by the bad persons, and the participant was tasked with actively maintaining the distance between the representations. Our findings were consistent with our expectations that morally congruent events would feel more causally related, and that this would reflect in spatial proximity placements. Specifically, participants positioned bad (good) people significantly closer in space to representations of their previous immoral (moral) actions when the outcomes occurred to the same person (Frank punched someone - Frank was in a car accident) more strongly than when the outcomes occurred to a different person (Frank punched someone - Joe was in a car accident).

Present Studies - Overview

The University of Essex Ethics Committee for Human Research approved all experiments. Participants within Experiments 1-3 and 5-7 were recruited via the Prolific Academic subject pool and screened for the following exclusion criteria: native language other than English, left handedness, recent use of illicit drugs, uncorrected visual or auditory impairment and history of psychiatric or neurological illness. Additionally, the following were employed as inclusion criteria: a participation approval rating of below 90% (based on prior study performance-approval scores) and aged between 18-65. Experiment 4 was conducted within laboratory conditions at the University of Essex, and Experiments 8—10 recruited participants via Amazon's Mechanical Turk. For all experiments except Experiment 4, minimum required sample sizes across experiments were fixed ahead of data collection, but the final sample sizes were not completely predetermined due to the unpredictable nature of online recruitment (e.g., because of slight over-recruitment and removing participants due to duplicate IP addresses). Power calculations for mixed-effects regressions can be difficult so we based our sample sizes on achieving at least 80% power to detect small-to-medium effects (dz = 0.35) in simpler, within-subjects t-tests. A consent form protocol was presented at the beginning of each study instructing that by continuing to the experiment, the participant confirms that they have understood the information provided and consent to participating in the experiment. Each experiment was presented via Inquisit v4.01 (Draine, 1998; Millisecond Software).

Chapter 2: How Much Does Emotional Valence of Action Outcomes Affect Temporal Binding?

Abstract

Temporal binding refers to the compression of the perceived time interval between voluntary actions and their sensory consequences. Research suggests that the emotional content of an action outcome can modulate the effects of temporal binding. We attempted to conceptually replicate these findings using a time interval estimation task and different emotionally-valenced action outcomes (Experiments 1 and 2) than used in previous research. Contrary to previous findings, we found no evidence that temporal binding was affected by the emotional valence of action outcomes. After validating our stimuli for equivalence of perceived emotional valence and arousal (Experiment 3), in Experiment 4 we directly replicated Yoshie and Haggard's (2013) original experiment using sound vocalizations as action outcomes and failed to detect a significant effect of emotion on temporal binding. These studies suggest that the emotional valence of action outcomes exerts little influence on temporal binding. The potential implications of these findings are discussed.

Keywords:
Temporal binding, emotional valence, facial expressions conveying emotion, voluntary action, self-serving bias, time interval estimation, replication study.

How Much Does Emotional Valence of Action Outcomes Affect Temporal Binding?

Temporal binding refers to the compression of the perceived time interval between voluntary actions and their sensory consequences (Haggard, Clark & Kalogeras, 2002). More specifically, an outcome (e.g., a tone) is experienced earlier when it is triggered by a voluntary action compared to when it occurs in isolation or is triggered by an involuntary movement. Similarly, actions that trigger an event are experienced later than actions with no discernible outcome (see Moore & Obhi, 2012, for a review). For example, Haggard et al. (2002) examined judgements of the onset time of both a voluntary action and a resulting tone using the Libet clock method (Libet, Gleason, Wright & Pearl, 1983), where one estimates the time of onset of an action or outcome via the position of a rotating clock-hand around a clock-face. These judgements were compared to those made when only the action was performed (i.e., with no outcome) and when a sound was heard in isolation (i.e., without a prior cause). Haggard et al. found that the perceived time of an action was later when the action produced a tone compared to when there was no outcome. Moreover, the perceived time of a sound was earlier when the sound had been produced by an action compared to when it was heard in isolation. In other words, temporal binding means that the time interval between an action and its outcome becomes perceptually compressed when we think there is a causal relationship between action and outcome. Temporal binding has also been observed with methods other than the Libet task, such as verbal or numerical estimates of the interval between action and outcome (Humphreys & Buehner, 2009; Humphreys & Buehner, 2010). Temporal binding has been shown to occur for both self- and other-generated actions (Moore, Teufel, Subramaniam, Davis & Fletcher, 2013; Poonian & Cunnington, 2013) and may be a general phenomenon linking causally related events (Buehner, 2012).

To date, researchers have mostly investigated the conditions required for temporal binding and the mechanisms that underpin it (Hughes, Desantis & Waszak, 2013), and they have done so using experimental tasks that often involve basic actions, such as a button press, producing sensory feedback, such as an auditory tone (David, Newan & Vogeley, 2008; Sato & Yasuda, 2005). These temporal binding tasks arguably lack any real-world complexity with which humans perform goal-directed actions to produce meaningful outcomes in everyday life (Moretto, Walsh & Haggard, 2011). Researchers have started to examine the generalizability of temporal binding effects to stimuli beyond simple and arbitrary outcomes, such as priming social cues (Aarts, Bijleveld, Custers, Dogge, Deelder et al., 2012), authorship of action cues (Desantis, Weiss, Schütz-Bosbach & Waszak, 2012), leader-follower cues (Pfister, Obhi, Rieger & Wenke, 2015) and economic and pain cues (Caspar, Christensen, Cleeremans & Haggard, 2016). For example, Aarts et al. (2012) found that, when primed with a positive picture (taken from the International Affective Picture System; Lang, Margaret & Bruce, 1999) that indicated a reward, temporal binding during the Libet clock task increased compared to neutral primes. Takahata et al. (2012) trained participants to associate two tones with either financial gain or loss. Using the Libet task, they found that the temporal interval between judgements of onsets for actions and outcomes of financial loss was significantly larger than for judgements of financial gain. In other words, negative outcomes reduced the effect of temporal binding. This points towards the possibility that temporal binding might be driven by self-serving biases, where one is more inclined to associate positive events with the self compared to negative events (Mezulis et al., 2004; Miller & Ross, 1975).

Yoshie and Haggard (2013) directly tested this idea by investigating whether temporal binding differed between outcomes that varied in terms of their intrinsic emotionality. They asked participants to make voluntary actions (a key-press) that produced auditory sounds that were either of positive or negative emotional vocalisations (e.g., laughter or disgust). Participants made temporal estimations of their actions and the ensuing sound via the Libet clock method. They found that positive sounds produced shorter estimations of onset-time between the action and sound compared to negative sounds (Experiment 1), with this effect being mostly driven by decreased binding to negative outcomes (Experiment 2).

Yoshie and Haggard's (2013) research provided promising evidence that negative emotional outcomes reduce temporal binding, which occurs presumably because people are less inclined to attribute negative outcomes to themselves. However, despite the potential importance of Yoshie and Haggard's (2013) findings, they have yet to be replicated using other temporal binding tasks and different emotionally-valenced action outcomes. Thus, answering Christensen, Yoshie, Di Costa and Haggard's (2016) call for more research exploring the emotional modulation of temporal binding using alternative methods, the goal of the current research was to conceptually replicate Yoshie and Haggard's (2013) temporal binding effects using an interval estimation procedure (vs. the Libet task; Moore et al., 2012) and images of faces conveying positive and negative emotions (vs. emotional vocalizations; experiments 1 and 2). Moreover, we conducted a separate study to validate the perceived valence of the face stimuli we used in Experiments 1 and 2 (Experiment 3), and we conducted a highly-powered direct replication of Yoshie and Haggard's first experiment (Experiment 4). On the basis of Yoshie and Haggard's findings, we expected that temporal binding would be smaller for negative outcomes (faces or vocalizations conveying negative emotions) than for positive outcomes (faces or vocalizations conveying positive emotions).

Experiment 1

We used an interval estimation procedure to gauge temporal binding (Ebert & Wegner, 2010; Engbert, Wohlschläger & Haggard, 2008; Moore, Wegner & Haggard, 2009). In this procedure, participants are asked to judge the time interval between an action and its sensory outcome (e.g., a button press and a sound). Using this procedure, Engbert et al. (2008) found that the interval between voluntary actions and visual, auditory, and somatic outcomes were compressed compared to the interval between passive actions and similar outcomes. For our task, participants were asked to press the space bar, which was followed by emotionally valenced action-outcomes—namely, emoticons depicting positive, neutral, or negative emotions (see Figure 1). Emoticons are prevalent throughout modern technological communication, and frequently used to convey emotion (Derks, Bos & von Grumbkow, 2011; Hudson, Nicolas, Howser, Lipsett & Robinson et al., 2015). Research has shown that emoticons elicit similar cortical responses to real faces (Churches, Nicholls, Thiessen, Kohler & Keage, 2014) and that emotions conveyed in emoticons are subject to similar behavioural biases (Öhman, Lundqvist, D., & Esteves, 2001) and neural processing disruptions (Jolij & Lamme, 2005) as real faces.

Method

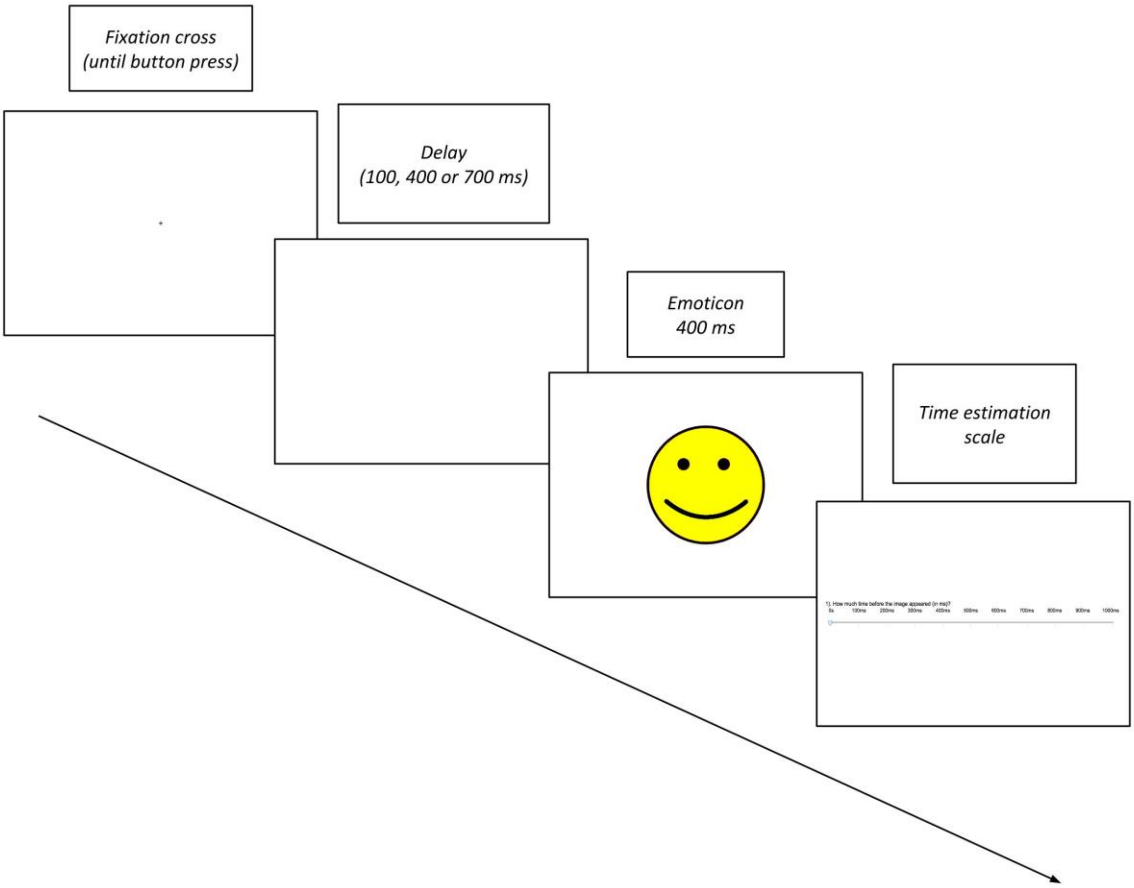
Participants. We recruited 80 native English-speaking participants (51 males, M_{age} = 33.91, SD_{age} = 11.27) through prolific.ac, an online crowdsourcing platform. Participants received monetary compensation. We screened participants for the following inclusion criteria: an approval rating of above 90% on prolific.ac (based on prior experiment performance/approval scores) and aged between 18-65. The required sample size was fixed ahead of data collection, and a power analysis showed we had 90% power to detect a small effect (Cohen's f = .10) of emotional valence on temporal binding (α = .05).

Materials and procedures. Experiment 1 consisted of 100 trials: 10 practice and 90 experimental trials. We used an interval estimation procedure to measure temporal binding (see Moore & Obhi, 2012). For each trial, participants saw a fixation cross on-screen, and in their own time, pressed the spacebar. In the practice block participant actions produced a neutral stimulus, which was a green circle with a diameter equal to the emoticon images. During practice trials, the green circle appeared after a randomly selected time interval from either 0ms or a multiple of 100ms up to 900ms. We used all intervals in the practice block, to encourage participants to expect the full range of durations in the experimental block. During the practice block, feedback was provided to participants after they made their time estimations. Feedback consisted of both the participant's estimated time and the actual time of stimulus onset to enhance familiarity with estimating time in milliseconds.

In the experimental condition, an emoticon appeared after either 100, 400 or 700ms (Moore et al., 2009), which remained on-screen for a further 400ms. We varied the delay intervals to increase participants' uncertainty regarding the interval between action and outcome to allow for variation in judgement times (cf. Ebert & Wegner, 2010). The emotional expressions of the emoticons were manipulated by orienting the lines representing the mouth: curved upwards for positive, curved downwards for negative, and a straight line for neutral. The emoticons were genderless, varied only in the shape of the mouth, and were presented on a white background in the center of the screen (see Figure 1).

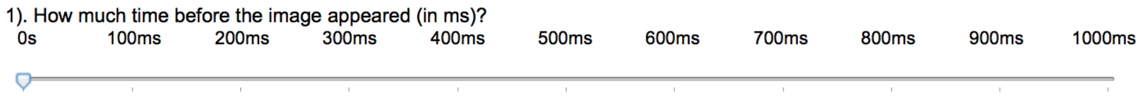
Chapter 2, Figure 1. Emoticons used in Experiment 1.

Participants underwent two blocks of 45 trials, allowing for 30 presentations of each emoticon image in total. Participants were instructed that they would not receive feedback for their time estimations during the experimental trials. A schematic display of the sequence of trial events is shown in Figure 2.



Chapter 2, Figure. 2. Schematic display of the sequence of trial events for Experiment 1.

Both the time intervals and emoticons (either positive, negative or neutral) were pseudo-randomized across trials, such that there was the same number of trials in each condition at each time interval. A blank screen then followed the emoticon for 400ms, replaced by a horizontal time estimation scale in the center of the screen (see Figure 3). The scale ranged from 0-1000ms, with demarcation lines every 100ms. Participants were instructed to scroll the slider along the bar to the time that they believed it took the image to appear since their action (in multiples of 100ms). Once selected, participants confirmed their selections by clicking on a 'finish' button and proceeded to the next trial.



Chapter 2, Figure. 3. Time estimation scale

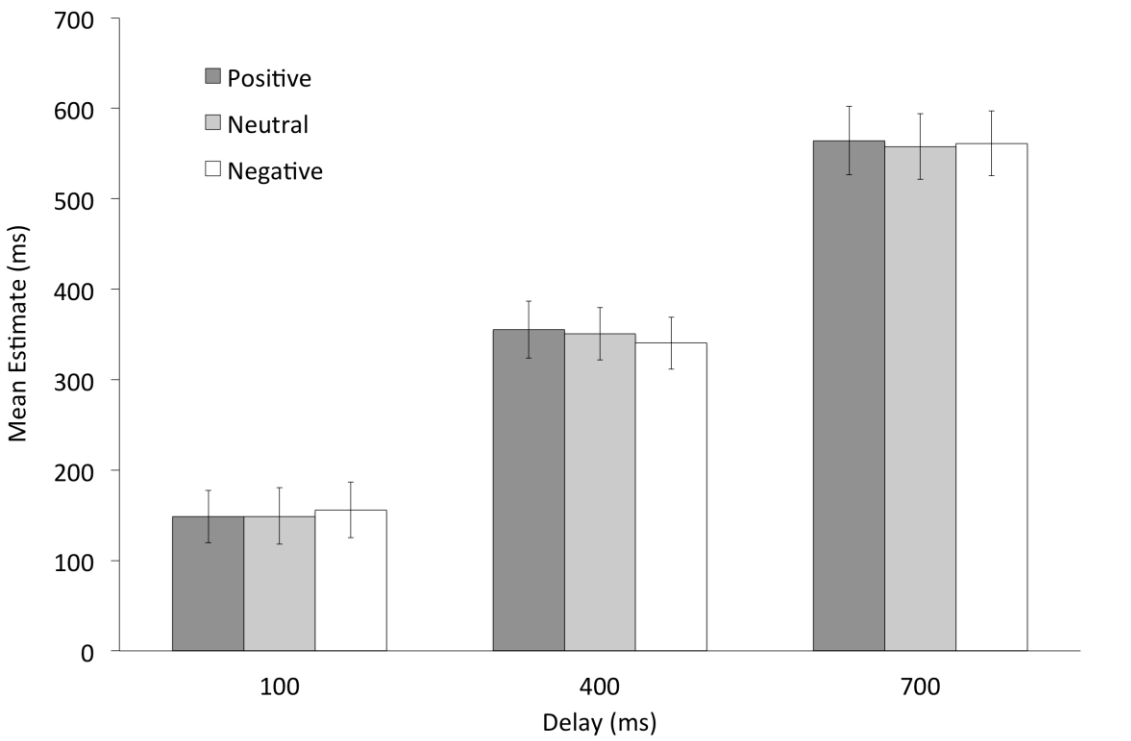
Results

Participants' mean time estimations for each of the three onset times (100, 400 and 700ms) and the three emoticons (positive, neutral and negative) were subjected to a 3 (emotional valence: positive, neutral, and negative) X 3 (temporal delay: 100, 400 and 700ms) fully within-subjects ANOVA (see Figure 4). Analysis revealed a significant main effect of Temporal Delay, $F(2, 158) = 56.54$, $p < .001$, $\eta p^2 = .77$, showing that even under less controlled experimental contexts (i.e., within an online testing platform), participants perceived distinct time intervals corresponding to their actual length (see Dewey & Knoblich, 2014, for comparable findings within a laboratory context). There was virtually no effect of emotional valence on time estimations, $F(2, 158) = 0.22$, $p =.80$, $\eta p^2 = .003$, nor was there an interaction between temporal delay and emotion, $F(4, 316) = 1.47$, $p = .21$, $\eta p^2 = .018$.

Discussion

In Experiment 1, the emotional valence of action outcomes did not affect temporal binding. One potential limitation of Experiment 1 is that although previous research has shown that emoticons can have the same affective consequences as real faces do (Yuasa, Saito & Mukawa, 2006), the emoticons we used might not have elicited enough of an emotional response to modulate temporal binding. Thus, rather than using emoticons for action outcomes, In

Experiment 2 we replicated our Experiment 1 procedure using images of real human faces expressing either negative or positive emotions.



Chapter 2, Figure 3. Mean time estimations of interval delay by emotional expression (Experiment 1). Error bars show 95% confidence intervals of the means.

Experiment 2

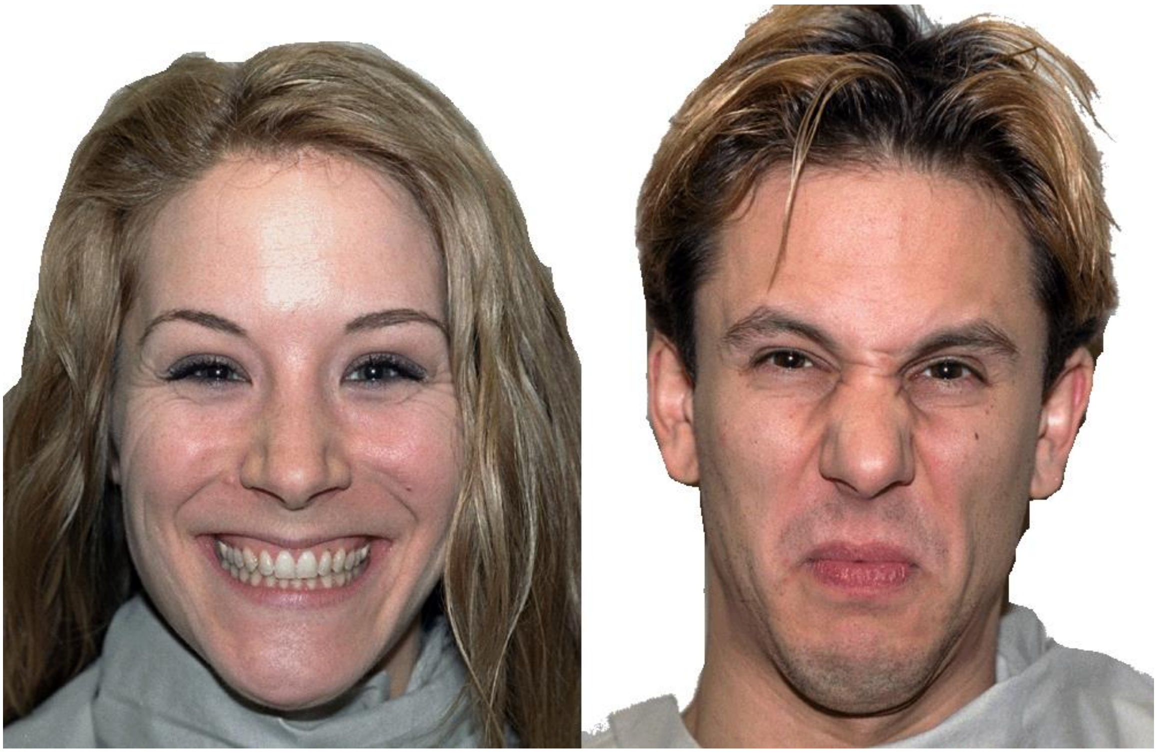
In Experiment 2, we used real-face images as the outcomes to participants' actions. Real face images have been well-documented to elicit electrocortical responses, and emotional expressions are typically rated along the dimensions of valence and arousal: Smith, Weinberg, Moran and Hajcak (2013), using the NimStim collection of face-images (NimStim, Tottenham, Tanaka, Leon, McCarry & Nurse et al., 2009), found that emotional expressions (e.g., happy, fearful, sad), elicited greater cortical responses than neutral face images. Generally, both negative and positive emotions invoke stronger emotional responses than faces with neutral expressions (Ito, Cacioppo & Lang, 1998), however the current literature suggests negative emotions elicit stronger cortical responses than positive emotions (Leppänen, Kauppinen, Peltola & Hietanen, 2007; Smith, Cacioppo, Larsen & Chartrand, 2003).

Method

Participants. We recruited 89 participants (55 males: M_{age} = 33.73, SD_{age} = 10.74) for Experiment 2. An additional participant was excluded due to a technical problem. Participants received monetary compensation. A power analysis showed that we had 95% power to detect a small effect (Cohen's f = .10) of emotional valence on temporal binding (α = .05).

Materials and procedures. Experiment 2 consisted of 110 trials: 30 practice trials, and 80 experimental trials (see Figure 5). To prepare participants for the experimental procedure, we asked participants to initially perform a practice task consisting of 10 trials where participant actions produced a neutral stimulus (the green circle). Similar to Experiment 1, during practice trials the time interval for the stimuli to appear was randomly selected from either 0ms, or a multiple of 100ms, up to 900ms. Participants were provided with feedback per Experiment 1.

Outcome stimuli consisted of 80 face images of young adults either portraying positive or negative expressions, taken from a widely used and validated set of face stimuli (NimStim, Tottenham et al., 2009). The facial images were balanced for gender, such that 10 males and 10 females were randomly chosen from the set (see Figure 5). Four facial images per male/female were chosen: two depicting positive facial emotions, and two depicting negative facial emotions (80 images in total, 4 x 20). The positive facial emotions included 40 images of a happy expression comprised the positive facial emotions, and 36 images of disgust and 4 images of fear expressions for the negative. Images were presented on a white background in the center of the screen. For the initial practice trials, we used the same neutral stimulus (green circle) as Experiment 1.



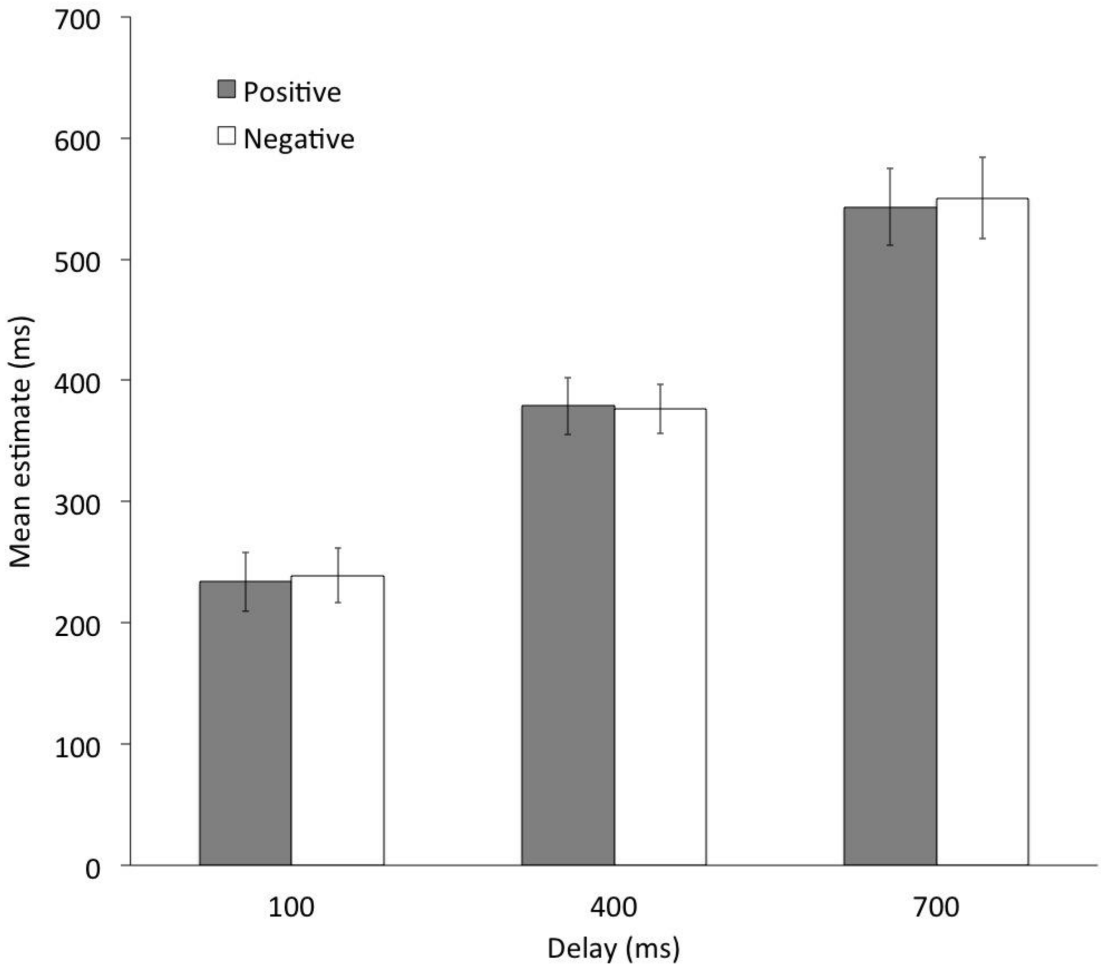
Chapter 2, Figure. 5. Example stimuli used in Experiment 2.

Participants underwent two experimental task blocks of 40 trials each, with a break between blocks. Each block was dedicated to either solely positive expressions or negative expressions, and the order of task blocks was counterbalanced between participants. Therefore, action-effects were predictable within their own blocks. Furthermore, participants were instructed that they would not receive feedback for their time estimations. The time interval for face images to appear was randomised at 100ms, 400ms, or 700ms (Moore et al., 2009), with the same number of trials in each condition at each time interval. A practice block of 10 trials that contained stimuli of the related task block preceded each experimental block. Upon block completion, participants were instructed that they would be asked to complete another practice task where they will see a different set of images, receiving feedback with their time estimations.

To incentivize participant to attend to the face stimuli, we also implemented catch-trials by informing participants that they would also be occasionally asked a question about the image they had just seen (specifically, "Was the previous face male or female?"). If they were correct, then they would be awarded an extra 10 pence per correct question. There were six catch trials in total - three trials per experimental condition. Seventy-six participants (84%) scored correctly on all catch trials, 8 participants (9%) scored correctly on 5 catch trials, and the remaining 6 participants scored correctly on 4 catch trials.

Results

We averaged time estimations for each of the three onset times (100, 400 and 700ms) and for each of the two levels for face-expressions (happy and disgust). We conducted a 2 (emotional valence: positive and negative) X 3 (temporal delay: 100, 400 and 700ms) fully within-subjects ANOVA. Analysis revealed a significant main effect of temporal delay, $F(2, 176) = 225.75$, $p < .001$, $\eta p^2 = .72$ (see Figure 6). Consistent with Experiment 1, there was virtually no effect of emotional valance on time estimation, $F(1, 88) = 0.092$, $p = .76$, $\eta p^2 = .001$. There was also no significant interaction between temporal delay and emotion, $F(2, 176) = .63$, $p = .53$, $\eta p^2 = .007$.



Chapter 2, Figure. 6. Mean time estimations of interval delay by emotional expression (Experiment 2). Error bars show 95% confidence intervals of the means.

Discussion

Similar to Experiment 1, the findings from our second experiment indicated no modulation of negative versus positive emotions on temporal binding. This is despite the use of real facial images depicting emotional expressions (as opposed to emoticons), and the predictability of which emotion-expression (either positive or negative) would result from the participant's action.

For both Experiments 1 and 2, we failed to find any meaningful effect of emotion on temporal binding, which seems inconsistent with earlier findings. One potential issue with our first two experiments, however, is that the stimuli we used for the positive and negative action outcomes (emoticons and real faces) might be perceived as less positively and/or negatively valenced than the sound vocalisations that Yoshie and Haggard (2013) used and therefore produce weaker temporal binding effects. To validate our stimuli, in Experiment 3 participants rated the emotional valence and arousal of the emoticons and faces we used in Experiments 1 and 2 and the positive and negative sound vocalisations that Yoshie and Haggard used.

Experiment 3

Method

Participants. Forty-nine participants were recruited via Amazon's Mechanical Turk (25 males, M_{age} = 34.80, SD_{age} = 11.56). To ensure data independence, one additional participant was not included in the analyses because they had a duplicate IP address.

Materials and procedures. Participants were informed that they would rate several images of faces and sound vocalizations in terms of how negative-to-positive and emotional arousing they appeared or sounded, respectively. Participants first performed a sound check that asked them to identify three different sounds (e.g., a cow mooing) from three choices (e.g., a pig's oink, a cow's moo, or a chicken's cluck) in order to ensure participants both could hear the sounds properly and were paying attention. All respondents saw the emoticons used in Experiment 1, all 80-face expressions used in Experiment 2, and heard 24 sounds (three repetitions of the 8 different sounds). The sounds were the same as those used by Yoshie and Haggard (2013), which were a selection of 8 different non-verbal emotional vocalizations: four negative vocalizations (screams expressing fear or retches expressing disgust, each with both male and female voices) and four positive vocalizations (cheers expressing achievement or laughs expressing amusement, each with both male and female voices). The block order of which type of stimulus the participants rated was randomly determined, and the stimuli presented within those blocks was randomised. Using the same rating scales as Yoshie and Haggard, after seeing/hearing the stimulus, participants judged the extent to which each stimulus looked (for the images) or sounded (for the vocalisations) negative-to-positive, on a 7-point scale ranging from 1 (*highly negative*) to 7 (*highly positive*). Participants also rated the extent to which they believed each stimulus sounded or looked emotionally arousing (1 = *not arousing at all* to 7 = *highly arousing*).

Results

Ratings of valence and emotional arousal were averaged across the different positive and negative faces and sounds. Because we were primarily interested in determining whether the different stimuli were perceived to be of equivalent valence, we conducted a one-way ANOVA with stimulus type on three levels (emoticons, faces, and vocalisations) separately for positive and negative stimuli. Shown in Table 1, there was a significant main effect of stimulus type in terms of perceived valence for both positive stimuli, $F(2,96) = 15.21$, $p < .001$, $\eta p^2 = .24$, and negative stimuli $F(2,96) = 22.44$, $p < .001$, $\eta p^2 = .32$. Paired sample *t*-tests revealed that the happy emoticon was rated as significantly more positive than the positive vocalizations, $t(48) = 3.52$, $p = .001$; there was no significant mean difference between the positive faces and positive vocalisations in terms of perceived valence, $t(48) = 1.95$, $p = .057$. For the negative stimuli, the negative vocalizations were rated as more positive (less negative) than both the sad emoticon, $t(48) = 5.76$, $p < .001$, and the negative faces, $t(48) = 2.70$, $p = .01$. Thus, the emoticon and face stimuli we used in Experiments 1 and 2 were perceived as either the same or more emotionally-valenced than the sound vocalisations used by Yoshie and Haggard (2013).

Chapter 2, Table 1. Mean (SD) ratings of the perceived emotional valence and arousal across the emoticons, face images and emotion vocalizations.

Stimulus type	Emoticons	Faces	Vocalisations
Negative stimuli			
Valence	1.45 (.58)	1.85 (.49)	2.18 (.64)
Arousal	2.41 (1.67)	2.71 (1.47)	2.81 (1.55)
Positive stimuli			
Valence	6.08 (.70)	5.46 (.52)	5.68 (.77)
Arousal	4.43 (1.83)	4.15 (1.32)	4.28 (1.45)

Note. Means that do not share subscripts across rows are significantly different ($p < .05$).

We also conducted a one-way ANOVA with stimulus type on three levels (emoticons, faces, and vocalisations) separately for positive and negative stimuli for perceived emotional arousal. There were no significant differences among the types of positive stimuli for the ratings of emotional arousal, $F(2, 96) = 1.44$, $p = .24$, $\eta p^2 = .03$. For the negative stimuli, $F(2, 96) = 3.70$, $p = .028$, $\eta p^2 = .07$, the negative vocalizations were rated as more arousing than the sad emoticon, $t(48) = 2.31$, $p = .025$, but were no more arousing than the negative faces, $t(48) = .79$, $p = .43$.

Discussion

The findings of Experiment 3 indicate that the visual stimuli used within Experiments 1 and 2 and the audio stimuli of Yoshie and Haggard (2013) were by and large rated similarly across dimensions of perceived valence and emotional arousal. More specifically, the positive emoticon was rated as more positive and more emotionally arousing than those of real faces and emotionally valenced vocalisations. Similarly, the negative emoticons and the negative faces were rated as more negative than the vocalisations. As such, the failure to find the predicted modulation of temporal binding by emotion in Experiments 1 and 2 does not seem to be driven by differences in the emotional appraisal of the stimuli.

Experiment 4

Because we did not find an effect of emotion on temporal binding in Experiments 1 and 2, we conducted a direct replication of Yoshie and Haggard (2013) to investigate the replicability of their findings.

Method

Participants. We recruited 24 participants to achieve 95% power to detect Yoshie and Haggard's reported effect size for their Experiment 1 (dz = .77): 12 males and 12 females (aged 18-23: M age = 21), one for each of the 8 (2 x 2 x 2) possible orders of conditions. Participants were paid for their time. Following Yoshie and Haggard (2013), we screened for the following exclusion criteria: native language other than English, left handedness, recent use of illicit drugs, uncorrected visual or auditory impairment, and history of psychiatric or neurological illness.

Materials and procedures. Experiment 4 used the exact same auditory stimuli as Yoshie and Haggard (2013). The stimuli were a selection of non-verbal emotional vocalizations, previously validated in the native English population to significantly differ in perceived valence, but not in perceived arousal (Sauter, Eisner, Calder, & Scott, 2010). In the negative condition, each participant's keypress was followed by one of four negative vocalizations (screams expressing fear or retches expressing disgust). In the positive condition, these were replaced by positive vocalizations (cheers expressing achievement or laughs expressing amusement). The auditory stimuli in each condition were carefully matched for pitch (peak frequency) and duration.

This experiment faithfully replicated the same procedure used by Yoshie and Haggard (2013). We presented the experiment via Macintosh computers (OS X 10.9.5), and used a customised program running in Inquisit v4.01 (Draine, 1998; Millisecond Software) to present participants with the temporal binding task on a 27-inch flat screen. We used the Libet clock task to measure the perceived timing of actions and sounds. During the experiment, participants viewed a Libet clock. In *agency* conditions, the participant was instructed to press a key on a computer keyboard with the right index finger at a time of his/her choosing, which caused a sound to appear 250ms later. The participant was then prompted to report where the clock hand was at the onset of their key-press or (agency action condition), in a separate block, at the onset of the sound (agency sound condition). In the single-event *baseline action* condition, the participant pressed a key at a time of his/her choosing. This keypress did not cause a sound, and the participant was asked to judge the time of his/her keypress. In the single-event *baseline sound* condition, the participant heard sounds at random intervals, which mimicked time intervals of participant key-presses, and judged the times of sound onsets. To make sure that participants understood the task, we asked participants to perform 5 practice trials before each condition.

Participants underwent four task blocks of 32 trials each (baseline action, baseline sound, agency action, and agency sound) for both the negative and positive conditions, or 256 (32 trials x 8 blocks) trials in total. In each block four different sounds of an emotional condition were presented in a randomized order (4 sounds x 8 repetitions). Since each block contained only positive or negative sounds, the four different vocalisations consisted of either the disgust and fear sounds, or the achievement and amusement sounds (each in both male and female voices). Each block was further divided into two sub-blocks of 16 trials each, with the stimuli randomised across the two sub-blocks, such that each sub-block could contain an uneven distribution of sounds. To ensure attention to the auditory stimuli, at the end of every sub-block we asked participants which of the four sounds they heard most frequently during that sub-block. Participants gained a reward of 25 pence for each correct answer to this question. The whole experiment was divided into two sessions of four blocks each. Each session was devoted to action judgments (baseline action and agency action) or sound judgments (baseline sound and agency sound) only. Half of participants (n = 8) judged the times of action in the first session and of sound in the second session, while in the other half (n = 8) the order was reversed. A 10-min break was inserted between the two sessions. To maximize the effects of emotional valence, within each session the baseline and agency blocks of one emotional condition (e.g., negative) were presented successively, and after a 5-min break the blocks of another emotional condition (e.g., positive). Both the order of emotional conditions (negative first or positive first) and the order of task types (baseline first or agency first) were consistent across the two sessions for each participant, and counterbalanced between participants (see Yoshie & Haggard, 2013).

Results

We used Yoshie and Haggard's (2013) protocol for extracting binding scores. Judgement errors were calculated individually for each block by subtracting the actual onset of the event with the perceived onset. Positive values reflect a delayed judgement, and negative outcomes reflect an anticipatory (early) judgement. Action binding (shift) was calculated by subtracting the mean judgement error of the action in the baseline condition from the mean judgement error in the agency condition. Similarly, sound binding (shift) was calculated by subtracting the mean judgement error of the sound in the baseline condition from the mean judgement of the sound in the agency condition. Composite binding was calculated by subtracting the mean shift in sound judgements from the mean shift in action judgements. Per Yoshie and Haggard (2013), paired *t*-tests (negative vs. positive) were used to assess the effects of emotional valence on temporal binding. We performed a Grubbs test for outliers (Grubbs, 1950), and no participant met the criteria for exclusion (all *ps* > .05). Additionally, we compared scores between positive and negative vocalisations on an attention task asking participants to state the most frequent sound within the preceding sub-block. A paired-samples t-test revealed no difference in participants' attention to sounds between negative (M = 3.83, SD = 1.34) and positive (M = 3.63, SD = 1.21) vocalisations, $t(23) = .96$, $p = .35$; $dz = .20$.

Chapter 2, Table 2. Mean (SD) judgement errors and shifts relative to baseline conditions between different emotion conditions.

	Action judgements			Sound judgements		
	Baseline (ms)	Agency (ms)	Shift (ms)	Baseline (ms)	Agency (ms)	Shift (ms)
Negative	-69.28 (110.56)	25.76 (114.54)	95.04	-206.02 (71.63)	-345.67 (154.51)	-139.66
Positive	-89.54 (129.84)	33.70 (141.43)	123.24	-190.55 (101.79)	-347.69 (158.95)	-157.14

Table 2 shows the mean judgment errors and shifts relative to baseline conditions for different emotional conditions. The presence of action binding was confirmed by a shift in judgement errors that was significantly different from zero for action judgements in both the negative, $t(23) = 2.94$, $p = .007$, dz = .60, and positive conditions, $t(23) = 3.78$, $p = .001$, $dz = .77$. Similarly, sound binding was also significant for both negative, $t(23) = 5.47$, $p < .001$, dz = 1.12, and positive vocalisations, $t(23) = 5.27$, $p < .001$, dz = 1.08. Composite binding did not differ significantly between the negative (M = -234.68, SD = 174.71) and positive conditions (M = -280.38, SD = 134.20), $t(23) = 1.20$, $p = .24$, $dz = .24$. Similarly, paired t-tests revealed no significant difference in sound binding, $t(23) = .64$, $p = .53$; $dz = .13$, or action binding, $t(23) = 1.16$, $p = .26$, $dz = .24$, between the positive and negative conditions.

Discussion

The findings of Experiment 4 suggest that temporal binding, as measured using the Libet clock method, was not significantly modulated by positive versus negative sound vocalisations as action outcomes. These results taken together with those of Experiments 1 and 2 suggest that the emotional valence of action outcomes exerts little influence on temporal binding.

General Discussion

The objective of this series of experiments was to investigate the degree to which temporal binding is modulated by emotional valence. Experiments 1 and 2 found no significant difference in temporal binding between positive and negative emoticons (Experiment 1) or positive and negative real facial expressions (Experiment 2). Experiment 3 revealed that the stimuli used in Experiments 1 and 2 were equivalent in valence and arousal to stimuli that have previously been observed to modulate temporal binding (Yoshie & Haggard, 2013). Furthermore, in a highly powered replication study (Experiment 4), we observed no significant modulation of temporal binding by emotionally valenced vocalisations (Yoshie & Haggard, 2013). Taken together, these finding cast doubt on whether temporal binding is influenced by outcome valence.

Despite showing no significant modulation by valence, temporal binding itself was clearly present in Study 4. Indeed, the binding scores were overall somewhat larger than Yoshie and Haggard’s (2013). This suggests that the absence of a valence effect in our study was not due to reduced sensitivity to detect emotional modulation. Although not significant, the effect of valence on binding was in the predicted direction in the current study. However, it is worth noting that this was largely driven by greater action binding to positive tones, whereas Yoshie and Haggard’s (2013) effect was more strongly localised on outcome binding. More recently Christensen et al. (2016) investigated the effect of outcome valence on prospective and retrospective components of action binding (see

Moore & Obhi, 2012) with the same vocalisations used here and in Yoshie and Haggard (2013). They observed significantly increased retrospective action binding only when the valence of the outcome was unpredictable. However, for predictable outcomes (as used in the current study) there was reduced action binding for both positive and negative outcomes compared to neutral outcomes. Taken together with the current findings, a complex picture emerges whereby the precise effect of emotion on temporal binding cannot be clearly attributed to a simple self-serving bias such that positive outcomes increase binding. This may reflect a genuine complexity in the precise mechanisms driving the emotional modulation of binding, or it might reflect the fact that the underlying effect is small or unreliable. The absence of an effect of valence in Experiments 1 and 2 suggest that any effect, if present in the population, does not generalize to other measures of binding. Future work should attempt to replicate and extend other examples of self-serving bias in temporal binding (Aarts et al., 2012; Takahata et al., 2012) and sensory attenuation (Gentsch, Weiss, Spengler, Synofzik & Schütz-Bosbach, 2015; Hughes, 2015) to further advance our understanding of how (or if) outcome valence influences implicit agency.

Assessing the degree to which binding is modulated by factors that also modulate explicit agency reports is important to determine the relationship between implicit and explicit agency. Recent evidence suggests that neither sensory attenuation (Dewey & Knoblich, 2014) nor temporal binding (Dewey & Knoblich, 2014; Saito, Takahata, Murai & Takahashi, 2015) correlate with explicit reports of agency. While explicit and implicit measures will never show total convergence, positive evidence of covariation is important to argue that conscious reports and unconscious biases are indeed measuring the same underlying process. The current studies provide new evidence that questions the degree to which temporal binding is modulated by self-serving biases.

Chapter 3: The Use of Temporal Binding as an Implicit Measure of Causality in Immanent Justice Reasoning

Preface

The experiments contained within Chapter 2 were conducted in order to obtain a baseline effect that temporal binding, and by extension sense of agency, could be modulated by social factors. We strongly expected, based on prior findings (e.g., Yoshie & Haggard, 2013), that we would find increased temporal binding for positive outcomes relative to negative outcomes. However, no modulatory effects were found. Despite this finding, we decided to persevere with the original aims of this research for three reasons: 1) the relationship between emotional valence of outcome and causality between events was not specifically tested, and thus there remained scope to explore how personal bias might influence proposed implicit measures of causality (e.g. temporal binding); 2) compared to emotional valence, the immanent justice reasoning literature contains a wealth of research demonstrating how explicit causal attributions can be modulated by the presence of morally congruent events occurring to the same person (i.e., bad people causing bad things to happen to them; Callan et al., 2006; Callan et al., 2012). Thus, with a genuine implicit measure of causality, we might expect similar modulatory effects; 3) where Chapter 2 concerned physical causality between the participant's button press and a visual outcome occurring, we wanted to explore whether temporal binding reflects biased judgments about the timing or spatial position of abstract representations of socially relevant acts. In this line of research, we expand upon notions that temporal binding is a measure of general causality (Buehner, 2012).

The Use of Temporal Binding as an Implicit Measure of Causality in Immanent Justice Reasoning

Understanding causal relations within our environment is a fundamental skill for the human species that enables, for example, social interaction (Bodner, Engelhardt, Minshew & Williams, 2015; Moskowitz, 2005), learning (Gopnik, Sobel, Schulz & Glymour, 2001), reasoning (Goswami & Brown, 1990) and self-awareness (Duval, Silvia & Lalwani, 2012; Duval & Wicklund, 1973). Understanding causal relations, put simply, is the process of identifying causality between events; the reason as to why a certain event or events has occurred. Such an understanding elicits the ability to alter the course of events to ensure unwanted outcomes are avoided; knowing a lack of oil in the family car would cause the engine to seize, we actively maintain the oil gauge to a healthy level.

Data collection and analysis allows humans, already capable computing complex causal matrices of probabilistic causes that contribute to events or outcomes (Cheng, 1997), further enhances our knowledge of causal relations. For example, medical research has shown smoking accounts for an estimated 72% of lung cancer cases (Cancer Research UK, Lung Cancer Risk, 2018).

However, when such data is missing, we are left with events of ambiguous origin. This can lead us to explain events from our learned experiences, and such a causal reasoning process is subject to biases in favour of less-informed, if not whimsical, explanations of events (Callan, Sutton, Harvey, & and Dawtry, 2014). As humans are infamously susceptible to biases and ideologies, particularly in conjunction with illiteracy of scientific or logical principles, we may infer or adopt erroneous causal beliefs to explain phenomena, even at the expense of physical plausibility. Despite the lack of logical/physical connection between two events, the notion that one's moral behaviour not only influences, but is a direct cause of, subsequent events in our lives persists in the form of Immanent Justice Reasoning (IJR; Callan, Ellard, & Nicol, 2006; Callan, Sutton, & Dovale, 2010; Callan, Harvey, Dawtry, & Sutton, 2013; Harvey & Callan, 2014).

IJR is defined as a lapse in naturalistic causal reasoning where actions bring about deserved outcomes, be they rewards or punishments, even when there is no physically plausible means by which they might have done so (Callan et al., 2006; Callan et al., 2014). Although linked with concerns for the deservingness of the outcome (e.g., bad people deserve bad things to happen to them), IJR takes this notion a step further where the moral worth of the person or behaviour is the direct cause of the outcome in the face of more naturalistic explanations (Callan et al., 2006; Harvey & Callan, 2014). Thus, IJR is a form of causal reasoning bias that enables us to preserve the belief that the world is predictable, just and orderly (Callan et al., 2014).

According to Lerner's (1980) just world theory, IJR serves as a function to maintain just world beliefs. Not only are we biased in our causal judgements, but we particularly are motivated to prefer explanations that help defend against threats to our belief in a just world. Recent research shows that construing events to be consistent with a just world allows us to make long-term goals (Bal & van den Bos, 2012; Callan et al., 2013; Xie, Liu, & Gan, 2011), avoid self-defeating behaviours (Callan, Kay, & Dawtry, 2014), and avoid smaller, immediate rewards in order to obtain larger, delayed rewards (Callan, Shead, & Olson, 2009). However, IJR research often utilises self-report methods to measure the strength of the causal link between someone's prior (mis)deeds and a future event. For example, reading a vignette detailing a fatal accident had occurred to either a good or bad person, followed by the question ‘‘to what extent do you feel that what happened to xxx was a result of his/her conduct?” (Callan et al., 2013). Callan, Ferguson and Bindemann (2013) showed that motivations to perceive justice exist beyond retrospective causal judgements using eye-tracking. Participants listened to a recorded vignette describing either a good or bad person through headphones, and their eye-gaze was recorded whilst two possible outcome images (good or bad) were displayed partway through the vignette. Callan et al. found that participant gaze shifted towards the morally congruent image, such that when listening to the bad (good) behaviour of a person, eye-gaze focused onto the bad (good) outcome. These findings suggest not only are we motivated to retrospectively assert morally congruent reasons for specific outcomes, we also anticipate morally congruent outcomes when justice is concerned.

The current research aims to extend previous IJR findings by exploring whether valence-laden sentences within a scenario can impact our temporal perception of the interval between two events. This notion stems from a phenomenon known as temporal binding, where causally linked events are perceived closer in time (Engbert, Wohlschläger & Haggard, 2008; Haggard, Clark, & Kalogeras, 2002; Moore & Fletcher, 2012; Moore, Wegner & Haggard, 2009). Temporal binding is rooted within David Hume’s (1739/1888) Treatise of Human Nature, where temporal contiguity between events underlies a causal relationship. In other words, the closer two events are in time, the more likely they are to be causally related. However, recent research demonstrates that not only does contiguity promote inferences of causality, but also the perception of a causal relationship leads one to perceive greater contiguity (Buehner et al., 2009).

Typically, in the time estimation version of temporal binding, two events are displayed sequentially, and the time between two events is estimated. This version of temporal binding is measured by asking participants to estimate the interval between the two events (such as the action and the outcome), using either a verbal estimate (Engbert et al., 2008; Kumar & Srinivasan, 2017; Moore et al., 2009) or an estimate via a time scale (Moreton, Callan, & Hughes, 2017).

Interval estimation offers a direct method of measuring temporal binding and has been shown to depend on causality: for example, Humphreys and Buehner (2009) proposed that our internal pacemaker (Wearden, 2001) slows during periods in which we predict the consequences to intentional, causal action. Indeed, the authors showed that the temporal interval between a causal event (pressing an on-switch and an ensuing sound) is significantly underestimated when compared to a non-causal event. Buehner and Humphreys (2009), also demonstrated binding when events are believed to be causally related, and suggest that mentally connected constructs, in that they are causally related, and perception, are inter-related.

However, temporal binding research typically involves the causal relation between an individual's action causing an outcome stimuli to occur, such as the participant performing a button press and causing an auditory tone or visual image to appear after an interval between 50-1000ms (Buehner, 2012; Engbert et al., 2008; Haggard & Clark, 2003; Haggard et al., 2002; Humphreys & Buehner, 2010; Moore & Haggard, 2008; Takahata et al., 2012). In other words, such research has revolved around physical events occurring within a short time frame of one another. Conversely, events within IJR research are typically abstract and occur after long periods (a common connecting phrase being "Later that day…").

The current research aims to extend temporal binding research beyond physical events. The temporal binding phenomenon is arguably rooted in causal beliefs; for example, Buehner (2012) demonstrated that by believing a machine to be its own causal agent, the outcomes of its actions were anticipated comparably to human actions. If a causal relationship exists between two events, then, given temporal contiguity, the two events are more likely to follow each other closely in space and time than unrelated events. Thus, our research affords a test of whether causally connected scenarios (via notions of deservingness) would be subject to the same consequences to temporal contraction as action/outcome pairings.

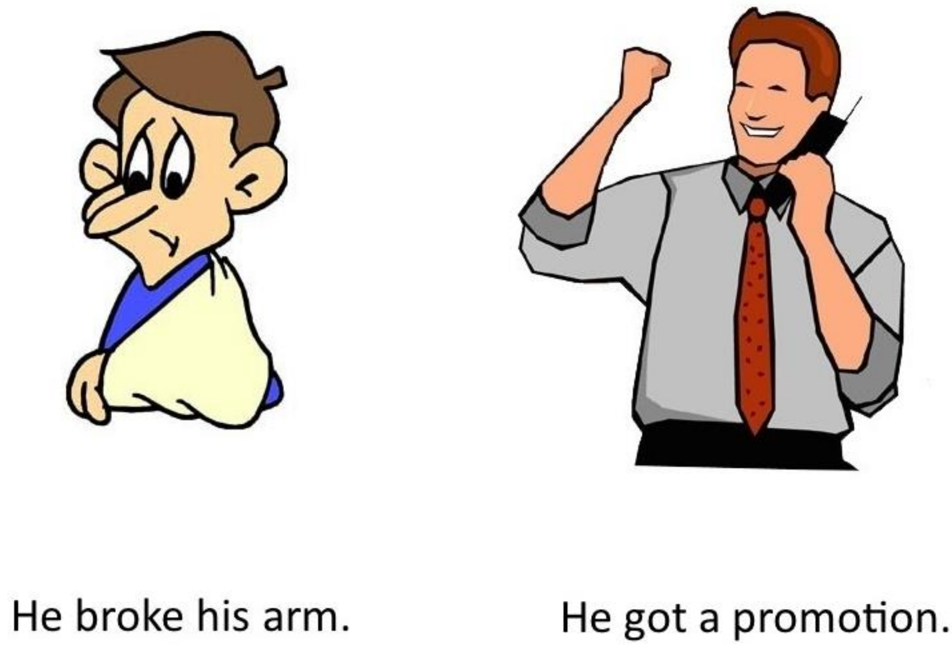
Our expectations were weighted upon an overarching causal theory of temporal binding and research within IJR, where descriptions of a good(bad) person followed by a good(bad) outcome will elicit a causal relation, generating contiguity and estimated closer in time. Conversely, morally incongruent sentences, and hence less causally related, should be estimated more relatively, temporally, distant.

Present Study

We used an interval estimation procedure to gauge temporal binding (Ebert & Wegner, 2010; Engbert et al., 2008; Moore, Wegner, & Haggard, 2009). In this procedure, participants are asked to judge the time interval between an action and its sensory outcome (e.g., a button press and an image). For our task, participants were presented with the two possible outcomes to a scenario, each with their own short description of what 'had happened'. This served to give participants prior knowledge of what could happen following the events of the scenario. Prior knowledge of the possible outcomes meant that participants would retain the outcomes consciously during their reading of the scenario. In turn, reading the scenario would build up "just world" expectations about the forthcoming event (Callan et al., 2012). Thus, in terms of immanent justice reasoning, we would expect that prior knowledge would motivate the desire for the specific morally congruent outcome to appear post-scenario. This desire would then generate outcome identity predictions consistent with immanent justice reasoning, providing both the predictive and inferential processes that generate temporal binding (Moore & Ohbi, 2012).

Then, participants were presented with a scenario that depicted a person with either a positive or negative character (with an equal male-to-female ratio throughout the scenarios). For example, "While walking home after work, Simon offered to carry an elderly woman's bags and helped her cross the road." for positive, and "While walking home after work, Simon yelled at an elderly woman to get out of his way and shoved her into the gutter as he passed by." for negative. We used 36 scenarios in total – that is, an event that occurred to the same person (e.g., "Simon"). As there were both positive and negative versions of each scenario, however, there were a total of 72 possible sub-scenarios (see Appendix A). However, participants only saw one version of each scenario and pseudo-randomised to ensure equal presentation of valence.

Every scenario ended with "Later that day…" in order to produce a complete sentence with the outcome-stimuli's descriptions. The participant was then asked to press the space bar, which was followed by either positively or negatively valenced action-outcome—specifically, outcomes depicting something positive or negative happening to the protagonist (see Fig. 1 for examples). For example, one possible outcome was for the protagonist to receive a promotion, therefore the completed, “Later that day he got a promotion.”



Chapter 3, Figure 1. Example outcome stimuli.

We used five images to represent a positive outcome, and four images to represent a negative outcome. This slight difference was due to the availability of images depicting a clear, unambiguous outcome. Equally, this was due to the availability of gender-related images. Thus, for positive outcomes, there were four gender-related images, (two for each gender; for example, a man cheering), and one neutral image (an image of a sum of money). For negative outcomes, there were two gender-related images (one for each gender; for example, a woman slipping on a wet floor), and two neutral images (a traffic accident and snake poised to strike a victim; for a full visual list of images). Furthermore, the scenarios were adapted from scenarios used previously within IJR research and have been shown to be effective at eliciting IJR expectations of a just world (Callan, Ferguson & Bindemann, 2013). We expected time estimations of the temporal interval between scenario and outcome to reflect explicit measurements of IJR: scenario-outcome combinations congruent in their valences (for example, a scenario describing a good person receiving an unrelated good outcome) would produce greater temporal binding, whereas incongruent combinations would produce less.

Method

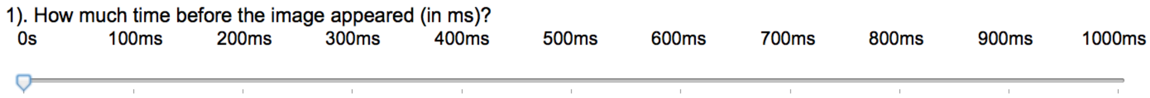
Participants. We recruited 89 participants (42 males: M_{age} = 36.23, SD_{age} = 11.68) through prolific.ac, an online crowdsourcing platform. Participants received monetary compensation. We screened participants for the following inclusion criteria: an approval rating of above 90% on prolific.ac (based on prior experiment performance/approval scores) and aged between 18 and 65. The required sample size was fixed ahead of data collection, and a power analysis showed we had 90% power to detect a small effect (Cohen's f = 0.10) of emotional valence on temporal binding (a = 0.05).

Materials and Procedure. Experiment 6 consisted of 77 trials: 20 practice trials, 9 exposure task trials, 9 memory task trials, 3 pre-experimental practice trials, and 36 experimental trials. We used an interval estimation procedure to measure temporal binding (see Moore & Obhi, 2012). For each practice and experimental trial, participants saw a fixation cross on the screen, and in their own time, pressed the spacebar. In the first practice block of 10 trials participant actions produced a neutral stimulus, which was a green circle, and in the second practice block of 10 trials actions produce an image of an emotionally valenced emoticon, either positive (smiling) or negative (angry). During practice trials, the green circle/emoticons appeared after a randomly selected time interval from either 0 ms or a multiple of 100 ms up to 900 ms. We used all intervals in the practice blocks, to encourage participants to expect the full range of durations in the experimental block. During the practice blocks, feedback was provided to participants after they made their time estimations. Feedback consisted of both the participant's estimated time and the actual time of stimulus onset to enhance familiarity with estimating time in milliseconds.

Participants were then randomly presented with the four negative and five positive outcome stimuli. Each stimulus appeared individually onscreen for four seconds, and afterwards was subsequently replaced with two Likert scales, ranging from 1-7, asking participants to "Please rate how positive/negative the image was." and "Please rate how emotionally arousing the image was." After completion, participants then completed a memory task. In this task, participants saw three outcome stimuli onscreen without their descriptions, and one description at the bottom that corresponded to one of the three stimuli. Participants were asked to select which of the images the description corresponded to. The mean score of remembering correct description to images was 6.07 for the nine images (*SD* = 2.39).

A pre-experimental practice block of three neutral scenarios, each with their own two neutral outcomes, preceded the experimental block, in order to introduce the mechanics of the outcome-scenario-time estimation task presentation. However, the task sequence was the same for both the pre-experimental practice trials and experimental trials: participants initially saw two possible outcomes to a forthcoming scenario. In experimental trials this was a positive and negative outcome, and either randomly presented on the left-hand/right-hand sides of the screen, including their description.

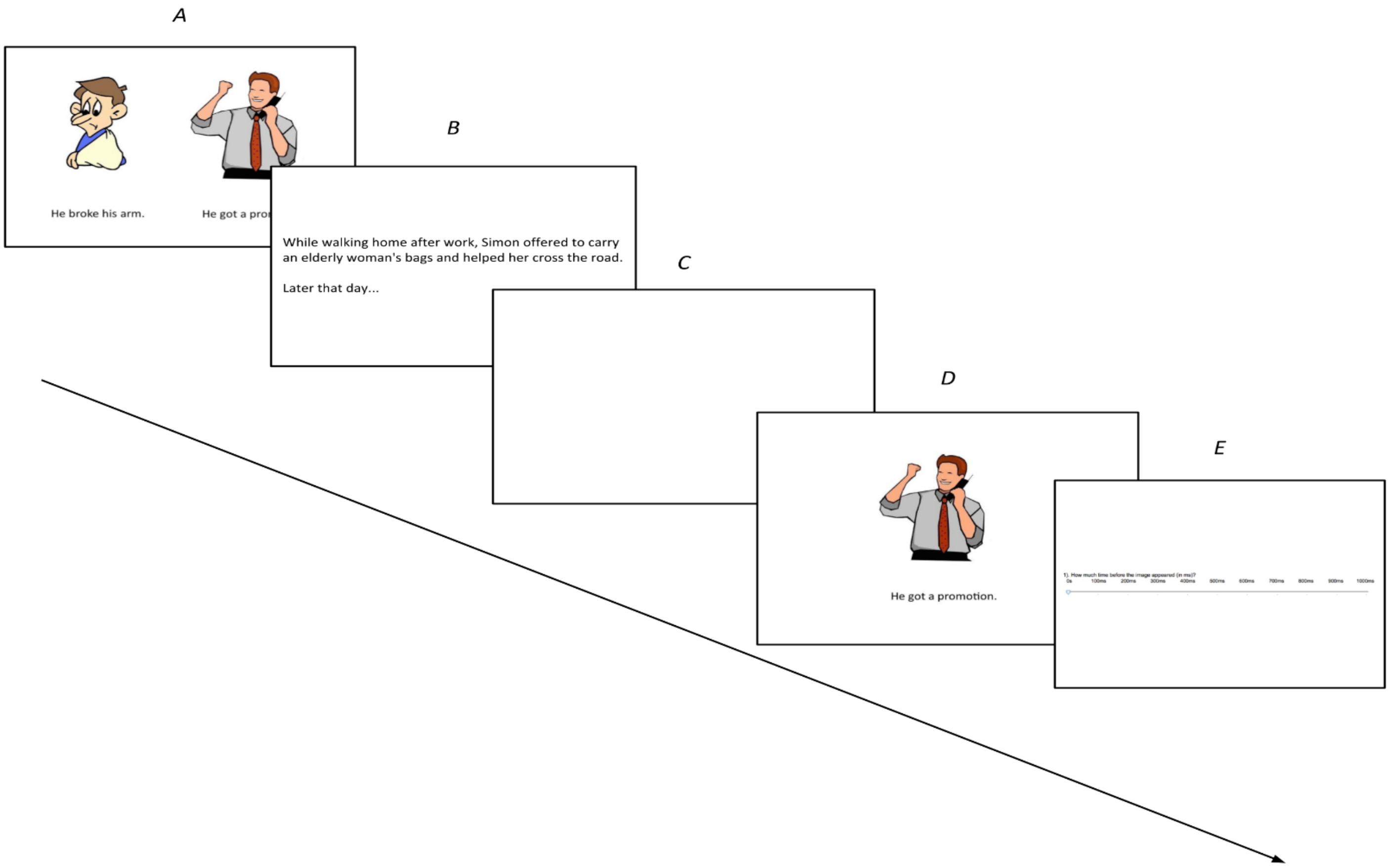
Next, participants saw one of thirty-six scenarios depicting an individual with either a good or bad character. The participant was able to respond by pressing the spacebar to continue after 2500ms after scenario presentation. This was set to prevent participants from either accidentally clicking past the stimuli, or from speeding through trials without paying attention to the scenario. To indicate when a participant was able to proceed, a grey button appeared beneath the scenario stating ‘Press the spacebar to continue’. After pressing the spacebar, one of the two outcomes presented directly before the scenario appeared after either 100, 400 or 700 ms (Moore et al., 2009), which remained on the screen for a further 400 ms. We varied the delay intervals to increase participants’ uncertainty regarding the interval between action and outcome to allow for variation in judgement times (cf. Ebert & Wegner, 2010). A blank screen then followed the outcome stimulus for 400 ms and was replaced by a horizontal time estimation scale in the centre of the screen (see Fig. 2). The scale ranged from 0-1000 ms, with demarcation lines every 100 ms. Participants were instructed to scroll the slider along the bar to the time that they believed it took the image to appear since their action (in multiples of 100 ms). Once selected, participants confirmed their selections by clicking on a ‘finish’ button and proceeded to the next trial. Participants were instructed that they would not receive feedback for their time estimations during the experimental trials.



Chapter 3, Figure 2. Time estimation scale.

During experimental trials participants were presented with the 36 scenarios, 1 scenario per trial (hence, 36 experimental trials). These scenarios were divided evenly, such that 18 scenarios were the version that depicted an individual with a positive character, and 18 with a negative character, randomly presented across trials. Valence of outcome stimuli, presented after the scenario, was also pseudo-randomised to be presented equally in the two types of trials (positive and negative characters). This resulted in a within 2 (positive vs. negative valence of the character in the initial scenario) x 2 (positive vs. negative valence of outcome stimuli) factorial design. Scenarios/stimuli were counterbalanced across participants, such that across all 36 trials each participant saw an equal number of the four possible scenario-outcome combinations (i.e., 9 trials per Good scenario-Good outcome, Good-Bad, Bad-Good, and Bad-Bad). For example, two participants presented with “While walking home after work, Simon offered to carry an elderly woman's bags and helped her cross the road.” (i.e. a good scenario) then being presented with the good or bad outcome (Good-Good/Good-Bad). Subsequently, another two participants would be presented with “While walking home after work, Simon yelled at an elderly woman to get out of his way and shoved her into the gutter as he passed by.” (i.e. the same protagonists, but now depicting a bad scenario) and be presented with either the good or bad outcome (Bad-Good/Bad-Bad).

Time intervals were pseudo-randomised across trials, such that there was the same number of trials in each condition at each time interval. A schematic display of the sequence of trial events is shown in Figure 3.



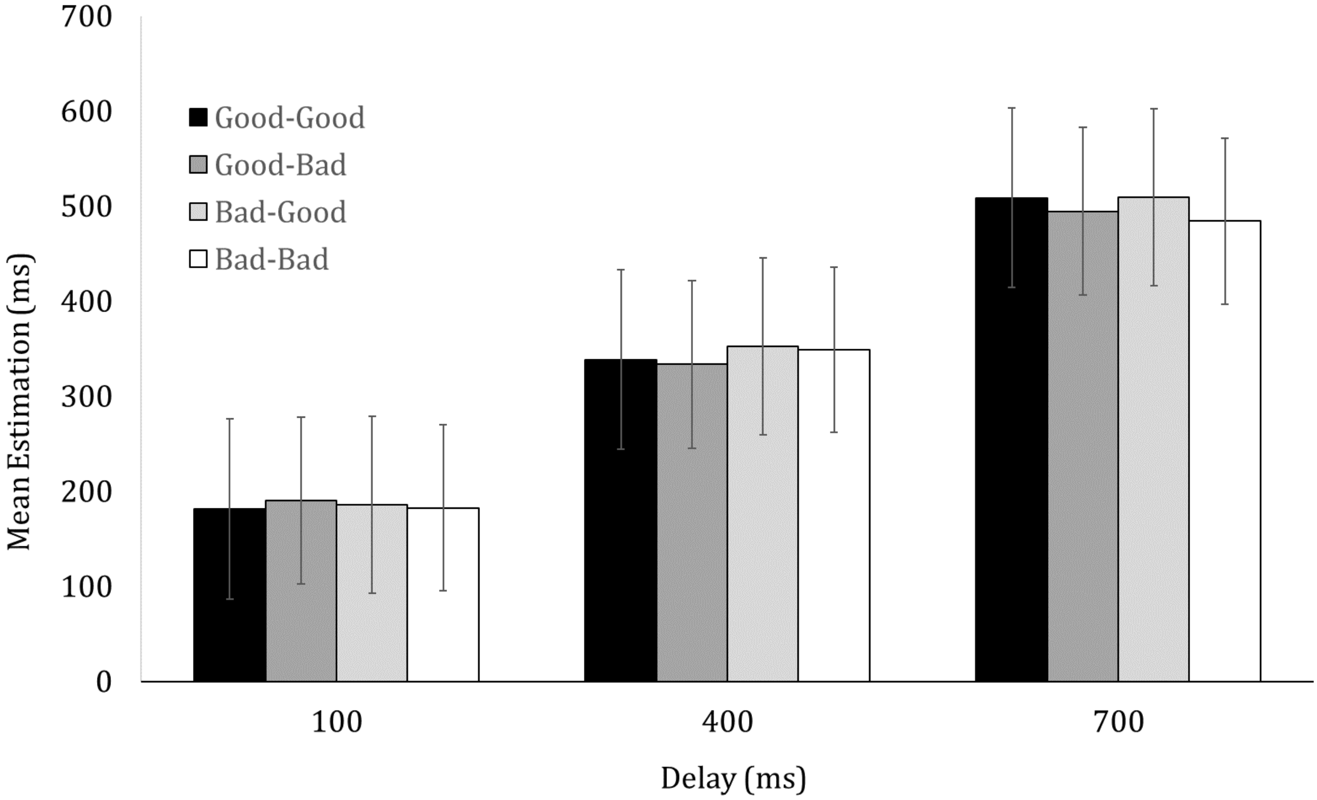
Chapter 3, Figure 3. Schematic display of the sequence of trial events: A) Potential outcomes (bad or good), B) Scenario (bad or good), C) Delay (100, 400 or 700ms), D) Outcome (400ms), E) Time estimation scale.

To incentivize participant to attend to the scenarios, we also implemented catch-trials by informing participants that they would also be occasionally asked a question about the scenario-vignette they had just read (for example, "Where did the scenario take place?"). There were six catch trials in total. 10 participants (11.2%) scored correctly on all catch trials, 39 participants (43.8%) scored correctly on 5 catch trials, 33 participants (37.1%) scored correctly on 4 catch trials, and 7 participants (7.9%) score correctly on 3 catch trials or lower.

Results

Firstly, ratings of valence for our stimuli were averaged across the positive and negative outcome stimuli, and a paired-samples t-test revealed that positive outcomes (M = 6.29, SD = .23) were rated significant more positively than negative outcomes (M = 2.06, SD = .39), t(89) = -26.54, p < .001, *d* = 13.21.

Participants' mean time estimations for each of the three onset times (100, 400 and 700 ms) and the four scenario-outcome valence combinations: good(scenario)-good(outcome), good-bad, bad-good, bad-bad were subjected to a fully within-subjects ANOVA (see Fig. 1). Analysis revealed a significant main effect of temporal delay, $F(2, 176) = 183.11$, $p < .001$, $\eta p^2 = .68$, showing that even under less controlled experimental contexts (i.e., within an online testing platform), participants perceived distinct time intervals corresponding to their actual length (see Dewey & Knoblich, 2014, for comparable findings within a laboratory context). There was no statistically significant effect of scenario-outcome valence congruency on time estimations, $F(2,264) = 0.51$, $p = 0.67$, $\eta p^2 = 0.006$, nor was there an interaction between temporal delay and scenario-outcome valence congruency, $F(6,528) =.62$, $p = 0.71$, $\eta p^2 = 0.007$.



Chapter 3, Figure 4: Table depicting the mean temporal estimations for each scenario-outcome valence combination by temporal delay. Error bars show 95% confidence intervals of the means.

Discussion

Events that are causally linked appear closer together in time (Faro et al., 2005; Buehner & Humphreys, 2009). However, within our study, we failed to demonstrate any temporal perception variability across our experimental conditions. Despite the widespread usage of temporal binding as a measure of implicit causal beliefs, the congruency of scenarios used did not influence temporal estimates. Our findings suggest that not only does our perception of time fail to be modulated by concerns for justice. Given the wealth of research showing the reliability of immanent justice reasoning on these types of scenarios, these findings also point the possibility that temporal binding may not be suitable as an implicit measure of causal beliefs.

Three possible conclusions explain the results of our experiment from a theoretical standpoint: 1) the experimental mechanics, i.e., use of a temporal binding measure between a vignette and visual outcome, were not appropriate to detecting implicitly held causal beliefs; 2) IJR is not susceptible to implicit measures of causal relations between two events, or at least, the effect size is vastly small in comparison to its relative success using explicit measures (Callan, Sutton & Harvey, 2014); 3) temporal binding, as has been purported (Hughes, Desantis & Waszak, 2013), is less of a measure of causal relations, but is rather much more susceptible to low-level processes such as temporal prediction and control. However, from a practical standpoint, our experiment may have had a limitation in that the temporal interval may have been estimated as the time between the button press and the outcome image, rather than from the scenario. This would render any modulatory effects due to causal beliefs unmeasured. As a follow-up test, future experiments may benefit from an observational, rather than action, oriented procedure, whereby participants estimate the interval between two events whilst passively observing the sequence. This would also fit in with previous research (Buehner, 2012), that suggests that temporal binding occurs due to the causal relation between events, rather than any action involved.

One foreseeable limitation to using abstract events concerning justice is that, due to the nature of the stimuli, a plausible connection, despite lacking explicit information, may still arise between some scenarios. For example, where "Jamie saw a homeless man with a bucket for change, and stole the homeless man's money", and later that day "He broke his arm", one could construe that, possibly while feeling excitement at his achievement, guilt over his actions, or perhaps simply whilst fleeing the scene of the crime, Jamie became careless or distracted and consequently broke his arm. Naturally, we would expect those who possess a need to believe in a just world, and are sensitive to threats that challenge this belief, would be more prone to adopting the IJR account of events rather than construing a more rational explanation (Hafer & Bègue, 2005; Lerner, 1980). Moreover, such a construal would be more probable with morally congruent events given the conceptual similarity between stimuli (Casasanto, 2008). Although near-impossible to remove every possible chain of events that one may conjure to link the two events, future work should make efforts when controlling for probabilistic links in their stimuli. For example, external raters could score the plausibility of one event causing the second event. Additionally, mixed effects modelling including fixed effects for perceived similarity between events could assess the unique predictive power of similarity upon temporal estimation over and above moral congruence.

A further limitation to our study concerned, without use of an explicit causal measure, the difficulty to be fully confident that any differences in interval estimations would have been due to engagement in immanent justice reasoning. Future work, where exploring novel implicit measures, should include explicit measures to enable correlational tests to improve confidence that the implicit measures are measuring the same underlying construct.

One experiment alone cannot provide enough strength to counter prior published research showing the modulatory effects of causal beliefs. However, what we do present is that, given the conflicting findings that surround temporal binding as to its underlying mechanisms (e.g., Ebert & Wegner, 2010; Moreton, Callan, & Hughes, 2017; Yoshie & Haggard, 2013; Zopf, Polito, & Moore, 2018); further inquiry may be required to ascertain temporal binding's accuracy as a measure of implicit causal beliefs. Spatial binding, the spatial equivalent to temporal binding, may procure more promising results after previous evidence demonstrates its relative consistency as a measure of causal beliefs (Buehner & Humphreys, 2010; Woods, Lehet, & Chatterjee, 2013). Given the strength of evidence behind IJR with explicit measurements (Callan et al., 2006; Callan et al., 2010; Callan et al., 2013; Harvey & Callan, 2014), we expect that a robust measure of implicit causal beliefs will yield similar results.

Chapter 4: The Role of Causal Knowledge in Temporal Binding Measured via Stimulus Anticipation

Preface

Although temporal binding is routinely employed as a measure of SoA, several studies question its proposed validity (Buehner & Humphreys, 2009; Moreton et al., 2018; Stetson et al., 2006). Chapter 2 suggests that, not only is temporal binding immune to modulatory effects of emotionally valenced outcomes, but also contributes to the ongoing discussion of conflicting findings within the literature (for example, regarding learned action-effects to congruent/incongruent actions, see Desantis et al., 2012, Ebert & Wegner, 2010, Hughes, 2018, and Zopf, Polito, & Moore, 2018). Chapter 3, though not without its limitations, makes a preliminary suggestion that binding may not result from causal mechanisms; or at least, abstract causality without a direct, physical mechanism. Thus, at this juncture of my research, we decided to metaphorically take a step back and approach the aim of this research by investigating the role of causal belief in temporal binding. Chapter 4 attempts to extend two previous studies (Experiment 6: Buehner, 2012; Experiment 7: Desantis, Roussel and Waszak, 2011) to further clarify whether knowledge of or belief in a causal mechanism modulates temporal binding. Specifically, whether temporal binding would increase when: a) a causal agent (the participant) produced an outcome relative to conditions without any apparent causal mechanism (Experiment 6), and b) authorship of the outcome is believed to have been caused by the participant relative to an external agent.

The Role of Causal Knowledge in Temporal Binding Measured via Stimulus Anticipation

Understanding causal relations within our environment is key for social interaction (Bodner, Engelhardt, Minshew & Williams, 2015; Moskowitz, 2005), learning (Gopnik, Sobel, Schulz & Glymour, 2001), reasoning (Goswami & Brown, 1990) and self-awareness (Duval, Silvia & Lalwani, 2012; Duval & Wicklund, 1973).

However, causal relations (for example, that event 'A' is the cause of event 'B') can never be explicitly known given our sensory organs' inability to detect causal relations, nor does sensory input provide explicit information as to causal relations (Cheng, 1997; Nadel & Hardt, 2011; Shanks, Holyoak, & Medin, 1996; Scholl & Tremoulet, 2000). Instead, we infer causal relations from the extent to which they fulfil criteria first proposed by Hume in his Treatise of Human Nature (1739/1888).

One such criterion, that of temporal and spatial contiguity, identifies that events occurring closer together in time and space (respectively) are more likely to be inferred as causally related. However, research suggests this conjecture is only one half of a bi-directional constraint of Bayesian causal inference (Eagleman & Holcombe, 2002; Fereday & Buehner, 2015), where events causally related to an ensuing event are often perceived as more temporally/spatially contiguous (Buehner & Humphreys, 2009, 2010; Callan, Moreton, & Hughes, *submitted;* Cravo, Claessens, & Baldo, 2009; Engbert, Wohlschläger, & Haggard, 2008; Haggard & Clark, 2003; Haggard, Clark, & Kalogeras, 2002). In other words, when we know two events to be temporally/spatially contiguous, we infer causality. Conversely, when we know two events are causally related, we infer temporal/spatial contiguity.

In terms of temporal contiguity, either the individual events (e.g., action/outcome) are perceived closer to one another or the temporal interval between the events is perceived as shorter, relative to non-causally related events. As such, this temporal contraction, otherwise known as temporal binding, has been used as a measure of implicit causal inference (e.g., Blakey et al., 2019; Buehner, 2012; Buehner & Humphreys, 2009), and is employed throughout experiments 6 and 7 reported here.

Temporal binding, operationalised as the reduced perceived temporal interval between an action and an ensuing outcome (compared to the actual interval), has evoked a wealth of research (see Moore & Obhi, 2012, for a review). However, several interpretations of how the binding phenomenon arises exist: as stated, previous research suggests that binding, both spatial and temporal, results from knowledge of causal mechanisms between events, where the knowledge that one event will cause another renders our perceptual faculties to temporally and spatially bind the two events together (Buehner, 2012; Buehner & Humphreys, 2010). Conversely, the forward model promotes that the human motor system gives rise to temporal binding when intentional actions are met with expected motor identity predictions, where the latter refers to prediction of the learned identity of a sensory event based on an action performed by the participant (Engbert & Wohlschläger, 2007; Engbert, Wohlschläger, Thomas, & Haggard, 2007; Haggard et al., 2002; Hughes, Desantis, & Waszak, 2013) Moore & Fletcher, 2012; Moore, Wegner & Haggard, 2009; Obhi & Hall, 2011). However, multiple accounts have arisen that challenge the proposition that motor identity-prediction is a necessary component of temporal binding (Barlas, Hockley, & Obhi, 2017; Desantis, Hughes and Waszak, 2012; Hughes et al., 2013). For example, Desantis et al. (2012) found that temporal control (i.e., the ability to control the temporal onset of an outcome) produced greater temporal binding than the ability to predict the identity of the outcome. Buehner and Humphreys (2009) equally showed that, despite identical outcomes, , actions that caused a tone, compared to actions timed to the onset of a signal indicating that the tone will occur, produced the most binding. In other words, regardless of knowledge over what specific outcome will occur, understanding that there is a causal relationship between a given action and some event occurring afterwards may be the primary modulatory factor that temporally attracts outcomes towards their cause (and vice versa). Thus, given our understanding requires further clarification over temporal binding's underlying mechanism, we further investigated whether knowledge of a causal relationship between two events was sufficient to produce binding relative to non-causally related events across two experiments.

Experiment 6 extended Buehner's (2012) research that measured binding via stimulus anticipation, where participants are asked to time a button press with the onset of an outcome stimulus (e.g., a visual or auditory cue). In this method, the participant presses a button that either produces the outcome (causal condition) or a signal indicates that the outcome will occur after a given time interval (baseline condition). Hence, temporal binding is quantified as the anticipatory error of the timed action to the onset of the outcome. As such, greater anticipation reflects greater temporal contiguity: causally related events are perceived closer together in time, where this perceptual bias induces the notion that causally produced outcomes will appear sooner relative to unrelated or merely correlational secondary events (Buehner, 2012). Indeed, using stimulus anticipation, and consistent with the causal binding hypothesis, research has shown causally produced outcomes are anticipated more than non-causal outcomes (Blakey et al., 2019; Buehner, 2012; Buehner & Humphreys 2009).

However, Buehner (2012) introduced a further condition, where across two experiments he compared self-caused actions to machine-caused actions in order to test whether intentional action or causal agency (i.e., an agent performing a specific action that produces an outcome) were sufficient to produce binding. Thus, Buehner employed three conditions: self-causal, machine-causal, and baseline. In the self-causal condition, participants pressed a button on the left-hand side of a response box that would cause an LED bulb to flash after an interval of either 500, 900, or 1300ms (blocked) on the right-hand side of the response box. The participant's task was to press a second button on the right-hand side of the box when they thought the LED was going to flash. In the machine causal condition, participants initiated a machine by pressing a button on a machine, which, after a period of 2-5s, would activate a lever that pressed down on the left-hand button on the response box. Participants again pressed the right-hand button on the response box when they thought the LED would flash. As a point of emphasis, Buehner (2012) strongly advocated to participants that the machine was separate from any computer influence and acted on its own behalf within the given time frame. A baseline condition involved a signal LED flash on the left-hand side of the box after a period of 2,300 to 2,800ms after trial onset, indicating that the response LED would flash after the interval. The results showed that, compared to the baseline condition, both self-causal and machine-causal conditions showed greater stimulus anticipation of the LED flash, without differing from each other.

Thus, contrary to the forward model where intentional action being met with expected consequences produces temporal binding, Buehner (2012) provides evidence that an intentional action is not strictly necessary to produce binding. Rather, the author proposed an account of causal binding to explain how a machine's action could equally produce temporal binding, and that the temporal binding phenomenon occurs when a perceiver understands the underlying causal relation linking events. The fundamental principle of temporal binding, as proposed, is the predictive relation between two perceivably causally linked events. In other words, when one understands that event A causes event B, the time interval between the two events contracts.

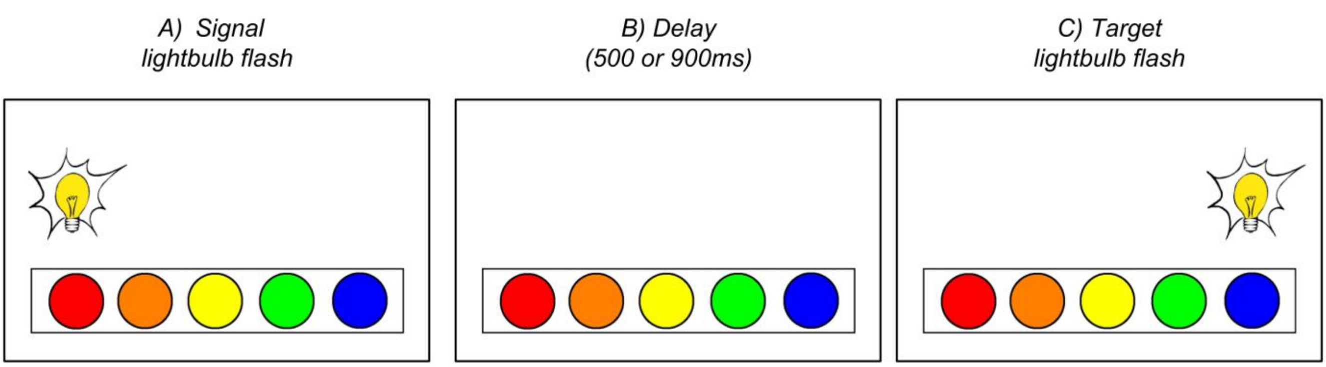
Across 2 experiments we tested the degree to which temporal binding, as measured using the stimulus anticipation paradigm devised by Buehner (2012), required a causal relationship between events. In experiment 6, we tested whether the explicit instruction to the participants regarding the computer being "an autonomous mechanical causal agent" (p.1491), was necessary to produce the binding effect in the machine causal condition. Buehner (2012) reasoned that previous studies using machine-caused conditions (e.g., Wohlschläger, Haggard, Gesierich, & Prinz, 2003) had not produced temporal binding because both the target tone, and the machine press were caused by the computer running the task. This means that the computer would be the common cause of both events rather than the machine causing the tone.

Buehner (2012) represents a departure from the traditional view that intentional action is a key mechanism for temporal binding of actions and outcomes given that, according to the forward model, it is a necessary component for generating sensorimotor predictions in order to compare actual vs. expected sensory feedback of our actions (Engbert et al., 2008; Haggard & Clark, 2003; Haggard et al., 2002; Moore, 2016). Nonetheless, other studies (e.g., Cravo, Claessens, & Baldo, 2009; Moore & Haggard, 2008) suggest that *both* action (vs. passive observation) and the impression of a causal relation between two events are required for binding, with neither sufficient to produce the effect alone.

Experiment 6 thus sought to further clarify the roles of intention and causality in temporal binding. In our extension of Buehner (2012), we replaced the machine-causal condition with a timer condition. Effectively, this condition maintains the components of the machine-causal condition, but removes the machine's presence as an agent, and thus any causal component to Buehner's condition. The timer condition was named so for the visual presence of a timer counting upwards, which was present upon the machine used in the original experiment. Buehner (2012) found greater stimulus anticipation for self-caused and machine-caused outcomes compared to those outcomes preceded by a temporally predictable signal at 500 and 900ms. Thus, if participants anticipate outcomes to a comparable degree between the self-causal and timer conditions (compared to baseline) this would suggest that a causal element between events is not solely necessary. Conversely, if outcomes within the timer condition are anticipated closer to baseline levels, this would suggest the reverse. As such, Experiment 6 was designed to test the degree to which actions produced by an autonomous agent modulated stimulus anticipation.

Experiment 6

In Experiment 6 we extended previous research (Buehner, 2012) in order to test whether knowledge of a causal mechanism to produce a stimulus change would elicit temporal binding. We modified our experiment in three ways compared to Buehner (2012): firstly, given the similarity in results of stimulus anticipation for temporal intervals 900ms and 1300ms in the original study, we only employed prediction time intervals of 500ms and 900ms. Secondly, we replaced the machine-causal condition with a time-sequence condition, which displayed a clock counting upwards in seconds on the left-hand side of the screen. Both experiments took place online, and so Experiment 6 utilised a virtual control-box instead (see Fig. 1) of the physical control box used in Buehner (2012), and instead of LED flashes to indicate Event A (signal of cause) and Event B (outcome), we used images of light bulbs to indicate events. Thus, a signal light bulb image appearing on the left-hand side (Event A) of the screen preceded a target light bulb image appearing on the right-hand side (Event B). As in the original study, the participant's task was to press the appropriate button when they thought the target light bulb would appear after a stimulus prediction interval (SPI) after the signal light bulb.



Chapter 4, Figure 1. Virtual control box with: A) Signal lightbulb flash, B) Delay screen, C) Target lightbulb flash

Method

Participants. We recruited 80 participants (43 males: M_{age} = 36.41, SD_{age} = 11.26) through prolific.ac, an online crowdsourcing platform. An additional participant was excluded due to a technical problem. Participants received monetary compensation. We screened participants for the following inclusion criteria: an approval rating of above 90% on prolific.ac (based on prior experiment performance/approval scores) and aged between 18 and 65. Additionally, we applied the following exclusion criteria: native language other than English, left handedness, recent use of illicit drugs, uncorrected visual or auditory impairment and history of psychiatric or neurological illness. The required sample size was fixed ahead of data collection, and a power analysis showed we had 90% power to detect a small effect (Cohen's f = 0.10) of causal relations and stimulus predictivity on temporal binding (a = 0.05).

Materials and Procedure. Participants were told they would partake in a stimulus-anticipation experiment, in which their task would be to anticipate an image of a target light bulb appearing on the right-hand side of their computer screen by pressing a button (P) then they thought the light bulb would appear. This target light bulb was preceded by a signal light bulb that flashed for 100ms on the left-hand side of the screen, separated from the target light bulb with either a short (500ms) or long (900ms) SPI. SPI was blocked, and counter-balanced across participants. Participants were tasked with initiating each trial when presented with a message instructing them to press the spacebar to begin the trial (throughout all conditions, as per Buehner, 2012, Experiment 7).

During the baseline condition, after trial-initiation with the spacebar, the signal light bulb appeared on the left-hand side of their computer screen after a randomly determined interval between 2,300 and 2,800ms. Then, after the SPI, the second light bulb flashed for 100ms on the right-hand side of the screen. Participants were asked to press the ‘P’ button on their keyboard when they expected the target light bulb to appear.

The self-causal condition procedure was identical to the baseline condition. However, instead of the first light bulb appearing after a random interval, participants activated the first light bulb by pressing the ‘Q’ button at a time of their choosing. After the SPI, the target light bulb appeared, where participants were asked to press the ‘P’ when they expected the target to appear.

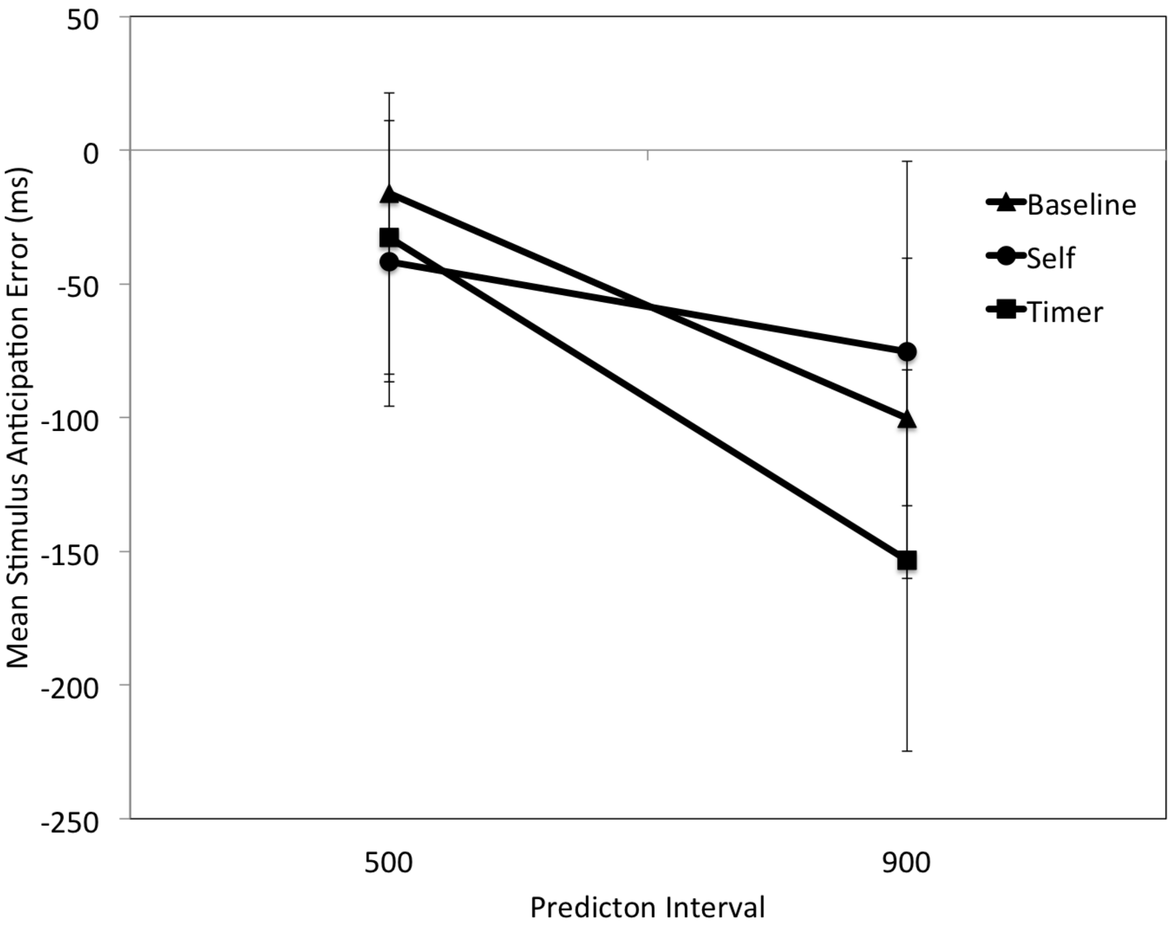
During the time-sequence condition, a timer began 1000ms after the trial began. Starting from 0, the timer counted upwards in seconds at the location of the signal light bulb. After a randomly determined interval between 2-5 seconds, the signal bulb flashed, followed by the target light bulb after the SPI. Specifically, participants were informed, “After 2 to 5 seconds, you will see a light bulb flash on the LEFT-hand side of the screen. This indicates that another light bulb will flash on the RIGHT-hand side of the screen. Your task is to press the 'P' button on your keyboard when you think the second light bulb will appear”. In other words, that the timer would not cause the target light bulb to flash, but only to indicate when the signal light bulb would appear. Once again participants were tasked with pressing ‘P’ when they expected the target to occur.

Experiment 6 consisted of 60 trials: 20 baseline trials, 20 self-causal trials, and 20 time-sequence trials. Each condition was blocked, and the order sequence randomised across participants. Equally, the prediction interval (500ms or

900ms) was blocked within each condition, and randomly presented across participants. Thus, participants experienced both time intervals before advancing to the next condition.

Results

Employing Buehner's (2012) data analysis protocol, median stimulus anticipation times of the last 10 trials of each block were calculated for the two prediction intervals (500 and 900ms) and for each of the three conditions. Four participants whose data included scores three times the inter-quartile range from the median were excluded from data analysis. We then conducted a 3 (condition: baseline, self-causal and time-sequence) X 2 (prediction interval: 500 and 900ms) fully within-subjects ANOVA (see Fig. 2).



Chapter 4, Figure 2. Mean stimulus anticipation error by condition (Experiment 6).

Analysis revealed a significant main effect of prediction interval, $F(1, 74) = 53.97$, $p < .001$. Analysis also revealed a main effect of condition on stimulus anticipation, $F(2, 148) = 4.31$, $p = .015$, $\eta p^2 = .055$. There was also a significant interaction between prediction interval and stimulus anticipation, $F(2, 148) = 16.15$, $p < .001$, $\eta p^2 = .179$. Paired sample t-tests revealed that the self-causal condition produced more temporal binding than the baseline condition at a prediction interval of 500ms, $t(74) = 2.65$, $p = .01$, $dz = 0.03$. However, there was no difference between the time-sequence condition and the other conditions (*p*s > .05). At a temporal interval of 900ms, the time-sequence condition produced more temporal binding than the both baseline condition, $t(74) = 3.26$, $p = .002$, $dz = 0.10$, and the self-causal condition, $t(74) = 4.24$, $p < .001$, $dz = 0.42$, with no differences observed between the latter two ($p > .05$).

Experiment 6 - Discussion

Experiment 6 found greater binding via stimulus anticipation at 500ms for self-caused actions than passively observing a sequence of two events (baseline). Our timer condition did not significantly differ from either self-causal or baseline conditions. Conversely, at 900ms, we observed greater stimulus anticipation of the second stimulus during the timer condition than both other conditions. This complex pattern of results does not provide clear support for the assumption that perceived causality drives temporal binding in this task. Unlike in Buehner et al., (2012) we did not instruct participants that the outcome was caused by an independent agent (a machine in their case), we simply presented participants with a sequence that was not initiated by them and asked them to anticipate the presentation of the second stimulus. At least at 900ms, this was sufficient to produce greater intentional binding than in the baseline condition. It is also noteworthy that although self-caused and baseline conditioned significantly differed at 500ms, they did not differ at 900ms. Although a similar pattern was observed in Buehner et al. (2012), they did not show a significant interaction between interval and condition, whereas we did (although see Humphreys and Buehner, 2009, also discussed below).

Experiment 7 aimed to further extend these findings by directly manipulating perceived causality by instructing participants that either they or someone else triggered the outcome. Such a strategy has previously been successfully shown to modulate temporal binding using the Libet Clock method (Desantis et al., 2011), so here we aimed to further assess the stimulus anticipation paradigm as a measure of causal binding by conceptually replicating another experiment, Desantis et al. (2011), using a stimulus anticipation measure of temporal binding.

In their paper, Desantis et al. (2011) investigated whether prior causal beliefs influence temporal binding via the Libet Clock method (Libet, Gleason, Wright, & Pearl, 1983). In typical temporal binding experiments employing the Libet clock method (Engbert & Wohlschläger, 2007; Haggard et al., 2002; Moore & Haggard, 2008, 2010; Wohlschläger, Haggard, Gesierich, & Prinz, 2003) self-generated actions are estimated to have occurred later, relative to baseline measurements, whilst the reverse happens for outcomes, and thus actions and outcomes are perceptually drawn together in time for self-generated actions.

In Desantis et al. (2011), during a set of belief implementation phases, participants were instructed that they would participate alongside the confederate as a second participant. Furthermore, that their task was to press the space bar and estimate the onset times of either their action or the ensuing tone (occurring 150, 300 or 450 ms after action) via the Libet clock when their name appeared at the top of the screen. During these phases, participants also observed a confederate estimating onset times of their (the confederate's) actions (with a different button)/tones when the confederate's name appeared. During experimental trials, participants (and the confederate) pressed their button to produce the sound regardless of the name appearing at the top of screen, although were informed that only the action given by named person would produce the sound. However, during experimental trials, the participant's actions caused all outcomes to occur, regardless of the name at the top of the screen. Participants again reported onset estimations of actions and outcomes in separate conditions.

The findings of Desantis et al. (2011) showed that, despite being responsible for all outcomes, participants showed less temporal binding when they believed the sound to have been caused by the confederate's button press. From the perspective of the forward model, these findings suggest that causal beliefs (i.e., thinking you did not cause an outcome despite the opposite being true) can inhibit what would have otherwise been a predicted action-effect. In turn, the process in which actual/predicted sensorimotor effects is absent, hence actions and action-effects are not temporally drawn to one another. In other words, Desantis et al., (2011) provide evidence that the forward model is an insufficient account of temporal binding. Such an interpretation equally appears to be in line with research suggesting both intentional action and causal relation are both required to induce temporal binding (Cravo et al., 2009; Moore & Haggard, 2008).

From the causal binding perspective, the findings from Desantis et al. (2011) support the notion that causality, not intentional action, is at the root of binding. Desantis et al. (2011) proposed that causal knowledge indicates what outcomes to predict, from the position of the forward model, and thus unpredicted outcomes cannot be temporally bound to an action believed not to cause an outcome. However, this conjecture contends with Buehner's (2012) causal account of binding that eliminates intentional action from the temporal binding equation, where even mechanical agents enacting causally effective actions produce the observed reduction in the temporal interval between action and outcome. Indeed, Desantis et al. (2011) premise their study on prior research claiming that an intentional agent is required to produce binding (Wohlschläger, Engbert, & Haggard, 2003; Wohlschläger, Haggard et al., 2003), which Buehner (2012) directly criticised for lacking a genuine causal connection between machine actions and the ensuing outcome.

More directly, Desantis et al. (2011)'s findings equally contend with the strength of causal knowledge as a modulator of temporal binding in Experiment 6 presented here. Participants passively observing the sequence of stimuli in the timer condition anticipated the secondary flashbulb significantly more than outcomes produced in the self-causal condition at 900ms.

Desantis et al. (2011) employed the Libet clock method (Libet et al., 1983) to capture temporal binding. However, this method has been subject to criticisms that suggest the clock is an inadequate method of testing the interval between events (Humphreys & Buehner, 2009; Pockett & Miller, 2007), highlighting the rise in alternative measures such as interval estimation and stimulus anticipation. Thus, Experiment 7 of the present paper aimed to conceptually replicate Desantis et al.'s (2011) research on prior causal beliefs by employing stimulus anticipation to measure temporal binding. By doing so, we again renew any exploration of the mechanisms of temporal binding by replicating previous research.

As such, there are three potential hypotheses to Experiment 7: 1: Both self and other causality will lead to increased anticipation but will not differ (i.e. binding is generated by causality, independent of whether you caused the outcome or someone else). This hypothesis corresponds with the explanations of temporal binding within Buehner (2012), where there is a clear causal connection between action and outcome. 2: Binding is present in both conditions, but greater in the self-caused condition (i.e., binding is causality related, but is also boosted in reference to the self). 3) There is no binding in either condition, suggesting that stimulus anticipation does not accurately track causal belief or intentionality.

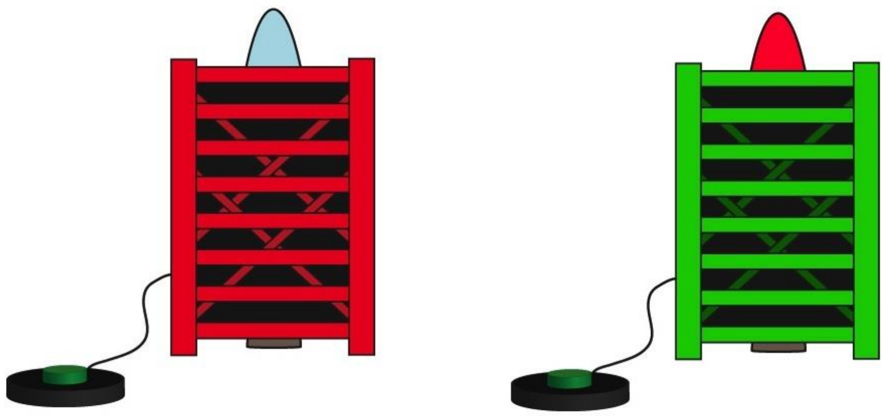
Experiment 7

Methods

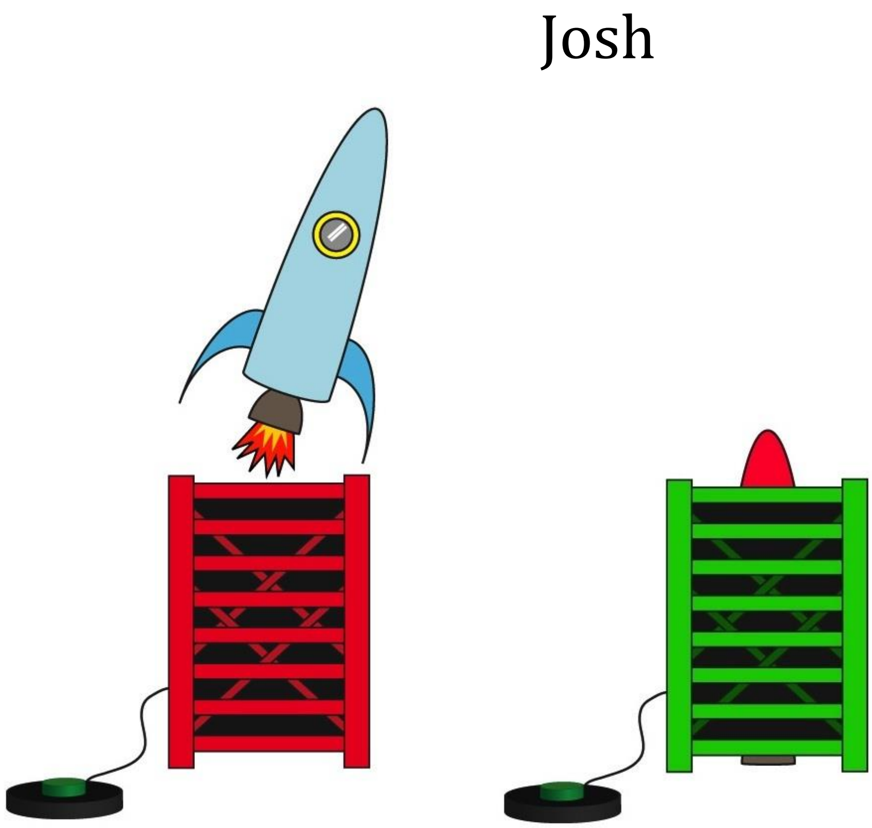
Experiment 7 conceptually replicated Desantis et al.'s (2011) work on the roles of intentional action and causal beliefs upon temporal binding. Instead of the Libet clock method, we used stimulus anticipation to measure temporal binding. As Experiment 7 was conducted online, participants were led to believe they were paired with another user of the crowdsourcing platform Prolific.ac (as opposed to the original study, which used a real-life confederate). Stimulus anticipation was achieved through similar methods to Desantis et al.: training phases reinforced the notion that self-action caused a visual stimulus to change (see procedure below), in the form of a rocket launching (see Fig. 3; stimuli acquired from Blakey et al., 2019), on the screen, and that the other 'user' was acting autonomously, performing their own actions and rendering changes to their own stimuli. However, during experimental phases, participants were solely responsible for all stimulus changes, and we expected stimulus anticipation scores, and thus temporal binding, to reflect those observed within the original study: perceived self-caused action-effects would induce greater temporal binding than perceived other-caused action-effects, despite all action-effects being self-caused.

A)

Josh



B)



Chapter 4, Figure 3. A) Depiction of both participant's and other 'user's stimuli at rest, name above indicating who must launch their rocket. B) Depiction of participant's rocket launching after SPI.

Participants. We recruited 95 participants (33 males: M_{age} = 39.19, SD_{age} = 10.12) through prolific.ac, an online crowdsourcing platform. Due to unclear reasons, many participants did not respond correctly to the task and 'timed out' during the time-sensitive conditions (see below for an explanation). We excluded participants who timed out more than 10 times, leaving 64 participants left with useable data for analysis (mean/SDs of timeouts for included participants: M = 4.60, SD = 2.99; for excluded participants: M = 20.41, SD = 9.76). The high exclusion rate is further discussed below. We screened participants for the same inclusion/exclusion criteria as Experiment 6.

Materials and Procedure. Participants were instructed they would partake in an experiment alongside another Prolific.ac user. In this experiment, participants had a specific rocket that belonged to them on the left-hand side of the screen, whilst the other 'user' owned a rocket on the right-hand side (Fig. 3a). The participants were tasked with initiating their rocket's launch sequence with the 'Q' button, and then anticipate the onset of the rocket launching by pressing the 'R' button when they thought the rocket was going to launch (Fig. 3b).

Participants were briefed on what to expect throughout the experiment, which involved two training phases and the main task, and asked to type a name they wished to be referred to as throughout the experiment. They then proceeded to a sham 'connecting' screen, which showed a box displaying that the experiment was attempting to connect to another user. After 30 seconds, participants were told they had successfully connected to another user named 'George', who would be their partner for the remainder of the experiment.

A pre-trial screen containing the name indicating whose turn it would be to launch their rocket in the coming trial, displayed for 2500ms at the top of the screen, preceded each trial throughout the experiment. This pre-trial screen was then followed by a fixation cross, displayed for a random interval between 1500 to 2700ms, before the participant's rocket (left) and George's rocket (right) appeared, whilst the name of whose turn it was to launch their rocket remained throughout the trial.

Two training phases preceded the main task. During the first training phase, participants were instructed that their name appearing during the pre-trial screen, as well as at the top of screen during the trial, indicated that it was their turn to press their launch button ('Q') to initiate the launch sequence of their rocket on the left-hand side of the screen, and anticipate when they thought their rocket was going to launch by pressing 'R'. During 'George's' trials, participants were told they did not have to perform any action. Participants completed two comprehension-check questions regarding what the correct button was for launching the rocket, and what was the correct button to press when they anticipated it to launch, before the training phase commenced.

Following the training phase instructions, participants advanced to the task, which involved the pre-trial screen indicating whose rocket would be launched in the following trial. Throughout participant trials, the two rockets appeared on the screen after the fixation cross and following the participant's action (pressing 'Q'), the participant's rocket launched. Participant's action and the rocket launching were separated by a stimulus prediction interval (SPI) of either 500ms or 900ms (blocked). Subsequently, stimulus anticipation was measured via the error between participant's button pressed indicating when they thought the rocket was going to launch ('R') and the trial's SPI.

In trials where George's name was present, the two rockets appeared on-screen after the fixation cross, and the rocket on the right-hand side of the screen launched after a random interval (600-1400ms for the 500ms SPI, and 1000-1900ms for the 900ms SPI). George's launched rocket displayed onscreen for 1000ms before the trial ended (see Fig. 2).

The first training phase lasted two blocks of 20 trials each, one for each SPI, and randomly presented between participants. Ten trials per block were dedicated to each 'participant' (naïve and false/non-existent). After the first block concluded, participants were instructed that they would be doing the same task, but with a different interval separating the launch and anticipation button presses.

The second training phase was similar to the first phase; however, participants were now encouraged to respond as quickly as possible to the rockets appearing onscreen after the fixation cross. During participant trials, failing to respond within one second of the rockets appearing resulted in an error message being displayed, and the trial repeated. Additionally, participants were also encouraged to press their launch button on the other user's trials as well, and to predict when their rocket would launch with the 'R' button, as this would aid them in the main experiment, even though this would have no impact on the trial.

Similar to the first training phase, the second phase lasted for two blocks of 20 trials each, one for each SPI, randomly presented, and 10 per 'participant'. To reinforce the presence of another user, 2 of George's trials ended in an error message stating that the other user had failed to respond within the time limit, and that their trial would repeat.

After the two training phases, participants were then told they would now move onto the main task, which would be very similar to the second training phase. As with the training phases, the participant's overall task was to initiate the launch sequence with 'Q' when the rockets appeared after the fixation cross, and press 'R' when they thought the rocket was going to launch. However, participants were instructed to respond to the rockets appearing with their launch button regardless of whose name appeared during the pre-trial screen/at the top of the screen during the trial, and that during the other user's trials this would not impact the timing of their rocket's launch. However, contrary to these instructions, and to the training phases, participants' action controlled the launch sequences of both rockets, regardless of whose name appeared during the pre-trial screen. In other words, the participant's action ('Q') always initiated the launch sequence of the rocket of the respective owner's trial (i.e., launched their own rocket when their name was displayed, and launched George's rocket when their name was displayed). Thus, similar to Desantis et al. (2011), unbeknownst to the participants, participant's action always produced the outcome (i.e., either rocket launching, depending on the designated author) and controlled when the outcome would occur, regardless of whose name designated authorship of the outcome. To encourage participants to respond to the rockets appearing in a timely fashion during 'George's trials, participants were told they could earn extra income by responding the quickest to the image of the two rockets appearing, at 2p per trial(for a total of £1.60) throughout the main task.

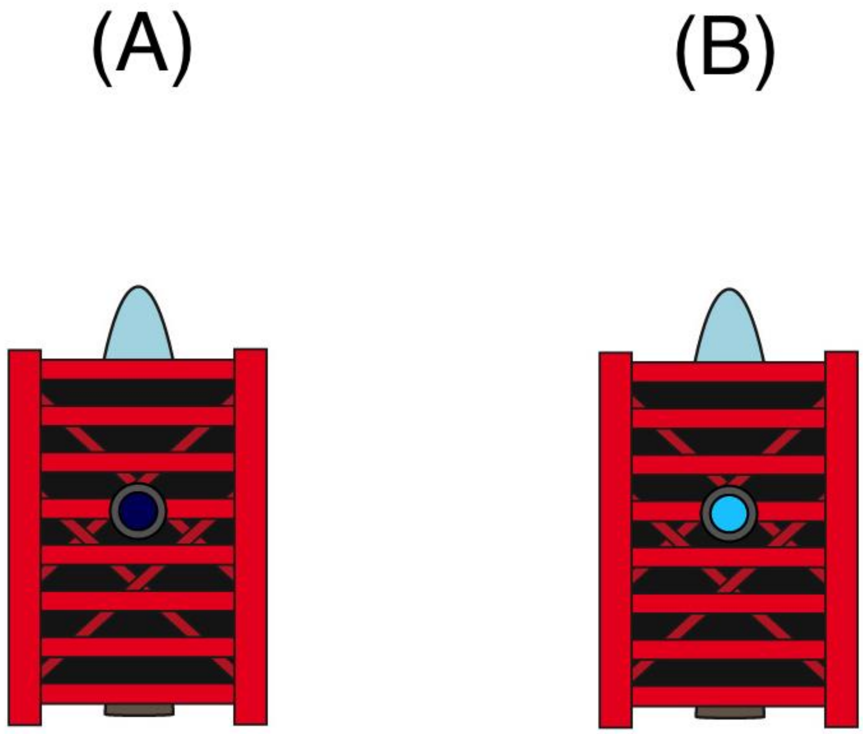
Additionally, participants were instructed to predict when either rocket would launch (i.e., both their own and the other user's rocket). Thus, the main task involved participants unknowingly initiating the launch sequences of both their own and George's rockets, and predicting when they thought said rocket would launch. As with Desantis et al. (2011), we expected that, when believing one to be the cause of the outcome, action/outcome would be temporally drawn together. Conversely, when under the belief one is not the cause, such temporal binding would be significantly diminished. After comprehension-check questions that asked what buttons they needed to press during the trial, participants proceeded on to the main task.

As per Desantis et al., (2011), if the participant failed to respond during 'other' trials with 'Q', the trial would display George's rocket launching, but repeat until the participant responded within one second of the stimuli appearing.

As mentioned during the previous participants subsection, some participants failed to follow the main task's instruction. Specifically, 31 participants managed to respond correctly with the launch button ('Q') to both their own and George's rockets, but failed to anticipate when the participant thought the rocket was going to launch with 'R'. Subsequently, these participants frequently timed out, and were removed from data analysis.

The main task consisted of two blocks of 40 trials each, one block per SPI, and 20 trials per block per 'participant'.

The baseline condition appeared either before or after the main task was completed. This involved a singular rocket appearing onscreen after 500ms. For those who performed the baseline condition before the main task, after a further interval between of 500ms a light would flash in the centre of the launcher-box (see Fig. 4), which indicated that after SPI of 500 or 900ms the rocket would launch. The participant's task was to time the launch of the rocket with the launch button ("R"). For those who performed the baseline condition after the main task, the interval between the rocket appearing onscreen and the flash in the centre rocket was set from the reaction times to the stimuli appearing onscreen during the main task. This allowed us test for order effects of condition, whilst attempting to avoid any effect of onset difference during stimulus appearing / target indicator / target appearing between the conditions. No difference was observed between the groups who experienced the baseline condition before the main task or after (p > .05).



Chapter 4, Figure 4. A) Depiction of participant's stimulus rocket at rest during baseline condition. B) The light flashing on the launcher-box indicating the rocket will launch.

Participants were told that after a certain amount of time, a singular rocket will appear, and that unlike the previous tasks (those performing the baseline condition after the main task), they were not reacting to this image. Instead, participants were instructed to wait for a light on the rocket to flash, which would indicate that, after a certain amount of time, the rocket will launch, and that their task would be to press the launch button ('R') when thought the rocket would launch.

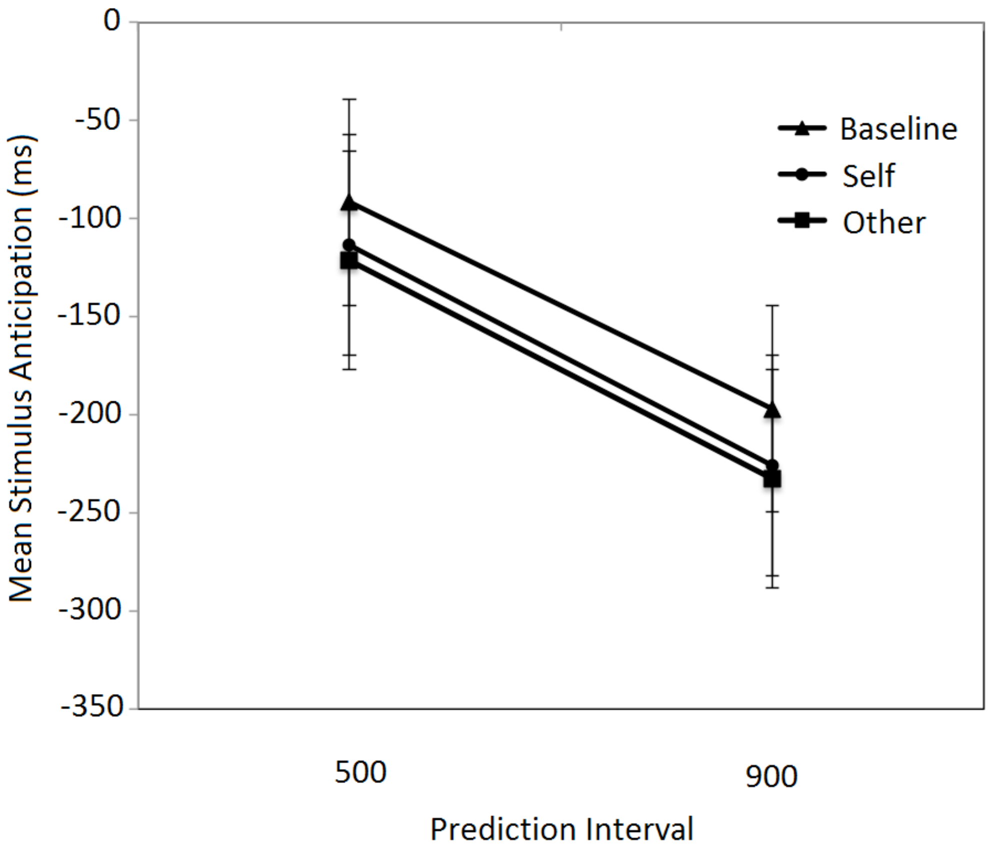
Additionally, questions regarding the causal belief of the outcome occurred after 10 trials randomly throughout the main task. Participants were asked 'Who triggered the outcome? Please make this judgement based upon your own feeling of you or your partner having caused the outcome.'

The SPI between launch button ('Q') and the rocket launching lasted either 500ms or 900ms. SPIs were blocked, such that participants completed training and experimental trials before moving on to the other SPI and were counterbalanced between participants. After participants completed the main experimental trials for one SPI, they then proceeded to the second training phase again, but with the remaining SPI between launch button press and the rocket launching. Afterwards, participants repeated the main experimental trials with the remaining SPI.

After completion of all experimental trials, participants were debriefed as to the use of deception used with the experiment and given the option to withdraw their data. As participants had been deceived as to the nature of the reaction time element of the main task, and subsequent possible monetary reward, each participant was paid an extra £1 on top of their participant payment.

Results

We used the same analysis procedure as Experiment 6, with a 3 (condition: baseline, self-causal and other-causal) X 2 (prediction interval: 500 and 900ms) fully within-subjects ANOVA (see Fig. 5).



Chapter 4, Figure 5. Mean stimulus anticipation error by condition (Experiment 7).

Similar to Experiment 1, analysis revealed a significant main effect of prediction interval, $F(1, 63) = 245.44$, $p < .001$, $\eta p^2 = .796$. However, there was no main effect of condition on stimulus anticipation, $F(1, 63) = 1.22$, $p > .05$, $\eta p^2 = .019$. There was also no significant interaction between prediction interval and stimulus anticipation, $F(2, 126) = 0.27$, $p > .05$, $\eta p^2 = .179$.

Results from the causal questions presented throughout the main task, equally distributed between the 'self 'and 'other' trials, showed participants responded with 'self-caused' (identifying themselves as the cause of the rocket launch) on 87.5% of trials for both the 'self' and 'other' trials. This significant lack in the belief that George's outcomes were indeed caused by him is in conjunction with post-experiment questions: when asked "Do you feel that, during the main experiment, your launch button 'Q' caused the other user's rocket to launch?", 56% of participants felt that their launch button did not cause George's rocket to launch.).

To reconcile these percentage differences, the time elapsed between the main task and post-experiment questions may have afforded participants reflective time over the authorship of outcomes. Moreover, the actual presence of the question itself may have biased participants' responses. The causal questions during the main task also did not allow any alternatives to 'self' or 'other' causes, and thus regarding the more open post-experiment question of whether participants thought their button press caused George's rocket to launch, "no" permitted non-George alternatives (for example, the computer program/software itself). To the question "What do you think is the purpose of participating alongside another Prolific Academic user?", 16% conveyed a lack of belief that there was another user, supporting the idea that participants likely did not know precisely what the mechanism behind George's rocket launching, but were nevertheless suspicious of his existence. No participant reported suspicion that they were the cause of the confederate's rocket launching.

Discussion

The aim of the current paper was to investigate the role of causal belief in temporal binding measured using stimulus anticipation. Experiment 6 advanced upon Buehner (2012) who found evidence for the notion that intention is not a necessary component to which temporal binding arises. We extended Buehner (2012), however with two distinctions: we modified the machine-causal condition to a timer condition using the same temporal parameters as Buehner, to test whether knowledge of a causal relation modulates temporal binding. Our second distinction was our use of an online platform, rather than a laboratory setting. Comparisons between online and laboratory settings have shown that participant performance is consistent across settings in terms of reliability of data (Dandurand, Shultz, & Onishi, 2008; Gould, Cox, Brumby, & Wiseman, 2015), therefore any differences between Experiment 6 and Buehner (2012) can reasonably be attributed to our modification of the machine-causal condition. Our findings showed that, at 500ms, we replicate Buehner (2012)'s finding of self-caused targets being anticipated earlier than those in the baseline condition. Our time-sequence condition that replaced Buehner's machine-causal condition, however, did not differ from either baseline or self-causal conditions, at least in the 500ms interval condition. At 900ms, outcomes in the time-sequence condition were anticipated more than the other two, with no distinction between self-causal and baseline conditions. This peculiar result is not easily explained; however, it is consistent with the trend shown in Buehner (2012) where the machine causal condition showed consistently larger mean stimulus anticipation at longer intervals (900/1300ms).

One explanation of this result is that temporal binding for self-causal actions may weaken beyond a certain threshold, where voluntary actions produce temporal shifts at short intervals (Engbert et al., 2008; Haggard et al., 2002). There exists no solid agreement over the extent of the temporal binding effect regarding the maximum length of the interval to which actions and outcomes become temporally drawn together. Haggard et al. (2002) reported the temporal binding effect lessening from 250ms to 650ms, whereas others have suggested that temporal binding may last up to 2 seconds (Shanks, Pearson, & Dickinson, 1989), if not 4 seconds (Buehner and Humphreys, 2009; Humphreys and Buehner, 2009). Moreover, Buehner (2012) suggested that intentional action might drive a 'boost' in action-outcome pairings at short intervals, despite not demonstrating this in a later experiment.

Whilst Experiment 6 partially replicated the findings of its original study, Experiment 7 failed to replicate Desantis et al. (2011)'s findings that causal beliefs modulate temporal binding. Specifically, in our study, there were no differences between self- and other-outcome stimulus anticipations. Examination of participant responses to causal authorship suggests this result is likely due to our failed manipulation of an online confederate. However, irrespective of our manipulation, one would still expect a difference to emerge between causal and non-causal events. Thus, given the lack of difference between the causal conditions and baseline condition, it is difficult to assert that Experiment 7 provides support for any particular explanation of temporal binding. Our findings, in reference to our hypotheses, tentatively support the possible outcome that there would be no binding in either condition, suggesting that stimulus anticipation does not accurately track causal belief or intentionality. However, we should be cautious to over-interpret our findings, given that we have departed from both original experiments by altering procedural elements (e.g., conducting the experiments online). Specifically, we might expect the non-significant trend in Experiment 7 of the causal conditions (compared to baseline) to become more pronounced within controlled laboratory conditions.

Buehner (2012) proposed that a causal connection between events would lead to greater binding. In other words, both the self and other conditions should have led to greater anticipation, as causality is generated independent of which agent was the cause. However, the latter conditions equally did not differ from the baseline condition, which contained no causal relation between events. This is also in contention with our findings from Experiment 6, which showed a significant difference between self-caused actions and passively observing events in the baseline condition, albeit only at 500ms. The inconsistency between experiments 6 and 7, and between Buehner (2012)'s proposal and our findings, questions whether stimulus anticipation is a reliable measure of temporal binding. Moreover, given that only two known prior publications have utilised stimulus anticipation (Buehner & Humphreys, 2009; Buehner, 2012), different methods of measuring binding, for example, interval estimation (Engbert & Wohlschläger, 2007; Engbert et al., 2008, Engbert et al., 2007; Fereday & Buehner, 2017; Humphreys & Buehner, 2009) may find success.

However, one failing may simply be the lack of believability of our study's paradigm to online participants, whereas Desantis et al. employed a real-life confederate. Any further exploration of this topic should ideally begin with a laboratory setting. Despite the previously stated research suggesting that data accrued from online studies is sufficiently reliable, our experiment may have not been designed adequately enough to convince participants of the genuine presence of another online user. This was evidenced by responses to catch trials asking participants ‘who was the cause of previous outcome’, and to post-experiment questions, suggesting the majority of participants did not believe that they participated alongside another user. Furthermore, a significant number of participants were excluded due to not responding within the correct time frame, and thus this suggests our instructions were not clear enough to ensure sufficient task comprehension. Again, a laboratory setting would enable the chance to clarify any participant confusion, and indeed others have suggested complex online studies should employ caution when it comes to task comprehension.

Assessing the degree to which binding is modulated by intentionality of action, causal knowledge of events, and low-level processes is important for clarifying the mechanism of a method used frequently in causal research Buehner, 2012; Buehner & Humphreys, 2010; Engbert & Wohlschläger, 2007; Engbert et al., 2007; Haggard et al., 2002; Moore & Fletcher, 2012; Moore et al., 2009; Obhi & Hall, 2011). Experiment 6 here demonstrates that causal knowledge may hold less influence than previously thought (Buehner, 2012). Although with potentially undermining weaknesses, Experiment 7 further also failed to show temporal binding for self-caused action outcomes using stimulus anticipation. Overall, the current studies provide new evidence that questions the degree to which stimulus anticipation is a reliable measurement of temporal binding, and the degree to which temporal binding, as measured via stimulus anticipation, is modulated by knowledge of causal relations, be it knowledge of the causal mechanism or the author of outcome.

Chapter 5: Bad People Feel Close to Bad Outcomes: Immanent Justice Reasoning by Spatial Proximity

Preface

The experiments contained within Chapter 4, extending previous research (Buehner, 2012; Desantis et al., 2011), failed to provide a clear picture of temporal binding modulation. Outcomes produced by a causal agent were more anticipated relative to baseline measures where outcomes occurred after a signal stimulus at 500ms, but not at 900ms, during Experiment 6. However, the self-causal/baseline distinction failed to occur during Experiment 7 for both time intervals. Our findings from Experiment 6, as well as those from chapters 2 and 3, challenged the notion that knowledge of a causal relationship between events influences the perceived temporal interval between the events. Yet, the experiments contained within chapters 3 and 4 were not without limitation, and thus we were cautious to accept such a finding without scepticism. To approach the aim of this PhD project from an alternative angle, and in conjunction with Buehner and Humphreys (2010) demonstrating that space appears to contract when causal beliefs are held about two events, we moved away from temporal binding by exploring whether the power of causal beliefs to modulate binding is reflected in biased judgments about the spatial position of abstract representations of moral acts. Thus, the research contained within chapter 5 of this thesis aimed to elucidate the role between IJR and an implicit measure of causality via spatial binding, where we expected morally congruent events to be positioned closer to each other relative morally incongruent events.

Abstract

Spatial contiguity is a key indicator of perceived causality, such that people perceive events that are causally linked as being closer together in space. Across 5 experiments, we tested the idea that immanent justice reasoning, which is this belief that moral actions bring about deserved outcomes, also influences spatial proximity. Participants positioned representations of people's fortuitous bad (vs. good) outcomes significantly closer in space to representations of their previous immoral actions when the outcomes occurred to the same person (Frank punched someone - Frank was in a car accident) more strongly than when the outcomes occurred to a different person (Frank punched someone - Joe was in a car accident). This occurred both when participants ordered the sentences in a list (Experiment 1) and when they were free to move the outcome anywhere on the screen (Experiment 2). In Experiment 3, the positive or negative outcomes were being "chased" across the screen by bad persons. Participants acted to keep good outcomes and bad people apart, as compared to bad outcomes and bad people. Taken together, these findings suggest that our desire to perceive that people get what they deserve biases the physical proximity of bad outcomes to previous immoral behavior.

Keywords:
Immanent justice; spatial proximity; spatial binding; social perception; deservingness

Bad People Feel Close to Bad Outcomes: Immanent Justice Reasoning by Spatial Proximity

A wealth of research has confirmed Hume's (1739/1978) assertion that "cause and effect must be contiguous in space and time" (p. 173). Perception of causality is influenced by both spatial proximity and temporal contiguity (Einhorn & Hogarth, 1986; Michotte, 1963; Rips, 2011). The converse is also true: events that are causally linked are perceived to be closer together in both time (Faro, Leclerc, & Hastie, 2005) and space (Buehner & Humphreys, 2010). For instance, actions and their ensuing sensory outcomes are bound towards one another in time (Haggard, Clark, & Kalogeras, 2002; Hughes, Desantis, & Waszak, 2013). Similarly, the distance between two moving balls is perceived as smaller when the movement of the two balls are causally linked than when they are not (Buehner & Humphreys, 2010).

Thus, Hume's proposed link between causality, space, and time is now well established. However, most previous research has focused on the degree to which events that are in fact causally related are subject to temporal or spatial binding. In the current research, we go beyond such mechanical causal relationships to investigate the spatial binding of social actions and outcomes in situations where plausible causal linkages between events are arguably missing but where perceivers are nonetheless motivated to perceive causality. Specifically, we investigated whether *immanent justice reasoning* influences the spatial proximity of actions and ensuing outcomes.

According to Lerner's (1980) just world theory, IJR serves as a function to maintain just world beliefs. Not only are we biased in our causal judgements, we are motivated to prefer explanations that help defend against threats to our belief in a just world. Recent research shows that construing events to be consistent with a just world allows us to make long-term goals (Bal & van den Bos, 2012; Callan, Harvey, Dawtry, & Sutton, 2013; Xie, Liu, & Gan, 2011), avoid self-defeating behaviours (Callan, Kay, & Dawtry, 2014), avoid smaller, immediate rewards in order to obtain larger, delayed rewards (Callan, Shead, & Olson, 2009), and maintain a commitment to justice in the face of threat (Hafer & Bègue, 2005). However, IJR research often utilises self-report methods to measure the strength of the causal link between someone's prior (mis)deeds and a future event. For example, reading a vignette detailing a fatal accident had occurred to either a good or bad person, followed by the question "to what extent do you feel that what happened to *xxx* was a result of his/her conduct?" (Callan et al., 2013). Callan, Ferguson and Bindemann (2012) showed that motivations to perceive justice exist beyond retrospective causal judgements using eye-tracking. Participants listened to a recorded vignette describing either a good or bad person through headphones, and their eye-gaze was recorded whilst two possible outcome images (good or bad) were displayed partway through the vignette. Callan et al. found that participant gaze shifted towards the morally congruent image, such that when listening to the bad (good) behaviour of a person, eye-gaze focused onto the bad (good) outcome. These findings suggest not only are we motivated to retrospectively assert morally congruent reasons for specific outcomes, we also anticipate morally congruent outcomes when justice is concerned.

The current research aims to extend previous IJR findings by exploring whether valence-laden sentences within a scenario are felt closer in space. Recent research demonstrates that not only does contiguity promote inferences of causality, but also the perception of a causal relationship leads one to perceive greater contiguity (Buehner et al., 2009). Thus, in terms of IJR, morally congruent sentences should be more contiguous to one another and felt closer in space. Conversely, morally incongruent sentences, and hence less causally related, should be felt further apart.

The goal of the current research was to extend previous IJR across three experiments using spatial tasks where participants varied the distance between sentences depending on how close they feel in space: a ranked choice task (Experiment 8), a free positioning task (Experiment 9) and a sentence-chasing task (Experiment 10). On the basis of previous findings (Buehner & Humphreys, 2009; Callan et al., 2006; Casasanto, 2008) we expected that morally congruent sentences would elicit smaller distances compared to incongruent sentences.

Experiment 8

Sampling

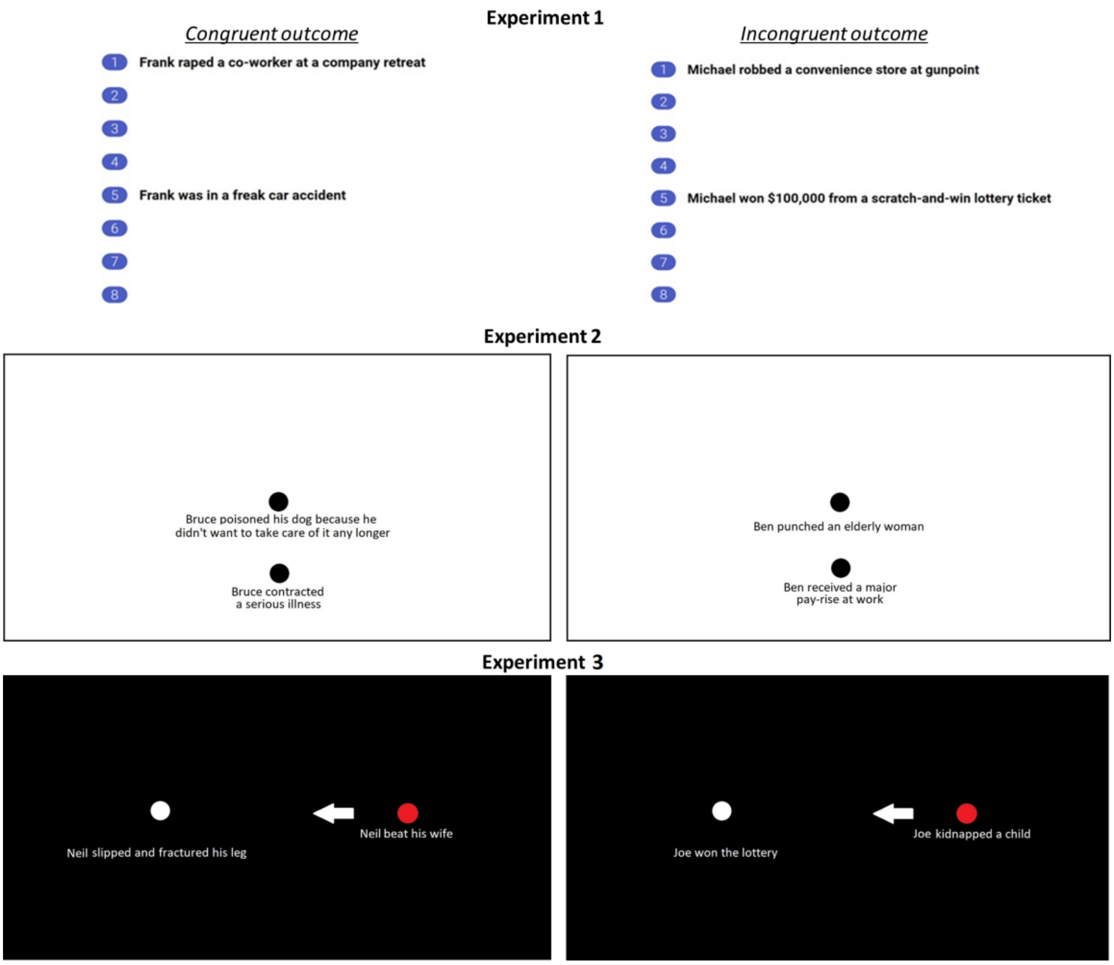
We recruited participants from the U.S.A. through Amazon's Mechanical Turk for all experiments. Unless otherwise noted, the required sample sizes across experiments were fixed ahead of data collection but the final sample sizes were not completely predetermined due to the unpredictable nature of online recruitment (e.g., because of slight over-recruitment and removing participants due to duplicate IP addresses). Power calculations for mixed-effects regressions can be difficult so we based our sample sizes on achieving at least 80% power to detect small-to-medium effects ($dz = 0.35$) in simpler, within-subjects t-tests for our within-subjects designs. The data and materials for all experiments are available at osf.io/37g9k/. We report all measures, manipulations, and exclusions in these experiments.

Method

Participants. In Experiment 8, 80 participants (44 males: M_{age} = 34.60, SD_{age} = 10.26) completed a brief online study about people's perceptions of physical closeness. We excluded an additional 3 participants due to duplicate IP addresses (here and throughout, we retained the earliest response).

Materials and procedures. We created a stimulus-manipulation procedure where participants moved two individual sentences, describing two events of a scenario, closer together or further apart depending on how close they felt the events were in space. There were 8 positions that the sentences could occupy (see Figure 1). At the beginning of each trial, the first event of the scenario (e.g., "Frank raped a co-worker at a company retreat") occupied the first position and the outcome event (e.g., "Frank was in a freak car accident") occupied the fifth position. Participants were instructed to click and drag the sentences closer together or further apart from each other than this original position. Specifically, for each trial, participants read: "How physically close do these two events feel to you? If the events feel closer to you than they are right now, drag them closer together. If they feel further away, drag them further apart."

Prior to the main block, participants completed 1 practice trial with neutral sentences ("David took his bicycle to the corner store" and "Ross bought a new jigsaw puzzle"), and they were shown images of example positions for sentences that could be as close together as possible (positions 1 and 2) or as far apart as possible (positions 1 and 8).



Chapter 5, Figure 1. Examples of congruent (left) and incongruent (right) trials/scenarios across experiments. The images depict the starting positions of the immoral actions and outcomes for each trial. For Experiment 10, the white arrow illustrates the starting path of the red circle's motion once the trial began.

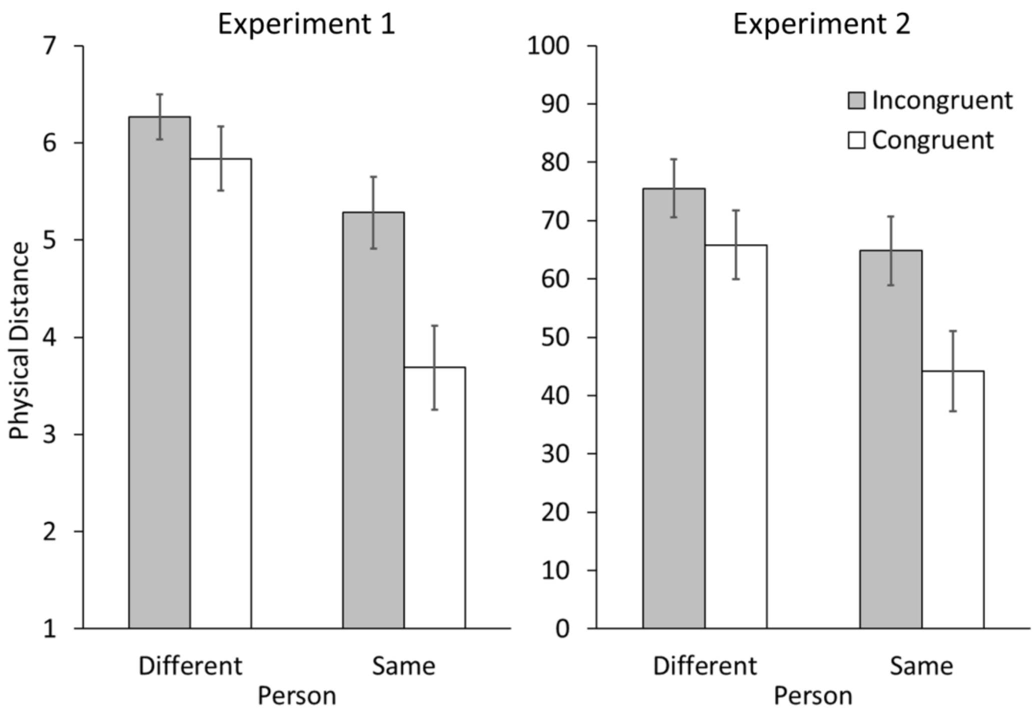
The scenarios we used for Experiment 8 represented the conditions of a fully-within 2 (Congruency: morally congruent vs. morally incongruent) X 2 (Person: same vs. different) factorial design (for a full list of scenarios used in each experiment, see Appendix B). In each scenario, the individual was a bad person. Congruency was manipulated by varying the value of the outcome (bad vs. good; congruent vs. incongruent, respectively). For the Person factor, we manipulated whether the outcome occurred to the same person or a different person by varying the names of the characters within the scenarios (e.g., Frank-Frank vs. Geoff-

Mark). We included the Person factor in our design to address one possible alternative explanation: that participants might spatially bind the bad (vs. good) outcomes to the immoral behaviours because they are similarly valenced, rather than because they are perceiving bad outcomes as consequences of the immoral actions *of the people on whom they befall*. Although Frank might deserve a bad outcome because *he* was a bad person, Geoff doesn't deserve a bad outcome simply because Mark was a bad person. Thus, although there is a congruency between the value of the actions and outcomes in both cases, we expected congruency to affect binding more strongly when the outcome occurred to the person who deserved it (i.e., the same person).

Participants were randomly assigned to one of two sets of 8 scenarios (2 scenarios per each condition). The scenarios were the same across sets (i.e., the good/bad action by the protagonist and the ensuing outcome) except that the names of the characters changed (e.g., the "John embezzled funds/John's apartment was flooded" scenario in set 1 became the "Mark embezzled funds/Tony's apartment was flooded" scenario in set 2). In other words, across sets the scenarios changed conditions, such that a 'same person' scenario in Set 1 became a 'different person' scenario in Set 2, in order to fully counter-balance any individual effect of scenario. Participants responded to the scenarios in succession and in a random order. Finally, participants reported their age and gender. We operationalised the perceived distance between the bad behaviors and outcomes as the absolute difference between the positions participants placed the two sentences. Thus, scores could range from 1 (the sentences were placed directly next to each other) to 7 (the sentences were placed as far apart as possible), with higher values indicating greater perceived spatial distances.

Results

Distances between the behaviors and outcomes were subjected to a 2 (Person) x 2 (Congruence) x 2 (Set) mixed ANOVA. Analyses revealed a main effect for Person, $F(1,78) = 74.59$, $p < .001$, $\eta p^2 = .489$, and Congruency, $F(1,78) = 41.83$, $p < .001$, $\eta p^2 = .349$, showing that sentences were placed significantly closer together when they were morally congruent and involved the same person. More importantly, as shown in Figure 2, there was a significant Person X Congruency interaction, $F(1,78) = 23.70$, $p < .001$, $\eta p^2 = .233$. Furthermore, the two sets did not differ in sentence placement, $F(1,78) = 0.18$, $p = .894$, and there was no significant interaction involving the sets and the within-subject factors: Person (2) X Set (2), $F(1,78) = 0.29$, $p = .589$; Congruence (2) X Set (2), $F(1,78) = 0.27$, $p = .606$; Person (2) X Congruence (2) X Set (2), $F(1,78) = 0.29$, $p = .591$. Referring to Figure 2, this effect appears to be driven by a combination of sentences being both congruent and involving the same person. Paired-samples t-tests revealed comparisons for movements of outcomes closer to the immoral actions when they were congruent (vs. incongruent) were significant when the events occurred to the same persons ($t(79) = 6.84$, $p < .001$). This effect was also present for different persons, ($t(79) = 2.88$, $p = .005$), but to a lesser degree.



Chapter 5, Figure 2. The effect of moral congruency on the distance between the immoral behaviors and outcomes as a function of whether the outcomes occurred to the same vs. different persons (Experiments 8 and 9). Error bars show 95% confidence intervals of the mean.

Discussion

In Experiment 8 we demonstrated that two morally congruent sentences that form a scenario are placed in closer proximity to each other than morally incongruent sentences. Additionally, this effect was bolstered when the scenario events involved the same person, although there was a significant effect of moral congruence for different people as well. The interaction between Congruence and Person was also significant. Inspecting the mean scores/Figure 2, there is a clear advantage for events occurring to the same person in terms of spatial proximity. This indicates that, while congruent valence of the sentences may contribute to an individual's desire to place them closer in space, a larger influence is whether the events involved the same person. In terms of immanent justice reasoning, morally congruent events involving the same person reflect events that fit a "Just World" narrative (Lerner, 1980), forming a stronger causal relationship (reflected via spatial contiguity) than events involving different people.

However, the presence of a main effect of congruence suggests that IJR may be a contributory factor, rather than single cause for spatial proximity placements. In other words, morally congruent sentences were placed significantly closer together when events occurred to the different people compared to incongruent sentences.

Experiment 9

In Experiment 9, we wished to replicate these findings to solidify our observations. We developed a more flexible paradigm within which, instead of ranks that may limit spatial selection, participants could, by using their mouse, freely place one sentence of the scenario around the computer screen.

Method

Participants. Sixty-two participants completed an online study (30 males, 31 females, 1 preferred not to say: M_{age} = 34.79, SD_{age} = 11.23).

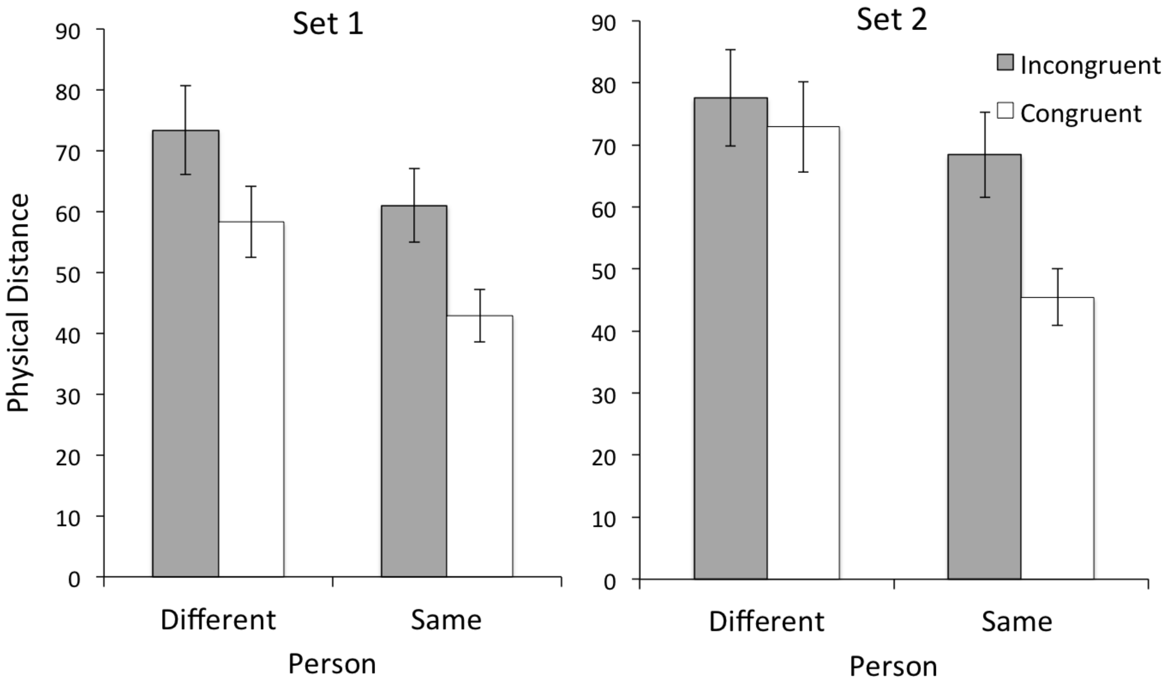
Materials and procedure. Participants first reported their age and gender. Experiment 9 was the same as Experiment 8 except that participants could freely move the "outcome" sentence to any position on the screen. The sentence describing the immoral action always occupied a fixed position in the centre of the screen. Each sentence was fixed to a black circle positioned directly above the middle of the sentence (see Figure 1). For each trial, the starting position of the moveable sentence was below the stationary sentence, two-thirds vertically of the display-screen. We asked participants to position the circles as close or as far apart as they wanted. Once they had clicked to select the position of the moveable circle, they could confirm their choice, or cancel to reposition the circle. We operationalised the distance between the deeds and outcomes as the absolute relative distance between the circles. We first calculated the Euclidean distance in pixels between the centre of the screen (the location of the fixed circle), and the selected position of the moveable circle. We then divided this by the maximum possible distance in pixels (i.e. from the centre of the screen to the corner of the screen), taken from the screen resolution of the participant's computer. We took the absolute value of this score, to give a positive score regardless of the direction participants chose to move the ball. We then multiplied this value by 100, to give a score representing the percentage of the maximum possible distance the two balls could be separated. Experiment 9 was created in Inquisit 4 (Millisecond Software, Seattle, WA) and based on Lyons-Warren, Rema, and Hershey's (2004) script for a spatial delayed response task.

Results

The distances between the circles (behaviours and outcomes) were analysed using a 2 (Person) x 2 (Congruence) x 2 (Set) mixed ANOVA as per Experiment 1. Analyses revealed main effects for Person, $F(1,60) = 51.01$, $p < .001$, $\eta p^2 = .460$, and Congruency, $F(1,60) = 37.68$, $p < .001$, $\eta p^2 = .386$. Consistent with Experiment 1, there was a significant Person X Congruency interaction, $F(1,60) = 7.96$, $p = .006$, $\eta p^2 = .117$ (see Figure 1).

Referring to Figure 2, similar to Experiment 8, this effect appears to be driven by a combination of sentences being both congruent and involving the same person. Paired samples t-tests revealed that participants positioned the circles associated with the immoral behaviours closer to the circles associated with the outcomes when they were congruent (vs. incongruent). This effect was significant both when the sentences depicted the same person, ($t(61) = 5.54$, $p < .001$), and when the sentences depicted two different people, ($t(61) = 4.05$, $p < .001$). Nonetheless, the presence of a significant interaction confirms that the effect was significantly stronger in the same person condition than the different person condition.

Furthermore, the two sets did not differ in sentence placement, $F(1,60) = 2.60$, $p = .112$, and nor did the sets significantly interact with person, $F(1,60) = 0.97$, $p = .329$, or congruence, $F(1,60) = 0.32$, $p = .574$. However, a three-way interaction, Set (2) X Congruence (2) X Person (2), was significant, $F(1,60) = 4.03$, $p = .049$. As shown in Figure 3, this effect appears to be driven by a difference between scenarios involving different persons when the action/outcome of the scenario were congruent.



Chapter 5, Figure 3. The effect of moral congruency on the distance between the immoral behaviors and outcomes as a function of whether the outcomes occurred to the same vs. different persons (Sets 1 and 2, Experiment 9). Error bars show 95% confidence intervals of the mean.

Discussion

Experiment 9, affording participants more flexibility in their spatial proximity placements, reinforced our findings from Experiment 8. Specifically, morally congruent sentences that form a scenario are placed in closer proximity to each other than morally incongruent sentences, particularly when these sentences involve the same person. Additionally, as the main effect of 'Person' ($F = 51.01$) was relatively much larger than the main effect of 'Congruence' ($F = 37.68$), spatial proximity placements appear to be more influenced by events that involve the same person rather than moral congruence, akin to Experiment 8, illustrated by Fig. 2.

As we consistently showed that sentences occurring to the same person generate more proximal placements compared to different persons, we removed the person factor from Experiment 10.

Experiment 10

The task we used in Experiment 10 was more active and dynamic than the tasks in Experiments 8 and 9. Participants controlled a moveable representation of a good or bad outcome and were asked to maintain its distance from a "chasing" ball representing a target's immoral behaviour. The rationale for this approach is that while experiments 8 and 9 represented spatial binding in the form of static decision making, phrasing within imminent justice reasoning literature often speaks of "maintaining" a commitment to justice in the face of threat (Callan et al., 2014). Experiments 8 and 9 demonstrated participants' spatial representations of their feelings of proximity towards the sentences. In Experiment 10, we wanted to explore the effort behind maintaining a commitment to justice in the form of spatial binding, where participants had to continuously move the outcome sentence of the scenario while the descriptor sentence "chased" the participant's sentence. The average distance between the two sentences would then quantify a participant's effort to maintain a specific spatial proximity between the two sentences based on the participant's feelings of how close the sentences should be.

Method

Participants. We recruited 68 participants (40 males, 27 females, 1 preferred not to say: M_{age} = 33.60, SD_{age} = 9.48; 66 participants were requested through MTurk).

Materials and procedures. In Experiment 10, after reporting their age and gender, participants controlled the movements of a circle representing the outcome (a white circle) while this was "chased" by a circle representing the initial deed (a red circle; see Figure 1). Both circles started each trial in the center of the screen on the vertical plane, and at 30% and 70% of the screen on the horizontal plane. Participants pressed one of the arrow keys to move the white circle. The circle moved one-step (4% of the screen) every 300 ms provided the participant had pressed during that time. The red circle moved one-step (2% of the screen size), in the direction of the white circle every 300ms. Each trial lasted 15 seconds, and participants could move the white circle away from or towards the red circle to keep them as close together or as far apart as they wanted them to be. We calculated the distance between the two circles on each sample (every 300ms) using their horizontal and vertical positions given as a percentage of the screen (i.e. 0%,0% for the top left corner, 50%,50% for the center, and 100%, 100% for the bottom right corner). The distance was calculated by summing the vertical distance and the horizontal distance on each sample. We calculated the average distance across the course of the trial by taking the mean distance over each of the 50 samples.

Given that our Experiments 8 and 9 showed that the predicted effect of congruency on perceived distance between the behaviours and outcomes was not simply due to them being similarly valenced (i.e., because the effects of congruency on perceived distance were weaker when the outcome occurred to a different vs. the same person), we did not include a Person factor in Experiment 10. Following a practice trial with neutral sentences (as per Experiment 8), participants responded to four congruent (e.g., "Joe robbed a store at gun point" and "Joe's apartment was destroyed by fire") and four incongruent (e.g., "Tim kidnapped a child" and "Tim received a major pay rise at work") scenarios. Participants were randomly assigned to 1 of 8 sets of 8 scenarios (see osf.io/37g9k/); the scenarios were presented sequentially in a random order across participants.

Results

The distances participants kept the circles associated with the outcomes from the "chasing" circles associated with the deeds were analysed with a 2 (Congruence) x 8 (Set) mixed ANOVA. Analyses revealed a significant effect for Congruency, $F(1,67) = 45.50$, $p = <.001$, $\eta p^2 = .404$, such that participants actively kept the outcome circles closer to the "chasing" immoral action circles when they were congruent ($M = 16.42$, $SD = 7.65$) than when they were incongruent ($M = 22.41$, $SD = 6.04$). Like the previous two experiments, between-subjects analysis showed that there was no significant difference between sets, $F(7,67) = 0.92$, $p = .445$.

General Discussion

Events that are causally linked appear closer together in space and time (Faro et al., 2005; Buehner & Humphreys, 2010). Across three studies, we show for the first time that even events with no logical causal association feel closer together when they satisfy our need to perceive that people get what they deserve (Lerner, 1980). Negative life events were positioned closer to previous immoral behaviours than were positive life events (Experiments 8 and 9). Furthermore, when controlling the movement of a positive or negative life event that was being chased by a bad person, participants kept positive outcomes further away from bad people (Experiment 10). This suggests that not only did participants position bad events and bad people closer together, they actively tried to keep good events away from bad people.

Taken together our findings show that the fallacy of immanent justice reasoning extends beyond a belief that bad behaviour causes negative life events, such that this belief generates a bias in the spatial proximity of these two events. This effect was not merely driven by participants being more inclined to group together two negative sentences, since the effect was significantly larger when the two events were described with the same protagonist. As such, although a cheating spouse and a car crash victim might feel closer together, this is particularly the case when the negative event has befallen the person committing the immoral act.

The current studies contribute to the growing literature of immanent justice reasoning by extending beyond self-report measures ordinarily employed (Callan et al., 2014). Where spatial binding research demonstrates the perceived contraction in space between causally linked events, we equally demonstrate that beliefs concerning justice can impact the spatial proximity between two complex, social events. In other words, the cognitive and motivational underpinnings of immanent justice reasoning extend to implicit measures, and our desire to maintain justice against threats not only affects our causal reasoning, but our sense of spatial contiguity. Additionally, given that, other than chance, no rational explanations to the sequences of events are possible, our findings highlight that our inferences of causality dependent upon spatial proximity may be as equally subject to non-rational causal systems as our explicit explanations of why bad outcomes happen to bad people.

From just-world theory, we would assume that this pattern occurred due to the perception that negative (positive) outcomes were more deserving of negative (positive) outcomes. The relationship between deservingness and immanent justice reasoning is tightly linked, where the more deserving of the outcome an individual is perceived to be, the more likely that events that befall them will be causally attributed to them if they are morally congruent (Callan et al., 2014). This presents the opportunity to solidify our findings: future studies could ask participants to rate the deservingness of scenarios and using mediational analysis could determine the extent to which spatial proximity of morally charged sentences is underpinned by concerns for deservingness.

However, an alternative explanation of closer proximity placements for morally congruent sentences may be confounded by the extent to which the sentences are perceived as similar. Casasanto (2008) found that participants rated stimulus word pairings (e.g., memory-hope) as more conceptually similar the closer together the words appeared on screen, arguing that people use spatial proximity as a cue for judging similarity. Thus, participants may have positioned morally congruent sentences closer together on the basis of similarity; for example, “Neil assaulted a child” and “Neil slipped and fractured his leg” are similar given that they are both negative and involve physical harm, rendering the decision to remove the person factor from Experiment 10 potentially problematic. Thus, future work could additionally acquire ratings of similarity of sentences and using scenarios as the unit of mediational analysis would allow us to address further the issue of whether congruence per se is associated with the spatial binding of bad outcomes to immoral actions or whether conceptual similarity between actions and outcomes might be confounding this association. Specifically, where we showed a significant main effect of congruence driving sentence placements, such an effect may divide between congruence and similarity, denoting which produces the most influence.

Consideration should also be given to the possibility that response bias (responding consistently with beliefs about what the experiment wants to find) may have contributed towards participants' decisions of sentence placement. In other words, closer spatial proximity between morally congruent sentences (and vice versa for incongruent sentences) may have been driven by participants presuming upon the experiment's aims and expectations. In order to minimise any contribution from response bias in future experiments, one option would be to conduct a between-subjects style design, where participants are divided by moral congruence, such that participants would only see morally congruent or incongruent scenarios. Alternatively, asking participants at the end of the experiment whether they could guess the hypothesis of the experiment would allow the removal of those participants who had guessed correctly.

A further caveat to our experiments, however, is that participants chose to position events based on how close they felt the events should be. Although participants were asked to place events according to how close they *felt* the events should be, participants actively chose the spatial proximity, rather than perceiving a particular distance, as is typical in spatial binding research (Buehner & Humphreys, 2009). Therefore, these experiments did not deal with perception per se, but rather the conscious decisions made that reflect spatial contiguity, and so our experiments cannot be fully labelled as an implicit measure. Nevertheless, even if inferences of causality drawn from spatial proximity cannot be relied upon in terms of perception of events, then at least we appear to spatially relate events closer together in terms of their causal relatedness in conscious decision-making. Subsequently, these experiments do extend previous IJR findings beyond the overtly explicit self-report measures of which are typically employed.

The next step in this line of research should be to ascertain whether the desire for spatial proximity between events extends to the perceptual level, closer in line with previous spatial binding research (Buehner & Humphreys, 2010). For example, a ball representing a good or bad action could contact a one side of an object, launching a second ball representing a morally (in)congruent outcome on the adjacent side, and participants could then replicate the size of the object, with smaller sizes suggesting a stronger causal connection between the two, as per the archetypal Michottean launching task (Michotte, 1963).

To extend this line of investigation even further, and, given that children demonstrate less immanent justice reasoning that adults (Raman & Winer, 2004), either measure of spatial binding (free placement or perceptual) could be compared between adults and children to explore whether the implicit bias we have found parallels the existing explicit causal reasoning bias shown throughout development. Naturally, as has been shown in previous research, we might equally expect that spatial binding would inflate with prior exposure to an unrelated, innocent victim's suffering (Callan et al., 2006) or to long-term goal focus (Callan, Harvey, et al., 2013).

Overall, we have shown that when making sense of the world around us, our experience of the spatial proximity between events is warped to fit the idea that the world is a fair and just place; bad people really do feel close to bad outcomes.

Chapter 6: General Discussion

General Discussion

The aim of this thesis was to investigate how attributions of causality, specifically the temporal binding between actions and their outcomes observed within sense of agency (SoA) research, translate to social, meaningful contexts. Moreover, a secondary aim of this thesis was to provide further input as to the specific mechanisms of temporal binding as a measure of SoA and implicit causal attribution in general.

As my PhD research progressed, the findings from my experiments motivated a shift in theoretical focus: not only did I fail to replicate previous research suggesting that temporal binding, and thus SoA, could be modulated by social factors (Chapter 2), this equally prevented the line of investigation I wished to pursue across the previous three years of academic research. Given that the aim of my thesis was to investigate implicit causal beliefs within the context of SoA first, and continuously adapt my experiments to further socially relevant settings and known explicit biases in causal reasoning/inference, again, the findings from Chapter 2 suggested our approach be modified.

As the specific outcomes used in Chapter 2 involved emotional valence, we shifted our approach by exploring temporal binding with immanent justice reasoning (IJR) in Chapter 3. However, similar to Chapter 2, we again found no modulation of temporal binding. The findings from both chapters 2 and 3 led to the experiments contained within Chapter 4, where our approach was further modified to examine under what conditions causal beliefs modulate the perceived time between causally related events. In Chapter 4 I extended/replicated two well-cited experiments in an attempt to clarify the mechanisms of temporal binding in terms of causal perception and low-level factors such as temporal prediction. The findings of Chapter 4, however, only provided slightly more clarity over why two causally related events become temporally attracted to each other.

Thus, chapters 2–4 construct a narrative where temporal binding, measured with a variety of methods, remained mostly impassive to modulation, regardless of physical or abstract causality, with a tentative conclusion drawn from Experiment 6 that causality is not wholly necessary to produce binding. As a result, my research transferred over to spatial binding, where I explored how otherwise causally unrelated events could influence feelings of spatial contiguity due to the moral valence of the events (Chapter 5). Chapter 2 has subsequently been published in the journal Consciousness and Cognition, and Chapter 5 has been submitted for publication at this time of writing.

Summary of Findings

Experiments 1 – 4 investigated if and how temporal binding is modulated by emotional valence. In accordance with the aims of my thesis, these experiments were conducted to establish a basic effect that temporal binding, and by extension, our SoA, could be modulated by social outcomes. However, Experiments 1 and 2 found no significant difference in binding between positive and negative emoticons (Experiment 1) or positive and negative real facial expressions (Experiment 2). Experiment 3 revealed that the stimuli used in experiments 1 and 2 were equivalent in valence and arousal to stimuli that have previously been observed to modulate temporal binding (Yoshie & Haggard, 2013). Finally, in a highly powered replication study (Experiment 4), we observed no significant modulation of temporal binding by emotionally valenced vocalisations. This occurred despite the original paper's (Yoshie & Haggard, 2013) reported findings that negative vocalisations produced significantly less binding.

Experiment 5 investigated the extent to which unconnected events could be causally linked due to their moralistic congruency, measured via temporal binding. This experiment sought to both further support previous explicit causal attributions within immanent justice research, as well as investigate the effectiveness and accuracy of temporal binding as a tool with which to measure implicit causal attribution. However, using estimations of the interval between a scenario depicting either a good or bad individual and a subsequent good or bad outcome elicited no modulations of temporal binding. Similar to Chapter 2, this experiment questioned temporal binding's usage as a measure of implicit causal beliefs; at least, to the extent that social outcomes can have a direct influence (e.g., Christensen et al., 2016; Yoshie & Haggard, 2013).

Experiments 6 and 7 investigated causal binding with regards to self- and other-actions. These experiments were conducted to explore whether top-down processes (knowledge of a causal relationship) were sufficient to produce the binding phenomenon when measured via stimulus anticipation. Moreover, these experiments directly follow those from Chapter 2, which questioned temporal binding as measure of SoA, and by extension, its ability to detect any modulations in implicit causal beliefs. Experiment 6, an extension of Buehner (2012), found significantly stronger temporal binding for self-caused actions at 500ms compared to a baseline condition. A time-sequence condition, where participants observed a time-sequence counting upwards and a signal occurred within a 2 – 5 second range, did not differ from either of the other two conditions. At 900ms, participants anticipated the target stimulus significantly more so than both baseline and self-causal conditions, with no difference observed between the latter two. Our findings partially replicate Buehner (2012), who found greater temporal binding for his self-causal and machine-causal conditions at 500ms (compared to baseline), and more temporal binding for the machine-causal binding at 900ms.

Experiment 7 failed to replicate Desantis, Roussel, and Waszak (2011)'s findings that causal beliefs modulate temporal binding. Specifically, in our study, there were no differences between self- and other-outcome stimulus anticipations.

Due to our inability to find sufficient evidence of temporal binding being modulated by either emotional valence, causal beliefs, or justice-based reasoning, experiments 8—10 investigated the extent to which explicit self-report measures used within immanent justice reasoning research replicates with spatial binding. Specifically, participants chose where to place two objects representing sentences that dictated the moral worth of an individual (e.g., a good or bad person) and an outcome occurring to them later on (good outcome or bad outcome). Per previous research (e.g. Callan et al., 2006; Buehner & Humphreys, 2009), we expected closer positioning of the two events when they were morally congruent, but only when the events concerned the same person. As a result, the findings of Chapter 5 showed that participants positioned representations of people's fortuitous outcomes significantly closer in space to representations of their previous immoral actions when the outcomes occurred to the same person (Frank punched someone - Frank was in a car accident) more strongly than when the outcomes occurred to a different person (Frank punched someone - Joe was in a car accident) throughout experiments 8 (ordering sentences in a ranked list) and 9 (freely placing one event around the screen). Experiment 10 showed that perceived deservingness predicted spatial position, with participants acting to keep good outcomes and bad people apart, as compared to bad outcomes and bad people.

Discussion of key variables – Sense of Agency and Temporal Binding

Research within the SoA literature suggests it is prone to influence from external factors (Aarts et al., 2007; Buehner & Humphreys, 2009; Engbert et al.,2007; Takahata et al., 2012). Specifically, Yoshie & Haggard (2013) suggested that, via the self-serving bias (Campbell & Sedikides, 1999), outcomes with a negative valence extend the perceived temporal interval between action and outcome. Conversely, positive outcomes induce actions and outcomes closer together. The findings from Chapter 2, however, seriously question this assertion. A more recent study, citing the research within Chapter 2, replicated our findings where valence had no impact on binding (Barlas, Hockley, & Obhi, 2017). Recent evidence by the same authors as those whose experiment was replicated in Experiment 4 (Yoshie & Haggard, 2013) conducted a further laboratory experiment investigating the effect of predictability of emotionally valenced outcomes (Yoshie & Haggard, 2017). Whilst no effect of unpredictable outcomes was found, predictable positive outcomes produced significantly more binding than negative outcomes. This effect was solely due to outcomes being perceived as earlier, rather than actions being perceived as later. This is slightly inconsistent with their previous paper (Yoshie & Haggard, 2013, supplementary materials), where they report both action and outcome shifts influenced the overall binding effect (though the effect was still predominantly driven by outcome binding). Again, this is also inconsistent with our data, where the majority of the composite binding observed was due to action binding, although did not reach a statistically significant level.

Thus, at least as far emotional valence goes, temporal binding does not appear to be modulated by such external factors. Future work should therefore attempt to replicate and extend other examples of self-serving bias in temporal binding (Aarts et al., 2012; Takahata et al., 2012) and sensory attenuation (Gentsch et al., 2015) to further advance our understanding of how (or if) outcome valence influences implicit agency.

Moving forward, not only do the results from Chapter 2 suggest that emotional valence does not influence temporal binding, but they also challenge the idea that the self-serving bias interacts with temporal binding (and by proxy, the SoA). However, some evidence suggests that positive outcomes do engender more SoA than negative outcomes using explicit causal judgements: Oishi et al.(2018) asked participants to move a controllable dot to a target whilst the target was of a specific colour. The dot contacting the target during the correct colour denoted a success trial. Conversely, an incorrect colour denoted a failed trial. On a 0-100 scale, participants significantly rated more control over the dot during success vs. failed trials. Thus, concerning sense of control over our actions measured explicitly, the self-serving bias may influence our judgement. That we failed to find any such influence with temporal binding methodologies leads to two interpretations: my thesis further contributes either 1) to the doubt that temporal binding measures the SoA, or 2) to the evidence that feelings (implicit) and judgements (explicit) of agency are distinct concepts, and that judgements of agency are more susceptible to cognitive biases such as the self-serving bias (Synofzik et al., 2008). Where feelings of agency refer to the implicit SoA that temporal binding is purported to capture, my thesis suggests that the self-serving bias may not interact with this aspect of SoA, and, as Oishi et al. (2018) demonstrate, may only hold influence over judgements of agency. Recent evidence also suggests that neither sensory attenuation (Dewey & Knoblich, 2014) nor temporal binding (Dewey & Knoblich, 2014; Saito et al., 2015) correlate with explicit reports of agency. While explicit and implicit measures will never show total convergence, positive evidence of covariation is important to argue that conscious reports and unconscious biases are indeed measuring the same underlying process.

Another important debate with regard to temporal binding is the extent to which this reflects agency per se., or simply causality. Chapter 4 of my thesis investigated how causal beliefs, and not intentional actions, modulate temporal binding, in order to shed some light on the matter, using stimulus anticipation. Specifically, we extended previous research (Buehner, 2012) in order to test whether knowledge of a causal mechanism to produce a stimulus change would elicit temporal binding. Buehner (2012) compared self-caused actions to machine-caused actions to investigate whether knowledge of a causal mechanism was sufficient to produce binding. Buehner found that both self- and machine-caused actions produced greater binding overall than passively observing a signal indicating the sequence of events. Unlike Experiment 4, we managed to partially replicate this original paper, where at 500ms self-caused targets were anticipated earlier than those in the baseline condition. Our time-sequence condition that replaced Buehner's machine-causal condition, however, did not differ from either baseline or self-causal conditions.

Thus, at a shorter time interval, we appear to reinforce Buehner's findings. However, at 900ms, outcomes in the time-sequence condition were anticipated more than the other two, with no distinction between self-causal and baseline conditions. Experiment 6 showed that, at least in part, the causal binding presented within Buehner (2012) might have been partially driven by processes other than an understanding of causality between events. The lack of self-causal binding at 900ms suggests that intentionality, despite that it may be neither necessary nor sufficient to solely produce binding (Buehner & Humphreys, 2009; Obhi & Hall, 2011), may still contribute to an overall binding effect at shorter intervals. As previous research suggests temporal binding for self-causal actions may weaken beyond a certain threshold, where voluntary actions produce temporal shifts at short intervals (Engbert et al., 2008; Haggard et al., 2002), intentional action might explain the findings for our self-causal condition between 500ms and 900ms.

Such a conclusion is reinforced by the quantity of studies suggesting that voluntary actions *do* produce temporal binding (David et al., 2008; Moore et al. 2009a; Moore and Haggard 2008; Tsakiris and Haggard 2003), and that intentional binding, i.e. temporal binding with intentional actions, is a subset of causal binding as Buehner (2012) suggests, where intentional actions contribute to causal predictivity of outcomes, drawn towards the action under the principles of temporal contiguity. This line of reasoning falls within Bayes theorem of ambiguity reduction, where judgement under uncertainty necessitates the use of multiple signals to arrive at the most probabilistic cause of an outcome.

However, the influence of intentionality has been found primarily through measuring temporal binding via the Libet clock or temporal estimation methods, as opposed to stimulus anticipation. Moreover, Buehner (2012) showed inconsistent patterns in the self-causal condition between his Experiments 1 and 2, finding less self-causal binding in the latter when participants initiated the trial with a preliminary button press. This may either indicate that stimulus anticipation is less susceptible to the influence of intentionality of action, or that this method may be less reliable than other well-tested measures.

Hence, another possible interpretation of Experiment 6's results portends to the reliability and/or effectiveness of stimulus anticipation serving as a measure of temporal binding. This was reinforced in Experiment 7 where we found no difference between the self-causal condition and baseline measures. Such a result conflicts with that of Experiment 6, which did show a significant difference between self-caused actions and passively observing events in the baseline condition, albeit only at 500ms. Furthermore, given that only three known prior publications have utilised stimulus anticipation to measure temporal binding (Buehner & Humphreys, 2009; Buehner, 2012; Blakey, et al., 2019), our findings, and lack of replication in Experiment 7, suggest that stimulus anticipation, as compared to other known methods (e.g., temporal estimation, Engbert & Wohlschläger, 2007), is not as reliable.

Speculation might also posit that stimulus anticipation's inability to detect differences between self-caused outcomes and baseline measures is due to stimulus anticipation's nature as an active task, meaning actions are actively timed to the onset of when an event is thought to occur. The Libet Clock/temporal estimation tasks, on the other hand, are retrospective judgements free from action planning, motor coordination and active time computation in order to accurately align action and event onset. Therefore, with increased cognitive load, stimulus anticipation may be less susceptible to response bias than other measures (e.g., Libet Clock/temporal estimation) that might enable biased temporal estimations for shorter/longer temporal interval durations based on experimental condition. However, given the limited publications on stimulus anticipation binding (Buehner & Humphreys, 2009; Buehner, 2012; Blakey, et al., 2019), none of which discuss the actual mechanism underlying stimulus anticipation beyond stating that if temporal binding is present, the interval between the two events in the causal condition relative to the non-causal condition should be underestimated, future work would need to establish the cognitive load potential of stimulus anticipation relative to other measures.

Conversely, our findings may otherwise suggest that top-down mechanisms, such as the existence of a causal relation between events, might not modulate temporal binding. Specifically, that efforts to convince participants that a machine acting as a causal agent, and causal beliefs regarding the author of outcome, are not necessary.

As has been discussed in Chapter 1, efferent motor signals, espoused by the comparator model of SoA, have been proposed to produce binding (Haggard & Clark, 2003), but may not be strictly necessary (Moore et al., 2013; Poonian & Cunnington, 2013). Intentional actions have been shown to elicit binding in some research (Haggard et al., 2002), but are also not strictly necessary (Buehner, 2012; Buehner & Humphreys, 2009). Knowledge of the causal mechanism between two events may influence temporal binding (Buehner, 2012), but, as Chapter 4 demonstrates, may also not be necessary, as is the case at 900ms where a simple visual counter, giving a predictable temporal window of the initial event, produced greater temporal binding for the following event than self-caused actions.

The difference between our self-causal condition and baseline can be attributed to the aforementioned notion that the effect of intentional action diminishes over time. However, what other mechanisms remain as to explain how temporal binding for our time-sequence condition was greater than self-caused actions at a later interval? In this condition, no voluntary action was performed, and therefore proponents of the comparator model would be unable to answer. Equally, no notion of causality was apparent to participants, who passively observed one event following another.

In their review Hughes et al. (2013) suggest four factors that influence prediction of a stimulus or action-effect: temporal prediction, temporal control, non-motor identity prediction, and motor identity prediction. Applying a contrast to our conditions in Experiment 6, the self-causal condition is left with both temporal control over the outcome (using one's action to control the point in time at which a stimulus will occur), temporal predictivity (the ability to predict the onset time of when the stimuli will occur) and motor identity prediction (prediction of the identity of a sensory event based on an action) over the time-sequence and baseline condition. However, both the baseline and time-counter conditions contained the same processes of temporal predictivity and motor identity prediction. One conclusion, therefore, is that temporal control over the onset of outcome may only generate temporal binding at shorter intervals. However, this still leaves the disparity between the time-sequence and baseline conditions at a longer interval of 900ms, which, despite the same processes outlined in Hughes et al., differed on two accounts: the relative temporal predictivity of the signal stimulus, and the visual representation of time. Specifically, the baseline condition was predictable within a 500ms margin, with the signal stimulus occurring between 2,300–2800ms after trial initiation. - Conversely, the time-sequence condition was predictable within a 3000ms margin, between 2–5s. Note, this difference between conditions is separate from the processes listed in Hughes et al. (2013), which concerned the relative effects on action-effect prediction, whereas the differences concerning the baseline and time-sequence condition only impacted the cue of the action-effect. Given that the baseline condition would appear more temporally predictable than the time-sequence condition, we may rule this out as to why, at a longer interval, the time-sequence yielded far greater binding (ignoring the line of reasoning that the more unpredictable the cue of a stimulus is, the more it is anticipated).

A remaining factor that differentiated between the time-sequence and baseline conditions was the visual depiction of the timer, placed where the signal would occur. This visual depiction of time may have employed more visual and temporal attention to the oncoming stimulus cue. Research measuring sensory attenuation, the counterpart to temporal binding when measuring SoA, has shown increased attention modulating sensory attenuation of auditory stimuli, such that more attention leads to lower N1 potentials (Lange, 2009). Sensory attenuation implies a sufficiently correct match between predicted and actual sensory feedback, which, although distinct, complements the perceived temporal contiguity between events. Thus, temporal attention may orient participants towards a sensory event that predicts a second event, causing the second event to appear closer in time. Furthermore, this influence would be more noticeable at longer intervals, where shorter periods enable easier prediction of stimulus timings (for a review, see Niemi & Näätänen, 1981). However, this is only speculation.

In a related study, Haggard & Cole (2007) found that attention to a specific event (intention, action, or action-effect) produced less binding than when participants were instructed to estimate the onset time of the specific event after the trial. However, the temporal interval for all trials was 250ms, and the authors employed the Libet clock method, which due to involving retrospective time estimations (as opposed to the aforementioned dynamic nature of stimulus anticipation), may engage different anticipatory properties to stimulus anticipation. For example, Spence, Shore and Klein (2001) showed, using a temporal order judgement task where an individual is asked to judge the temporal order of a sequence of stimuli, that attending to a particular modality speeds up the relative temporal processing of that modality. Hence, in Experiment 6, visually attending to the timer may have caused a shift in the temporal processing speed of visual information, leading to earlier anticipations of following visual stimulus events. Future research could explore this further by varying the duration of visual attention before presenting a following stimulus to be anticipated to observe any modulatory effects. One may also argue that the self-causal condition would have greater attention than passively observing a point on the screen, which may explain our finding at shorter intervals.

Overall, the role of attention in temporal binding still requires clarification. Experiment 6, although not primarily aimed at investigating attention, found results that may serve to base further research upon. Several possible options involve either using a visual stimulus unrelated to the counter in place of where the counter would be, or to have the counter in a different location, in order to distinguish between visual and temporal attention. A counter could equally be displayed in the self-causal condition in order to fully distinguish between the roles of temporal attention and control. Lastly, multiple onset times could also be utilised in order to observe a fuller picture of linear trends in binding across short and long intervals.

Discussion of key variables – Immanent Justice Reasoning

Despite the lack of success in observing modulation of temporal binding by socially salient outcomes (Chapter 2), or by modifying methods that manipulated causal beliefs (Chapter 4), we persevered with our original aim of this research project, that the perceptual error phenomena observed within SoA (Moore & Obhi, 2011) and causality (Buehner, 2012) research, in chapters 3 and 5. Given that our findings from Experiment 6 challenged the notion that knowledge of a causal relationship between events directly and significantly impacts binding, we were cautious to accept such a finding without scepticism. Moreover, such findings may have arisen due to stimulus anticipation's reliability as a measure of temporal binding, where causal events, such as intentional action producing a stimulus, invoke perceptual contraction of the temporal interval between the events (as has been observed via temporal estimation; Fereday & Buehner, 2017; Humphreys & Buehner, 2009; Kumar & Srinivasan, 2017). Thus, progressing from the well-established area of immanent justice reasoning (IJR; Callan et al., 2014), we sought to examine how the modulations of explicit causal judgements due to concerns for justice translated to temporal (using temporal estimation) and spatial binding.

Experiment 5 of Chapter 3 showed that the moral congruence of events does not modulate temporal binding. Put simply, it did not matter whether a person described as either good or bad encountered a fortuitous good or bad event – temporal binding was consistent with typical interval estimations of those intervals regardless of condition. As stated in the relevant discussion section, three possible conclusions explain the results of Experiment 5: 1) the experimental mechanics, i.e., use of a temporal binding measure between a vignette and visual outcome, were not appropriate to detecting implicitly held causal beliefs; 2) IJR is not susceptible to implicit measures of causal relations between two events, or at least, the effect size is vastly small in comparison to its relative success using explicit measures (Callan et al., 2014); 3) temporal binding, as has been purported (Hughes et al., 2013), is less of a measure of causal relations, but is rather much more susceptible to an amalgamation of processes including temporal prediction, temporal control, and non-/motor-identity prediction.

This further lack of success is no surprise given the previous three experiments utilising temporal binding. And, as notable point, this is despite using both interval estimation and the Libet clock as the measure (or stimulus anticipation in further experiments). As such, the most reasonable conclusion to draw is that temporal binding may result from several factors, with potential emphasis upon weighted integration of sensory evidence, temporal control and non-motor identity prediction processes (Desantis et al., 2011; Hughes et al., 2013). However, for all the wealth of research defining IJR's strength as an influence over explicit causal judgements (Callan et al., 2006; Callan et al., 2014; Maes, 1998; Raman & Winer, 2002, 2004; Woolley et al., 2011), this phenomenon has yet to be extended to implicit measures.

In order to rule out the second conclusion, that IJR is not susceptible to implicit measures of causal inference, experiments 8—10 employed spatial binding to test whether people would actively keep apart morally incongruent events, or conversely place congruent events closer together. We hypothesised that spatial binding, the spatial equivalent to temporal binding, may procure more promising results after previous evidence demonstrates its relative consistency as a measure of causal beliefs (Buehner & Humphreys, 2010; Woods et al., 2012). Spatial binding, although less studied than its temporal counterpart, nevertheless retains evidence that causally-related stimuli attract each other in space (Buehner & Humphreys, 2009; Scholl and Nakayama, 2002). Specifically, the spatial distance between two events becomes perceptually contracted when a causal relationship exists between them. Given the strength of evidence behind IJR with explicit measurements (Callan et al., 2006; Callan et al., 2010; Callan et al., 2013; Harvey & Callan, 2014), and that the distance between two moving balls is judged to be smaller when the movement of the two balls are causally linked (Buehner & Humphreys, 2010), we expected that a robust measure of implicit causal beliefs will yield similar results.

Contrary to our experiments until now, Experiments 8—10 showed that when controlling the position or movement of a positive or negative life event, relative to the position or movement of a bad person, positive outcomes were kept further apart from bad people. This suggests that not only did participants position bad events and bad people closer together, they actively tried to keep good events away from bad people. One caveat may arise in the form that these experiments did not deal with perception per se, but rather the conscious decisions made that reflect spatial contiguity. Thus, our experiments cannot be fully labelled as an implicit measure, but they do extend previous IJR findings beyond the overtly explicit self-report measures of which are typically employed. Future work, in order to ensure covariation between explicit and implicit IJR, should also employ explicit measures such as those used within IJR research (e.g., "To what extent do you feel XXX caused YYY", Callan et al., 2006). Covariation would validate implicit measures, which, within an area largely based upon explicit measures, is important in order to generate confidence that the implicit measure is truly measuring the same underlying construct as the explicit measure.

Additionally, consideration should also be given to the possibility that response bias (responding consistently with beliefs about what the experiment wants to find) may have contributed towards participants' decisions of sentence placement. In other words, closer spatial proximity between morally congruent sentences (and vice versa for incongruent sentences) may have been driven by participants presuming upon the experiment's aims and expectations. Although response biases should always be controlled, it is difficult, if not impossible, to completely remove such biases from non-natural experiments, especially in relation to those involving decisions pertaining to social stimuli (Nederhof, 1985; Paulhus, 1991). Moreover, social desirability has been shown to be a separate construct to just world beliefs (Dalbert, Lipkus, Sallay, & Goch, 2001; Loo, 2002). Cognitive measurements of IJR utilising eye-tracking have also been consistent with the self-report style of typical IJR research (Callan et al., 2012), providing evidence that people genuinely expect or look for outcomes morally congruent to prior events. As such, our findings were weighted on IJR as a genuine phenomenon, as the aforementioned evidence suggests. A further way to minimise any contribution from response bias in future experiments may involve a between-subjects style design, where participants are divided by moral congruence, such that participants would only see morally congruent or incongruent scenarios.

Taken together our findings show that the fallacy of immanent justice reasoning extends beyond a belief that bad behaviour causes negative life events, such that this belief generates a bias in the spatial proximity of these two events. This effect was not merely driven by participants being more inclined to group together two negative sentences, since the effect was significantly larger when the two events were described with the same protagonist.

What the research within Chapter 5 demonstrates is that, in contrast to temporal binding, spatial binding may yield more of a true implicit measure of causality, though the research is beyond making a direct comparison between Hume's tenants of temporal and spatial contiguity and their interaction with the binding phenomenon in causal contexts. Furthermore, these results add to the discussion of feelings vs. judgements of causality: although participants were asked to place events according to how close they *felt* the events should be, there is the additional component that participants actively chose the spatial proximity, rather than perceiving a particular distance, as is typical in spatial binding research (Buehner & Humphreys, 2009). Therefore, even if inferences of causality drawn from spatial proximity cannot be relied upon in terms of perception of events, then at least we appear to spatially relate events closer together in terms of their causal relatedness in conscious decision-making.

Thesis Overall Contribution

This thesis contributes to the discussion of temporal binding as a purported implicit measure of SoA and causality. Specifically, the findings question whether temporal binding is indeed a measure of these notions at all, and although not directly tested, suggests other mechanisms may explain the findings seen throughout previous research. There is no doubt that temporal binding, as a phenomenon whereby actions and outcomes are temporally drawn together, does occur within laboratory settings. However, the question of why the interval becomes contracted requires significant clarification and cannot be wholly attributed to one mechanism or another. Many articles continue to assume the reasons behind temporal binding; for example, binding reflects SoA due to the influence of voluntary action over event/time perception (relative to involuntary action/outcomes occurring in isolation), specifically when said actions produce learned motor-identity outcomes (Desantis, Roussel, & Waszak, 2010; Engbert & Wohlschläger, 2007; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Kumar & Srinivasan, 2017; Moore, Lagnado, Deal, & Haggard, 2009; Moore et al., 2012). This assumption remains, and is still cited within, recent articles (e.g., Haggard, 2017) despite several articles suggesting otherwise; for example, those that show neither agency (Buehner and Humphreys, 2009; Buehner, 2012; Dogge et al., 2012) nor motor-predictive processes (Desantis et al., 2012; Hughes et al., 2013) are required for binding to occur.

In going beyond the immediate interpretation of each of the previous chapters, this thesis overall indicates that temporal binding, albeit an interesting phenomenon, might poorly translate to contexts that extend to higher-level processes that involve socially meaningful outcomes, causal inference of mechanical agents, or non-physical causality between events.

In terms of temporal binding arising from weighted cue integration, whereby the sensorimotor system combines information from different sources such as multiple sensory modalities and external information (Desantis et al., 2011; Ernst & Banks, 2002; Hillis, Ernst, Banks, & Landy, 2002; Moore & Fletcher, 2012; Moore & Haggard, 2008; Wolpe, Haggard, Siebner, & Rowe, 2013), this thesis fails to find support for external cues, such as causal knowledge and the emotional ramifications of causing negative outcomes, modulating temporal binding.

Much evidence supports the strength of internal cues on our action awareness and perception of ensuing sensory events (Balslev, Nielsen, Lund, Law, & Paulson, 2006; Frith, 2005; Frith, Blakemore, & Wolpert, 2000; Tsakiris, Haggard, Franck, Mainy, & Sirigu, 2005). Where explicit SoA is concerned, previous research typically suggests that both internal and external cues are optimally combined: for example, Farrer, Valentin and Hupé (2013) showed greater SoA attributions when both premotor signals (in an active task) and contextual information (a tone occurring at the time of action) were present.

Further research additionally purports that cue integration is limited to action binding, as opposed to outcome binding (Wolpe et al., 2013). However, this is inconsistent with research showing the extent to which external factors that modulate temporal binding largely (such as identity of the outcome) impact outcome binding more significantly relative to action binding (e.g., Yoshie & Haggard, 2013). Such findings (e.g., Wolpe et al., 2013) are nonetheless consistent with the research presented throughout chapters 2 and 3.

Additionally, of note is the continued parallel usages of 'temporal' and 'intentional' binding, where in reference to binding as an implicit measure of SoA, it is typically referred to as intentional binding (e.g., Haggard, 2017; Wen, Yamashita, & Asama, 2015); conversely, temporal binding is seen more frequently in research that does not necessarily refer to SoA (e.g., Blakey et al., 2019; Fereday & Buehner, 2017), despite using the same methods (e.g., Libet clock, temporal estimation, stimulus anticipation) to measure the same perceived temporal shift. Moreover, the applicability of these terms and the situational appropriateness of when they apply (if they are separately valid terms) has not, to my knowledge, been explicitly discussed within the binding literature.

Subsequently, one possible avenue of thought draws from the lack of convergence between explicit and implicit measures of SoA (Dewey & Knoblich, 2014). In turn, one may infer that such measures indeed measure different aspects of SoA. The judgement of agency, and the feeling of agency, are empirically distinct; thus, the judgement of causal relationship, and the feeling that two events are causally related, might also be distinct, resulting in the lack of convergence between explicit and implicit measures, of which temporal binding fails to measure. Recent evidence employing transcranial direct current stimulation (tDCS) also showed a lack of convergence between implicit and explicit agency: Hughes (2018) asked participants to press the 'up'/'down' key to produce a loud/quiet tone, yet only half of experimental trials contained congruent action-effects to keypresses. A Libet clock was present throughout the trial, and participants were asked to estimate the onset times of their keypresses and the following tones in separate trials. Post-trial, participants then respond to seven-point scale that asked to what extent the participant felt their action caused the tone. Hughes also applied tDCS to participants' right temporoparietal junction (TPJ) in Experiment 1 (left TPJ, Experiment 2). The results showed that congruency of action to learned outcomes only impacted explicit ratings of agency; binding was unaffected by congruency. Additionally, and despite prior research (Khalighinejad & Haggard, 2015), temporal binding was unaffected by TPJ stimulation, whereas explicit ratings were reduced with tDCS applied to the right TPJ. Hughes (2018) highlights not only the mismatch between implicit and explicit agency, but also that congruency fails to modulate temporal binding despite the clear causal link between action and the outcome (up for louder sound, down for quieter sound). Equally, within the context of the review of temporal binding research by Hughes et al., (2013), no studies manipulating action identity prediction demonstrate binding modulation.

Throughout our experiments both congruency and valence failed to modulate temporal binding. Therefore, if binding is not modulated by socially relevant factors (Chapter 2), consistency of morally valenced action-outcomes (Chapter 3), and similar binding occurs between outcomes resulting from intentional actions and a time-sequence condition (Chapter 4), simpler attentional/temporal control mechanisms cannot be discounted as the cause of binding. For example, in their review, Hughes et al. (2013) demonstrate how previous studies have failed to isolate motor-identity prediction from temporal control. In conjunction, Desantis et al. (2012) also show that temporal control is sufficient to produce binding regardless of motor-identity prediction. Moreover, as this thesis has shown binding is not modulated by what we know to influence explicit agency (e.g., Oishi et al., 2018), we provide evidence that temporal binding is less a measure of implicit agency attribution or causal inference and may instead be a process evoked from attentional/temporal control mechanisms. Further research should seek to clarify the relation between temporal control, temporal attention, motor-identity prediction and causal beliefs to isolate their modulatory influence upon temporal binding.

Naturally, one possible conclusion that our findings from chapters 2 – 4, as well as the inconsistencies between explicit and implicit measures of agency/causality, indicate, is that temporal binding does not actually measure causality or agency. The assertion that binding measures SoA/causality is frequently stated (Buehner, 2012; Ebert & Wegner, 2010; Engbert, Wohlschläger & Haggard, 2008; Moore & Fletcher, 2012; Moore, Wegner & Haggard, 2009; Obhi & Hall, 2011a), but this has always been a theoretical assumption on the grounds of Hume's notion of temporal contiguity. One perspective to view this issue is from the ideomotor effect, the well-established notion that learned outcomes are mentally represented the same way as the actions that produce them (Elsner and Hommel, 2001; Greenwald, 1970). Strong evidence suggests we do internally activate specific effects that we predict on the basis of our selected actions (Kunde, 2003; Pfister, Kiesel, & Melcher, 2010; Waszak & Herwig, 2007). However, the evidence that binding arises from such effect-prediction, as many propose (de Vignemont & Fourneret, 2004; Evans, 1982; Farrer, Franck, Georgieff, Frith, Decety et al, 2003; Farrer et al., 2003; Marcel, 2003; Mechsner al., 2001; Saito et al., 2005), or even from general causal mechanisms (Buehner, 2012) becomes questionable given the lack replicability of findings, the lack of convergence between explicit and implicit measures of agency (Dewey and Knoblich, 2014; Hughes, 2018), and the lack of testing for low-level explanations of binding (Hughes et al., 2013).

The existence of temporal binding is unique and interesting, yet we are still in midst of fully understanding why we feel time is shorter between two related events. Although the research contained within my thesis cannot make any specific statements, it does cast doubt on existing explanations. Therefore, to the primary aim of my research, that of how the perceptual error phenomena observed within SoA research translate to more social contexts, our current understanding of the answer is: not very well. This is problematic given that SoA is defined as the relation between one's actions and our effects upon the world around us, not merely the interaction between our perception of relatively simple actions and the ensuing stimulus feedback within laboratory settings. Naturally, however, the research presented here would further need to be replicated in laboratory settings to fully extrapolate these assertions.

That being said, there are a few limitations to my research that should be recognised: Chapter 3, although a novel methodology in combining temporal binding as a measure of implicit causality and IJR, our procedure in Experiment 5 may simply not have been sufficient to produce a temporal binding effect. Such an occurrence could be explained by the explicit causal judgements made in typical IJR research, whereby the second of a pair of events often begins with a derivative of "later that day". In other words, although participants may believe that a negative outcome is caused by prior immoral behaviour, this does not necessarily reflect at relatively shorter time intervals.

Additionally, the research contained within Experiment 7 is unable to make any particular claims as to the effect of causal beliefs on temporal binding due to the lack of belief in the existence of a confederate. Although this is noted and subsequently my discussion does not utilise the findings from this experiment, future research would be prudent to carry out more careful procedures in order to ensure adequate deception.

Finally, the research contained with Chapter 5 requires extrapolation in order to carefully test the relationship between spatial binding and causal beliefs. As stated previously, although we found that participants placed morally congruent events close together, this does not theoretically test causality given the conscious decision-making aspect. Our results could be explained by conceptual similarity between the two events (Casasanto, 2008; Lakoff & Johnson, 1999). For example, drawing on work on mental metaphors (Lakoff & Johnson, 1980), Casasanto argued that people use spatial proximity as a cue for judging similarity. Thus, in our experiments people might have positioned the morally congruent (vs. incongruent) events closer together not because they were inferring causality from deservingness but simply because the events were conceptually similar (e.g., poisoning one's dog and contracting a serious illness are conceptually similar because they are both bad), and thus further research is required to distinguish between causal and implicit association factors.

One such method to explore the underpinning mechanism would be to specifically link judgements of physical closeness to beliefs of deservingness, rather than only assume this, given that if people generally perceive others positively, then they should perceive others' good outcomes as more deserving than their bad outcomes (Harvey et al., 2014). As an example, ratings could be obtained regarding the outcomes of each scenario as to its deservingness and similarity during the spatial positioning task. A linear mixed effects model could then be used to investigate the predictive power of both factors upon spatial distance.

Overall, future studies should follow the recommended steps set out in Hughes et al. (2013) in order to isolate the various components of temporal binding in order ascertain the accuracy of assumptions that binding is a measure of SoA, or causality in general. In temporal binding research, this is paramount in order to develop a model that correctly identifies the key contributions of the low-level processes (temporal prediction and control, and motor non-/identity prediction), as well as high-level processes such as intention and knowledge of causal relationships. Isolating these specific processes will enable a sorely needed model that could provide information as to the relative predictive power of each process. However, such a model may equally highlight the lack of influence that high-level processes have relative to low-level processes, as is the case in Chapter 2. Furthermore, the axiom that temporal binding is an implicit measure of SoA has been challenged by the research within this thesis, and, as has been stated, replication of key studies, specifically with the isolation of individual processes laid out within Hughes et al. (2013), is equally sorely needed if genuine insight into temporal binding is wanted.

Conclusion of Thesis

The temporal contraction between two events, otherwise known as temporal binding, has been shown to be elusive as to its underlying mechanism/s. This thesis casts further doubts as to the extent temporal binding, and its purported measurement of the sense of agency, is modulated by high-level processes such as emotional valence, the self-serving bias, intentional action, and known biases in explicit causal reasoning. Equally questioned is whether the findings from the sense of agency literature can be applied to real-world contexts.

Understanding causal relations between our actions and their effects upon the environment is indeed vital for the human species, but, as I have shown, measuring such a notion through temporal binding, as it stands, is highly questionable, and other implicit measures should be pursued. We have shown is that we can feel physically unconnected events to be closer together in space, specifically with an interesting interaction between concerns for justice and spatial contiguity, the inference of a causal connection based on spatial proximity. Therefore, spatial binding may offer a more suitable case in measuring our implicit causal belief, but that is research route yet to be fully explored and awaits the resources of another eager PhD student.

References

References

Aarts, H., Bijleveld, E., Custers, R., Dogge, M., Deelder, M., Schutter, D., & van Haren, N. E. (2012). Positive priming and intentional binding: Eye-blink rate predicts reward information effects on the sense of agency. *Social neuroscience*, *7*, 105-112.

Aarts, H., Custers, R., & Wegner, D. M. (2005). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and cognition*, *14*, 439-458.

Aarts, H., & van den Bos, K. (2011). On the foundations of beliefs in free will: Intentional binding and unconscious priming in self-agency. *Psychological Science*, *22*, 532-537.

Amundson, J. C., & Miller, R. R. (2007). Similarity in spatial origin of information facilitates cue competition and interference. *Learning and motivation*, *38*, 155-171.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Bäß, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*, 137-143.

Bal, M., & van den Bos, K. (2012). Blaming for a better future: Future orientation and associated intolerance of personal uncertainty lead to harsher reactions toward innocent victims. *Personality and Social Psychology Bulletin*, *38*, 835–844.

Balslev, D., Nielsen, F. Å., Lund, T. E., Law, I., & Paulson, O. B. (2006). Similar brain networks for detecting visuo-motor and visuo-proprioceptive synchrony. *Neuroimage*, *31*, 308-312.

Banks, W. P., & Pockett, S. (2006). Libet's work on the neuroscience of free will. In M. Velmans (Ed.), *The Blackwell Companion to Consciousness* (pp. 657-670). Cambridge, England: Blackwell.

Barlas, Z., Hockley, W. E., & Obhi, S. S. (2017). The effects of freedom of choice in action selection on perceived mental effort and the sense of agency. *Acta psychologica*, *180*, 122-129.

Baumard, N., & Chevallier, C. (2012). What goes around comes around: The evolutionary roots of the belief in immanent justice. *Journal of Cognition and Culture*, *12*, 67–80.

Baumeister, R. F. (2010). The self. *Advanced social psychology: The state of the science*, 139-175.

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

Beasley, N. E. (1968). The extent of individual differences in the perception of causality. Canadian Journal of Psychology, 22, 399-407.

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*, 551-559.

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

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*, 237–242.

Blakey, E., Tecwyn, E. C., McCormack, T., Lagnado, D. A., Hoerl, C., Lorimer, S., & Buehner, M. J. (2019). When causality shapes the experience of time: Evidence for temporal binding in young children. *Developmental science*, *22*, e12769.

Bodner, K. E., Engelhardt, C. R., Minshew, N. J., & Williams, D. L. (2015). Making inferences: Comprehension of physical causality, intentionality, and emotions in discourse by high-functioning older children, adolescents, and adults with autism. *Journal of Autism and Developmental Disorders*, *45*, 2721–2733.

Boyd, R. (1981). “Scientific Realism and Naturalistic Epistemology,” in P. Asquith and R. Giere (Eds.), *PSA 1980*, (vol. II), East Lansing MI: Philosophy of Science Association, pp. 613–662.

Boyd, R. (1984). “The Current Status of Scientific Realism,” in J. Leplin (ed.), *Scientific Realism*, Berkeley CA: University of California Press, pp. 41–82.

Bradley, G. W. (1978). Self-serving biases in the attribution process: A reexamination of the fact or fiction question. *Journal of personality and social psychology*, *36*, 56.

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication?. *Journal of Experimental Social Psychology*, *50*, 217-224.

Buehner, M. J. (2012). Understanding the past, predicting the future causation, not intentional action, is the root of temporal binding. *Psychological science*, *23*, 1490-1497.

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: a test of the assumption of causal power. *Journal of experimental psychology: learning, memory, and cognition*, *29*, 1119-1140.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological science*, *20*, 1221-1228.

Buehner, M. J., & Humphreys, G. R. (2010). Causal contraction: spatial binding in the perception of collision events. *Psychological Science*, *21*, 44-48.

Burt, D. L., & DeMello, L. R. (2003). Attribution of rape blame as a function of victim gender and sexuality, and perceived similarity to the victim. *Journal of Homosexuality*, *43*, 39-57.

Cahill, C., Silbersweig, D., & Frith, C. (1996). Psychotic experiences induced in deluded patients using distorted auditory feedback. *Cognitive Neuropsychiatry*, *1*, 201–211

Callan, M. J., & Ellard, J. H. (2010). Beyond derogation and blame of victims: Just world dynamics in everyday life. In D. R. Bobocel, A. C. Kay, M. P. Zanna, & J. M. Olson (Eds.), *The psychology of justice and legitimacy: The Ontario symposium: Vol. 11*, (pp. 53–77). New York: Psychology Press.

Callan, M. J., Ellard, J. H., & Nicol, J. E. (2006). The belief in a just world and immanent justice reasoning in adults. *Personality and Social Psychology Bulletin*, *32*, 1646-1658.

Callan, M. J., Ferguson, H. J., & Bindemann, M. (2013). Eye movements to audiovisual scenes reveal expectations of a just world. *Journal of Experimental Psychology: General*, *142*, 34–40.

Callan, M. J., Harvey, A. J., Dawtry, R. J., & Sutton, R. M. (2013). Through the looking glass: Focusing on long-term goals increases immanent justice reasoning. *British Journal of Social Psychology*, *52*, 377-385.

Callan, M. J., Kay, A. C., Davidenko, N., & Ellard, J. H. (2009). The effects of justice motivation on memory for self- and other-relevant events. *Journal of Experimental Social Psychology*, *45*, 614–623.

Callan, M. J., Kay, A. C., & Dawtry, R. J. (2014). Making sense of misfortune: Deservingness, self-esteem, and patterns of self-defeat. *Journal of Personality and Social Psychology*, *107*, 142.

Callan, M. J., Moreton, J., & Hughes, G. (2018). *Bad People Feel Close to Bad Outcomes: Immanent Justice Reasoning by Spatial Proximity*. Manuscript submitted for publication.

Callan, M. J., Shead, N. W., & Olson, J. M. (2009). Foregoing the labor for the fruits: The effect of just world threat on the desire for immediate monetary rewards. *Journal of Experimental Social Psychology*, *45*, 246–249.

Callan, M. J., Shead, N. W., & Olson, J. M. (2011). Personal relative deprivation, delay discounting, and gambling. *Journal of Personality and Social Psychology*, *101*, 955–973.

Callan, M. J., Sutton, R. M., & Dovale, C. (2010). When deserving translates into causing: The effect of cognitive load on immanent justice reasoning. *Journal of Experimental Social Psychology*, *46*, 1097-1100.

Callan, M. J., Sutton, R. M., Harvey, A. J., & Dawtry, R. J. (2014). Immanent justice reasoning: Theory, research, and current directions. In J. M. Olson, & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology*, Vol. 49. (pp. 105–161). London: Academic Press.

Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of general Psychology*, *3*, 23.

Cancer Research UK, Lung Cancer Risk. (2018, December 27). Retrieved from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/risk-factors#heading-Zero.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Casasanto, D. (2008). Similarity and proximity: When does close in space mean close in mind? *Memory & Cognition*, *36*, 1047-1056.

Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion changes the sense of agency in the human brain. *Current biology*, *26*, 585-592.

Caspar, E. A., Cleeremans, A., & Haggard, P. (2015). The relationship between human agency and embodiment. Consciousness and cognition, 33, 226-236.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367– 405.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, *58*, 545–567.

Cheng, P. W., & Novick, L. R. (2005). Constraints and nonconstraints in causal learning: Reply to White (2005) and to Luhmann and Ahn (2005). *Psychological Review*, *112*, 694 –707.

Christensen, J. F., Yoshie, M., Di Costa, S., & Haggard, P. (2016). Emotional valence, sense of agency and responsibility: A study using intentional binding. *Consciousness and cognition*, *43*, 1-10.

Churches, O., Nicholls, M., Thiessen, M., Kohler, M., & Keage, H. (2014). Emoticons in mind: An event-related potential study. *Social neuroscience*, *9*, 196-202.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, New Jersey: Lawrence Erlbaum.

Copi, I. M., Cohen, C., & McMahon, K. (2016). *Introduction to logic*. Routledge.

Correia, I., Salvado, S., & Alves, H. V. (2016). Belief in a just world and self-efficacy to promote justice in the world predict helping attitudes, but only among volunteers. *The Spanish Journal of Psychology*, *19*, E28.

Coulson, C. A. (1955). *Science and Christian belief*. Oxford: Oxford University Press.

Cravo, A. M., Claessens, P. M., & Baldo, M. V. (2009). Voluntary action and causality in temporal binding. *Experimental brain research*, *199*, 95-99.

Dalbert, C., Lipkus, I. M., Sallay, H., & Goch, I. (2001). A just and an unjust world: Structure and validity of different world beliefs. *Personality and Individual Differences*, *30*, 561-577.

Dalbert, C., & Stoeber, J. (2006). The personal belief in a just world and domain-specific beliefs about justice at school and in the family: A longitudinal study with adolescents. *International Journal of Behavioral Development*, *30*, 200-207.

Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior research methods*, *40*, 428-434.

David, N., Newen, A., & Vogeley, K. (2008). The ‘‘sense of agency’’ and its underlying cognitive and neural mechanisms. *Consciousness and Cognition*, *17*, 523–534.

Davidov, J., & Eisikovits, Z. (2015). Free will in total institutions: The case of choice inside Nazi death camps. *Consciousness and cognition*, *34*, 87-97.

De Vignemont, F., & Fourneret, P. (2004). The sense of agency: A philosophical and empirical review of the "Who" system. *Consciousness and Cognition*, *13*, 1-19.

Daprati, E., & Sirigu, A. (2002). Laterality effects on motor awareness. *Neuropsychologia*, *40*, 1379-1386.

Derks, D., Bos, A. E., & Von Grumbkow, J. (2008). Emoticons in computer-mediated communication: Social motives and social context. *CyberPsychology & Behavior*, *11*, 99-101.

Desantis, A., Hughes, G., & Waszak, F. (2012). Intentional binding is driven by the mere presence of an action and not by motor prediction. *PLoS one*, *7*, e29557.

Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition*, *20*, 1211-1220.

Desantis, A., Weiss, C., Schütz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: authorship belief modulates sensory attenuation. *PLoS one*, *7*, e37959.

Dewey, J. A., & Carr, T. H. (2013). When dyads act in parallel, a sense of agency for the auditory consequences depends on the order of the actions. *Consciousness and Cognition*, *22*, 155–166.

Dewey, J. A., & Knoblich, G. (2014). Do implicit and explicit measures of the sense of agency measure the same thing?. *PloS one*, *9*, e110118.

Dik, G., & Aarts, H. (2007). Behavioral cues to others' motivation and goal pursuits: The perception of effort facilitates goal inference and contagion. *Journal of Experimental Social Psychology*, *43*, 727-737.

Dogge, M., Schaap, M., Custers, R., Wegner, D. M., and Aarts, H. (2012). When moving without volition: implied self-causation enhances binding strength between involuntary actions and effects. *Consciousness and Cognition*, *21*, 501–506.

Douven, I. (2017). "What Is Inference to the Best Explanation? And Why Should We Care?" in T. Poston and K. McCain (Eds.), *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press.

Draine, S. 1998. *Inquisit [Computer softwarer]*. Seattle, WA: Millisecond Software.

Duval, T. S., & Silvia, P. J. (2002). Self-awareness, probability of improvement, and the self-serving bias. *Journal of personality and social psychology*, *82*, 49.

Duval, T. S., Silvia, P. J., & Lalwani, N. (2012). *Self-awareness & causal attribution: A dual systems theory*. Springer Science & Business Media.

Duval, S., & Wicklund, R. A. (1973). Effects of objective self-awareness on attribution of causality. *Journal of Experimental Social Psychology*, *9*, 17-31.

Dzuka, J., & Dalbert, C. (2006). The belief in a just world and subjective well-being in old age. *Aging and Mental Health*, *10*, 439-444.

Eagleman, D. M., & Holcombe, A. O. (2002). Causality and the perception of time. *Trends in cognitive sciences*, *6*, 323-325.

Eagleman, D. M., Peter, U. T., Buonomano, D., Janssen, P., Nobre, A. C., & Holcombe, A. O. (2005). Time and the brain: how subjective time relates to neural time. *Journal of Neuroscience*, *25*, 10369-10371.

Ebert, J. P., & Wegner, D. M. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness and cognition*, *19*, 481-489.

Eder, A. B., Müsseler, J., & Hommel, B. (2012). The structure of affective action representations: temporal binding of affective response codes. *Psychological Research*, *76*, 111-118.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*, 3.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of experimental psychology: human perception and performance*, *27*, 229.

Engbert, K., & Wohlschläger, A. (2007). Intentions and expectations in temporal binding. *Consciousness and Cognition*, *16*, 255-264.

Engbert, K., Wohlschläger, A., Thomas, R., & Haggard, P. (2007). Agency, subjective time, and other minds. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1261.

Engbert, K., Wohlschläger, A., & Haggard, P. (2008). Who is causing what? The sense of agency is relational and efferent-triggered. *Cognition*, *107*, 693-704.

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

Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.

Evans, J. S. B., Newstead, S. E., & Byrne, R. M. (1993). *Human reasoning: The psychology of deduction*. Psychology Press.

Falcon, Andrea, "Aristotle on Causality", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/aristotle-causality/>.

Faro, D., Leclerc, F., & Hastie, R. (2005). Perceived causality as a cue to temporal distance. *Psychological Science*, *16*, 673-677.

Farrer, C., Franck, N., Frith, C. D., Decety, J., Georgieff, N., d'Amato, T., & Jeannerod, M. (2004). Neural correlates of action attribution in schizophrenia. *Psychiatry Research: Neuroimaging*, *131*, 31-44.

Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage*, *18*, 324-333.

Farrer, C., Franck, N., Paillard, J., & Jeannerod, M. (2003). The role of proprioception in action recognition. *Consciousness and cognition*, *12*, 609-619.

Farrer, C., Valentin, G., & Hupé, J. M. (2013). The time windows of the sense of agency. *Consciousness and cognition*, *22*, 1431-1441.

Fein, D., & Stein, G. M. (1977). Immanent punishment and reward in six-and nine-year-old children. *The Journal of Genetic Psychology*, *131*, 91–96.

Feldman, G. (2017). Making sense of agency: Belief in free will as a unique and important construct. *Social and Personality Psychology Compass*, *11*, e12293.

Fereday, R., & Buehner, M. J. (2015). Temporal binding and internal clocks: Is clock slowing general or specific? In Proceedings of the 37th Annual Meeting of the Cognitive Science Society (pp. 686–691).

Fereday, R., & Buehner, M. J. (2017). Temporal binding and internal clocks: No evidence for general pacemaker slowing. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 971.

Fourneret, P., Paillard, J., Lamarre, Y., Cole, J., & Jeannerod, M. (2002). Lack of conscious recognition of one's own actions in a haptically deafferented patient. *Neuroreport*, *13*, 541-547.

Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Daléry, J., d'Amato, T., & Jeannerod, M. (2001). Defective recognition of one's own actions in patients with schizophrenia. *American Journal of Psychiatry*, *158*, 454-459.

Frith, C. (2005). The self in action: lessons from delusions of control. *Consciousness and cognition*, *14*, 752-770.

Frith, C. D., Blakemore, S., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. *Brain Research Brain Research Review*, *31*, 357–363.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Science*, *4*, 14-21.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us?: Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, *2*, 181-193.

Gentsch, A., & Synofzik, M. (2014). Affective coding: the emotional dimension of agency. *Frontiers in human neuroscience*, *8*, 608.

Gentsch, A., Weiss, C., Spengler, S., Synofzik, M., & Schütz-Bosbach, S. (2015). Doing good or bad: How interactions between action and emotion expectations shape the sense of agency. *Social neuroscience*, *10*, 418-430.

Georgieff, N., & Jeannerod, M. (1998). Beyond consciousness of external reality: a "who" system for consciousness of action and self-consciousness. *Consciousness and cognition*, *7*, 465-477.

Gledhill, R. (2009, May 22). Archbishop of Westminster attacks atheism but says nothing on child abuse. *The Times*. Retrieved from https://web.archive.org/web/20101115042653/http://www.timesonline.co.uk/tol/comment/faith/article6334837.ece

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental psychology*, *37*, 620-629.

Gopnik, A., Schulz, L., & Schulz, L. E. (Eds.). (2007). *Causal learning: Psychology, philosophy, and computation*. Oxford University Press.

Goodman, N. D., Ullman, T. D., & Tenenbaum, J. B. (2011). Learning a theory of causality. *Psychological review*, *118*, 110.

Gould, S. J., Cox, A. L., Brumby, D. P., & Wiseman, S. (2015). Home is where the lab is: a comparison of online and lab data from a time-sensitive study of interruption. *Human Computation*, *2*, 45-67.

Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and control*, *2*, 329-352.

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

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 299-323). Cambridge, MA, US: The MIT Press.

Goswami, U., & Brown, A. L. (1990). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, *35*, 69-95.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological review*, *116*, 661.

Haarmeier, T., Bunjes, F., Lindner, A., Berret, E., & Thier, P. (2001). Optimizing visual motion perception during eye movements. *Neuron*, *32*, 527-535.

Haering, C., & Kiesel, A. (2012). Mine is earlier than yours: Causal beliefs influence the perceived time of action effects. *Frontiers in psychology*, *3*, 393.

Hafer, C. L. (2000a). Investment in long-term goals and commitment to just means drive the need to believe in a just world. *Personality and Social Psychology Bulletin*, *26*, 1059–1073.

Hafer, C. L., & Bègue, L. (2005). Experimental research on just world theory: Problems, developments, and future challenges. *Psychological Bulletin*, *131*, 128–166.

Hafer, C. L., Bègue, L., Choma, B. L., & Dempsey, J. L. (2005). Belief in a just world and commitment to long-term deserved outcomes. *Social Justice Research*, *18*, 429–444.

Haggard, P., Aschersleben, G., Gehrke, J., & Prinz, W. (2002). Action, binding, and awareness. In: W. Prinz, & B. Hommel (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance.* (pp. 266-285). Oxford University Press.

Haggard, P., & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and cognition*, *12*, 695-707.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature neuroscience*, *5*, 382-385.

Haggard, P., & Tsakiris, M. (2009). The experience of agency: Feelings, judgments, and responsibility. *Current Directions in Psychological Science*, *18*, 242-246.

Haidt, J., & Joseph, C. (2007). The moral mind: How 5 sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind* (pp. 367–391). New York, NY: Oxford University Press.

Hallett, M. (2007). Volitional control of movement: the physiology of free will. *Clinical Neurophysiology*, *118*, 1179-1192.

Hare, E. H. (1962). Masturbatory insanity: The history of an idea. *Journal of Mental Science*, *108*, 1-25.

Harman, G. H. (1965). The inference to the best explanation. *The philosophical review*, *74*, 88-95.

Harvey, A. J., & Callan, M. J. (2014). The role of religiosity in ultimate and immanent justice reasoning. *Personality and Individual Differences*, *56*, 193-196.

Hassin, R. R., Aarts, H., & Ferguson, M. (2005). Automatic goal inferences. *Journal of Experimental Social Psychology*, *41*, 129–140.

Hauser, M., Knoblich, G., Repp, B. H., Lautenschlager, M., Gallinat, J., Heinz, A., & Voss, M. (2011). Altered sense of agency in schizophrenia and the putative psychotic prodrome. *Psychiatry research*, *186*(2-3), 170-176.anent justice reasoning. *Personality and Individual Differences*, *56*, 193–196.

Hillis J.M., Ernst M.O., Banks M.S., & Landy M.S. (2002) Combining sensory information: mandatory fusion within, but not between, senses. *Science*, *298*, 1627–1630.

Holland, P. W., Glymour, C., & Granger, C. (1985). Statistics and causal inference. *ETS Research Report Series*, *1985*, i-72.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal cognition*, *8*, 164-181.

Hudson, M. B., Nicolas, S. C., Howser, M. E., Lipsett, K. E., Robinson, I. W., Pope, L. J., Hobby, A. F., & Friedman, D. R. (2015). Examining how gender and emoticons influence Facebook jealousy. *Cyberpsychology, behavior, and social networking*, *18*, 87-92.

Hughes, G. (2015). ERP and behavioral evidence of increased sensory attenuation for fear-related action outcomes. *Biological psychology*, *111*, 8-13.

Hughes, G. (2018). The role of the temporoparietal junction in implicit and explicit sense of agency. *Neuropsychologia*, *113*, 1-5.

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological bulletin*, *139*, 133-151.

Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *Neuroimage*, *56*, 1632-1640.

Hume, D. (1739/1978). *A treatise of human natu*re. Oxford, UK: Oxford University Press.

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals temporal binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1542.

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental brain research*, *203*, 465-470.

Ito, T. A., Cacioppo, J. T., & Lang, P. J. (1998). Eliciting affect using the International Affective Picture System: Trajectories through evaluative space. *Personality and Social Psychology Bulletin*, *24*, 855-879.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*, 309.

Jolij, J., & Lamme, V. A. (2005). Repression of unconscious information by conscious processing: evidence from affective blindsight induced by transcranial magnetic stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 10747-10751.

James, W. (1890). *Principles of psychology*. New York: Holt.

Jeannerod, M., & Pacherie, E. (2004). Agency, simulation and self-identification. *Mind & language*, *19*, 113-146.

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, *79*, 1-17.

Josephson, J. R., & Josephson, S. G. (1994). *Abductive inference*. Cambridge: Cambridge University Press.

Joynson, R. B. (1971). Michotte's experimental methods. *British Journal of Psychology*, *62*, 293-302.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54-69.

Kalckert, A., & Ehrsson, H. H. (2014). The moving rubber hand illusion revisited: comparing movements and visuotactile stimulation to induce illusory ownership. *Consciousness and cognition*, *26*, 117-132.

Kawabe, T., Roseboom, W., & Nishida, S. (2013). The sense of agency is action-effect causality perception based on cross-modal grouping. *Proceedings of the Royal Society B: Biological Sciences*, *280*, 20130991–20130991.

Khalighinejad, N., & Haggard, P. (2015). Modulating human sense of agency with non-invasive brain stimulation. *Cortex*, *69*, 93-103.

Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the Stroop task. *Cognition*, *117*, 341-347.

Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., & Voss, M. (2011). Whodunnit? Electrophysiological correlates of agency judgements. *PLoS One*, *6*, e28657.

Kumar, D., & Srinivasan, N. (2017). Multi-scale control influences sense of agency: Investigating intentional binding using event-control approach. *Consciousness and Cognition*, *49*, 1-14.

Kunde, W. (2003). Temporal response-effect compatibility. *Psychological research*, *67*, 153-159.

Ladyman, J., Douven, I., Horsten, L., & Van Fraassen, B. (1997). A defence of van Fraassen's critique of abductive inference: Reply to Psillos. *The Philosophical Quarterly*, *47*, 305-321.

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, *4*, 195-208.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology*, *University of Florida*.

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, *69*, 127–137.

Legare, C. H., Evans, E. M., Rosengren, K. S., & Harris, P. L. (2012). The coexistence of natural and supernatural explanations across cultures and development. *Child Development*, *83*, 779–793.

Leppänen, J. M., Kauppinen, P., Peltola, M. J., & Hietanen, J. K. (2007). Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. *Brain research*, *1166*, 103-109.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum.

Lerner, M. J. (2002). Pursuing the justice motive. In M. Ross & Dale T. Miller (Eds.*), The justice motive in everyday life* (pp. 109–126). New York: Cambridge University Press.

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, *85*, 1030–1051.

Lerner, M. J., Miller, D. T., & Holmes, J. G. (1976). Deserving and the emergence of forms of justice. In L. Berkowitz & E. Walster (Eds.), *Advances in experimental social psychology, Vol. 9*, (pp. 133–162). New York: Academic Press.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain*, *106*, 623-642.

Linden, M., & Maercker, A. (Eds.). (2011). *Embitterment: Societal, psychological, and clinical perspectives*. Springer Science & Business Media.

Lindner, A., Haarmeier, T., Erb, M., Grodd, W., & Thier, P. (2006). Cerebrocerebellar circuits for the perceptual cancellation of eyemovement-induced retinal image motion. *Journal of Cognitive Neuroscience*, *18*, 1899–1912.

Lipton, P. (1992). Causation outside the law. *Jurisprudence: Cambridge Essays*, 127-148.

Lipton, P. (2009). Understanding without explanation. *Scientific understanding: Philosophical perspectives*, 43-63.

Loo, R. (2002). Belief in a just world: Support for independent just world and unjust world dimensions. *Personality and Individual Differences*, *33*, 703-711.

Lush, P., Roseboom, W., Cleeremans, A., Scott, R., Seth, A. K., & Dienes, Z. (2018). Intentional binding as Bayesian cue combination: testing predictions with trait individual differences. *Journal of Experimental Psychology: Human Perception and Performance*.

Lyons-Warren, A., Rema, L., & Hershey, T. (2004). Short- and long-term spatial delayed response performance across the lifespan. *Developmental Neuropsychology*, *26*, 661–678.

Maes, J. (1998). Immanent and ultimate justice: Two ways of believing in justice. In M. J. Lerner & L. Montada (Eds.), *Responses to victimizations and belief in a just world* (pp. 9–40). New York: Plenum.

Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological bulletin*, *132*, 895.

Marcel, A. (2003). The sense of agency: Awareness and ownership of action. *Agency and self-awareness*, 48-93.

McCurry, J. (2011, March 15). Tokyo governor apologises for calling tsunami 'divine punishment'. *The Guardian*. Retrieved from https://www.theguardian.com/world/2011/mar/15/tokyo-governor-tsunami-punishment

Mechsner, F. (2004). A psychological approach to human voluntary movements. *Journal of Motor Behavior*, *36*, 355-370.

Mechsner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature*, *414*, 69.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological bulletin*, *130*, 711-747.

Michotte, A. E. (1963). The perception of causality (Miles, T.R., Trans.). London: Methuen & Co. (Original work published 1946).

Metzinger, T. (2000), The subjectivity of subjective experience: A representationalist analysis of the first-person perspective, in T. Metzinger (Ed), *Neural Correlates of Consciousness. Empirical and Conceptual Questions*, Cambridge, MA: MIT Press.

Michotte, A. (1963). *The perception of causality* (T. R. Miles & Elaine Miles, Trans.). New York: Basic Books.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction?. *Psychological bulletin*, *82*, 213-225.

Monroe, A. E., Dillon, K. D., & Malle, B. F. (2014). Bringing free will down to Earth: People's psychological concept of free will and its role in moral judgment. *Consciousness and cognition*, *27*, 100-108.

Moretto, G., Walsh, E., & Haggard, P. (2011). Experience of agency and sense of responsibility. *Consciousness and cognition*, *20*, 1847-1854.

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Consciousness and cognition*, *21*, 59-68.

Moore, J. W., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and cognition*, *17*, 136-144.

Moore, J. W., & Haggard, P. (2010). Intentional binding and higher order agency experience. *Consciousness and cognition*, *19*, 490-491.

Moore, J. W., Middleton, D., Haggard, P., & Fletcher, P. C. (2012). Exploring implicit and explicit aspects of sense of agency. *Consciousness and cognition*, *21*, 1748-1753.

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and cognition*, *21*, 546-561.

Moore, J. W., Teufel, C., Subramaniam, N., Davis, G., & Fletcher, P. C. (2013). Attribution of intentional causation influences the perception of observed movements: behavioral evidence and neural correlates. *Frontiers in psychology*, *4*, 72-81.

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and cognition*, *18*, 1056-1064.

Moskowitz, G. B. (2005). Social cognition: Understanding self and others. Guilford Press.

Moreton, J., Callan, M. J., & Hughes, G. (2017). How much does emotional valence of action outcomes affect temporal binding?. *Consciousness and cognition*, *49*, 25-34.

Nadel, L., & Hardt, O. (2011). Update on memory systems and processes. *Neuropsychopharmacology*, *36*, 251.

Nahab, F. B., Kundu, P., Gallea, C., Kakareka, J., Pursley, R., Pohida, T., Miletta, N., Friedman, J., & Hallett, M. (2010). The neural processes underlying self-agency. *Cerebral Cortex*, *21*, 48-55.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European journal of social psychology*, *15*, 263-280.

Nemeroff, C., & Rozin, P. (1989). "You are what you eat": Applying the demand-free "impressions" technique to an unacknowledged belief. Ethos, 17, 50-69.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological bulletin*, *89*, 133.

Obhi, S. S., & Hall, P. (2011). Sense of agency and intentional binding in joint action. *Experimental brain research*, *211*, 655.

Ogawa, K., & Inui, T. (2007). Lateralization of the posterior parietal cortex for internal monitoring of self-versus externally generated movements. *Journal of Cognitive Neuroscience*, *19*, 1827-1835.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *Journal of personality and social psychology*, *80*, 381.

Oishi, H., Tanaka, K., & Watanabe, K. (2018). Feedback of action outcome retrospectively influences sense of agency in a continuous action task. *PloS one*, *13*, e0202690.

Pacherie, E. (2001). Agency lost and found: a commentary on Spence. *Philosophy, Psychiatry, & Psychology*, *8*, 173-176.

Pantelis, P. C., & Feldman, J. (2012). Exploring the mental space of autonomous intentional agents. *Attention, Perception, & Psychophysics*, *74*, 239-249.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes, Vol. 1. Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA, US: Academic Press.

Peirce, C. S. (1974). *Collected papers of charles sanders peirce* (Vol. 5). Harvard University Press.

Peirce, C. S. (1992). *The essential Peirce: selected philosophical writings* (Vol. 2). Indiana University Press.

Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review*, *14*, 577-596.

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a'low road'to'many roads' of evaluating biological significance. *Nature reviews neuroscience*, *11*, 773.

Percival, P., & Haviland, J. M. (1978). Consistency and retribution in children's immanent justice decisions. *Developmental Psychology*, *14*, 132–136.

Pfister, R., Obhi, S. S., Rieger, M., & Wenke, D. (2014). Action and perception in social contexts: intentional binding for social action effects. *Frontiers in human neuroscience*, *8*, 667.

Piaget, J. (1932/1965). The moral judgment of the child. London, England: Kegan, Paul, Trench, Trubner, & Co, Original work published 1932.

Pfister, R., Obhi, S. S., Rieger, M., & Wenke, D. (2015). Action and perception in social contexts: intentional binding for social action effects. *Frontiers in human neuroscience, 8,* 138-147.

Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition*, *16*, 241-254.

Poonian, S. K., & Cunnington, R. (2013). Intentional binding in self-made and observed actions. *Experimental brain research*, *229*, 419-427.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Assessing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185–227.

Prinz, W. (2003). Experimental approaches to action. In J. Roessler & N. Eilan (Eds.), *Agency and self-awareness* (pp. 165–187). Oxford: OUP.

Raftopoulos, A. (2016). Abduction, inference to the best explanation, and scientific practise: The case of newton's optics. In L. Magnani (Ed), *Model-Based Reasoning in Science and Technology* (pp. 259-277). Berlin: Springer-Verlag.

Raman, L., & Winer, G. A. (2002). Children's and adults' understanding of illness: Evidence in support of a coexistence model. *Genetic, Social, and General Psychology Monographs*, *128*, 325–355.

Raman, L., & Winer, G. A. (2004). Evidence of more immanent justice responding in adults than children: A challenge to traditional developmental theories. *British Journal of Developmental Psychology*, *22*, 255–274.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489*, 427–430.

Renes, R. A., & Aarts, H. (2017). The sense of agency in health and well-being. In D. de Ridder, M. Adriaanse, & K. Fujita (Eds.), *Routledge International Handbook of Self-Control in Health and Well-Being*, 193. New York: Routledge.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*, 151.

Ruess, M., Thomaschke, R., & Kiesel, A. (2017). The time course of intentional binding. *Attention, Perception, & Psychophysics*, *79*, 1123-1131.

Rigden, J. S. (2005). *Einstein 1905*. Harvard University Press.

Rigoni, D., & Brass, M. (2014). From intentions to neurons: Social and neural consequences of disbelieving in free will. *Topoi*, *33*, 5-12

Rigoni, D., Sammicheli, L., & Brass, M. (2011). Perspectives of the experience of will. *HumanaMente*, *15*, 139-158.

Riley, G. A., & Baah-Odoom, D. (2012). Belief in a just world, generalised self-efficacy and stigma may contribute to unsafe sexual intentions via a reduced perception of vulnerability to HIV/AIDS amongst young people in Ghana. *AIDS care*, *24*, 642-648.

Rips, L. J. (2011). Causation from perception. *Perspectives on Psychological Science*, *6*, 77-97.

Roethlisberger, F. J. & Dickson, W. J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.

Ross, L. (1977). The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process1. In *Advances in experimental social psychology* (Vol. 10, pp. 173-220). Academic Press.

Ryan, W. (1971). *Blaming the victim*. London: Orbach and Chambers.

Saito, N., Mushiake, H., Sakamoto, K., Itoyama, Y., & Tanji, J. (2005). Representation of immediate and final behavioral goals in the monkey prefrontal cortex during an instructed delay period. *Cerebral Cortex*, *15*, 1535-1546.

Saito, N., Takahata, K., Murai, T., & Takahashi, H. (2015). Discrepancy between explicit judgement of agency and implicit feeling of agency: implications for sense of agency and its disorders. *Consciousness and cognition*, *37*, 1-7.

Sato, A. (2009). Both motor prediction and conceptual congruency between preview and action-effect contribute to explicit judgment of agency. *Cognition*, *110*, 74-83.

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*, 241-255.

Sayen, J. (1985). *Einstein in America: The Scientist's Conscience in the Age of Hitler and Hiroshima*. New York, Crown Publishers, Inc.: New York.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in cognitive sciences*, *4*, 299-309.

Schwartz, B. (1989). *Psychology of learning and behavior*. WW Norton & Co.

Sedikides, C., & Strube, M. J. (1995). The multiply motivated self. *Personality and Social Psychology Bulletin*, *21*, 1330-1335.

Schlottmann, A., Ray, E. D., Mitchell, A., & Demetriou, N. (2006). Perceived physical and social causality in animated motions: Spontaneous reports and ratings. *Acta psychologica*, *123*, 112-143.

Shanks, D. R., Holyoak, K. J., & Medin, D. L. (Eds.). (1996). *The psychology of learning and motivation: advances in research and theory. Causal learning*. Academic Press.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgement of causality by human subjects. *The Quarterly Journal of Experimental Psychology*, *41*, 139-159.

Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review*, *3*, 314-321.

Smith, N. K., Cacioppo, J. T., Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia*, *41* 171-183.

Smith, E., Weinberg, A., Moran, T., & Hajcak, G. (2013). Electrocortical responses to NIMSTIM facial expressions of emotion. *International Journal of Psychophysiology*, *88*, 17-25.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*, 799.

Stetson, C., Cui, X., Montague, P. R., & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron*, *51*, 651-659.

Straube, B., & Chatterjee, A. (2010). Space and time in perceptual causality. *Frontiers in Human Neuroscience*, *4*, 28.

Sutton, R. M., Douglas, K. M., Wilkin, K., Elder, T. J., Cole, J. M., & Stathi, S. (2008). Justice for whom, exactly? Beliefs in justice for the self and various others. *Personality and Social Psychology Bulletin*, *38*, 528–541.

Sutton, R. M., & Winnard, E. J. (2007). Looking ahead through lenses of justice: The relevance of just-world beliefs to intentions and confidence in the future. *British Journal of Social Psychology*, *46*, 649–666.

Suzuki, K., Lush, P., Seth, A., & Roseboom, W. (2018, October 9). ‘Intentional binding’ without intentional action. https://doi.org/10.31234/osf.io/vaybe

Synofzik, M., Vosgerau, G., & Lindner, A. (2009). Me or not me–An optimal integration of agency cues?. *Consciousness and cognition*, *18*, 1065-1068.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and cognition*, *17*, 219-239.

Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's not my fault: Postdictive modulation of intentional binding by monetary gains and losses. *PloS one*, *7*, e53421.

Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*, 697.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological bulletin*, *103*, 193.

Taylor, S. E., & Brown, J. D. (1994). Positive illusions and well-being revisited: separating fact from fiction.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in cognitive sciences*, *10*, 309-318.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A, Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research*, *168*, 242-249.

Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. Cognition, 96, 215-231.

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of cognitive neuroscience*, *16*, 817-827.

Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neuroscience & Biobehavioral Reviews*, *36*, 943-959.

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends in cognitive sciences*, *7*, 65-69.

Wegner, D. M. (2004). Précis of the illusion of conscious will. *Behavioral and Brain Sciences*, *27*, 649-659.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American psychologist*, *54*, 480.

Wen, W., Yamashita, A., & Asama, H. (2015). The influence of action-outcome delay and arousal on sense of agency and the intentional binding effect. *Consciousness and cognition*, *36*, 87-95.

Wohlschläger, A., Engbert, K., & Haggard, P. (2003). Intentionality as a constituting condition for the own self—and other selves. *Consciousness and cognition*, *12*, 708-716.

Wohlschläger, A., Haggard, P., Gesierich, B., & Prinz, W. (2003). The perceived onset time of self-and other-generated actions. *Psychological Science*, *14*, 586-591.

Wolpe, N., Haggard, P., Siebner, H. R., & Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research*, *229*, 467-474.

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

Woods, A. J., Lehet, M., & Chatterjee, A. (2012). Context modulates the contribution of time and space in causal inference. *Frontiers in psychology*, *3*, 371.

Woolley, J. D., Cornelius, C. A., & Lacy, W. (2011). Developmental changes in the use of supernatural explanations of unusual events. *Journal of Cognition and Culture*, *11*, 311–337.

Wolfman, M., & Broderick, P. (1989). Batman Vol. 1, *Batman Year 3 - Changes Made*, *DC*, *437*.

Whorton, J. (2001). Looking back: The solitary vice: The superstition that masturbation could cause mental illness. *Western Journal of Medicine*, *175*, 66.

Xie, X., Liu, H., & Gan, Y. (2011). Belief in a just world when encountering the 5/12 Wenchuan earthquake. *Environment and Behavior*, *43*, 566-586.

Yoshie, M., & Haggard, P. (2013). Negative emotional outcomes attenuate sense of agency over voluntary actions. *Current biology*, *23*, 2028-2032.

Yoshie, M., & Haggard, P. (2017). Effects of emotional valence on sense of agency require a predictive model. *Scientific Reports*, *7*, 8733.

Zopf, R., Polito, V., & Moore, J. (2018). Revisiting the link between body and agency: visual movement congruency enhances intentional binding but is not body-specific. *Scientific reports*, *8*, 196.

Appendix

Appendix A

Scenarios used in Chapter 3:

Scenario 1:

While walking home after work, Simon offered to carry an elderly woman's bags and helped her cross the road.

While walking home after work, Simon yelled at an elderly woman to get out of his way and shoved her into the gutter as he passed by.

Scenario 2:

Jamie saw a homeless man drop his wallet. Jamie picked up the man's wallet and ran after him to give it back.

Jamie saw a homeless man with a bucket for change, and stole the homeless man's money.

Scenario 3:

Laura patiently allowed a man in a wheelchair to go ahead of her in the taxi rank.

Laura jumped the taxi queue while shoving a man in a wheelchair out of the way.

Scenario 4:

Carole saw a woman fall over, and ran immediately to check if she was okay.

Carole saw a woman fall over, and walked past while calling her a "fat cow".

Scenario 5:

While walking home, Maria saw a child looking lost, and helped them find their way back home.

While walking home, Maria saw a child looking lost, and ignored the child's request to help them find their way back home.

Scenario 8:

Emily helped a struggling elderly lady at the tube station with the ticket machine.

Emily pushed an elderly lady on the floor, fracturing the woman's hip, and called her "retarded".

Scenario 9:

Jack threw his water bottle in the public waste-bin, and decided to clean up the other rubbish on the ground, too.

Jack threw his water bottle at a squirrel that was eating a walnut on the bench next to him, killing the squirrel instantly.

Scenario 10:

Sonia spent more than twenty minutes patiently driving around the car park, even though the disabled parking space was available.

Sonia parked in a disabled parking space, forcing a genuinely disabled person was forced to park their car at the very end of the car park.

Scenario 11:

Harry told the woman that was flirting with him that he had a wife, and walked away.

Harry told the flirtatious woman to lead the way back to her place, even though Harry was married with children.

Scenario 12:

Carrie noticed that Giula's sandal was not fastened properly, and quickly knelt to fix it for her.

Carrie noticed that Giula's sandal was not fastened properly, but did purposely not tell her, leading to Giula breaking her ankle.

Scenario 13:

Miriam sacrificed her chance of acquiring the nursing job because she knew her friend Kate deserved it more.

Miriam told person in charge that Kate was a terrible nurse so that she could get the job herself.

Scenario 14:

Brian accidentally bumped into a parked car, and left a note offering to pay for the repair himself as an apology.

Brian was drunk-driving when he crashed into a car killing the other driver and the passenger.

Scenario 15:

Leonard made a donation to the local charity shop. He noticed the window was broken, so went home to collect his tools, and returned to fix the window.

Leonard was caught stealing from a charity shop. Although he got away, he returned to the shop, and threw a brick through the window.

Scenario 16:

Jordan risked his own life to save a drowning puppy in the river.

Jordan kicked a puppy into the river, where it drowned.

Scenario 17:

Samantha gave blood at the local hospital, although she doesn't like the sight of blood herself.

Samantha, a married mother of three, took a day off work yesterday to meet up with an attractive man at a seedy motel.

Scenario 18:

Richard went into his autistic son's room, took him in his arms, and promised him that he would always love him.

Richard went into his autistic son's room, slapped him in the face and left saying that he was never coming back.

Scenario 19:

Wendy told Janice that her house was absolutely lovely, and offered her some flowers in a very beautiful hand-made vase.

Wendy pretended to go to the restroom, but instead went into Janice's bedroom and stole one of her most valuable necklaces.

Scenario 20:

Frank paid for his items at the shop, and whilst no one was looking, put all of his spare change into a charity pot.

Frank paid for his items at the shop, and whilst no one was looking, stole all of the change from a charity pot.

Scenario 21:

David saw an elderly man fall over his walking cane down the street, and rushed over to help the elderly man to his feet.

David saw an elderly man fall over his walking cane down the street, laughed, and took a picture of the elderly man on the floor.

Scenario 22:

Although she was late for work, Rachel cycled slow on a school road full of children.

Rachel cycled at a fast and dangerous speed, leading her to crash into a child and knocking them onto the ground, breaking two of their teeth.

Scenario 23:

During the race, Ben ran back to pick up Peter, and they both crossed the finish line together.

During the race, Ben kicked Peter's ankle so that he could win the race and the gold medal.

Scenario 24:

Amanda stayed after class to help Monica go over the schoolwork she was finding difficult.

Amanda shouted to her students in the playground to point and laugh at Monica, calling her a "stupid little girl".

Scenario 25:

On a busy train, Mark gave up his seat for an elderly lady in crutches.

On a busy train, Mark scoffed at an elderly lady in crutches, refusing to give up his seat.

Scenario 26:

John bought some food for a homeless man.

John kicked and stole money from a homeless man.
Scenario 27:

Andrew volunteered to help care for sick children at the hospital.

Andrew, a married man with children, spent the night with a prostitute.

Scenario 28:

Marty returned home after spending 6 months of service in the army protecting civilians in Afghanistan.

Marty returned home after smuggling heroin from Afghanistan.

Scenario 29:

Nina defended a Turkish shop assistant from a racially abusive woman.

Nina joined in with another woman to racially abuse a Turkish shop assistant.

Scenario 30:

Jan donated a large sum of money to a children's cancer charity.

Jan embezzled a large sum of money from a children's cancer charity.

Scenario 31:

Lisa volunteered to spend the afternoon at a school teaching young children of the dangers of drugs and alcohol.

Lisa spent the afternoon near a school selling drugs to teenage students.

Scenario 32:

Aisha freely took the blame for her colleague who had broke a store item.

Aisha accidentally broke a store item and blamed her colleague.

Scenario 33:

Philip stopped by a stranger's car to help give them a jumpstart, even though it was very cold and raining.

Philip stopped by a stranger's car to help give them a jumpstart, but instead assaulted them, stole their belongings, and drove off.

Scenario 34:

Whilst shopping, Margaret noticed she had been significantly undercharged, and informed the cashier straight away.

Whilst shopping, Margaret noticed she had been significantly undercharged, but decided not to tell the cashier.

Scenario 35:

Alistair signed up to volunteer at a children's care home.

Alistair thought it was a complete waste of time to volunteer at a children's care home.

Scenario 36:

Dawn gave her coat to a shivering homeless woman.

Dawn stole the coat from a shivering homeless woman.

Appendix B

Scenarios used in Chapter 5, Experiments 1 and 2:

SET 1 [SET 2]

S = same person experiencing the outcome
D – different person experiencing the outcome
C = congruent outcome
I = incongruent outcome

SC1 [to DC3, Geoff-Mark]

Frank raped a co-worker at a company retreat
Frank was in a freak car accident

SC2 [to DC4, Mark-Tony]

John embezzled funds from a children's cancer care organization
John's ground-floor apartment was destroyed by flooding

DC1 [to SC3, Frank-Frank]

Geoff fled the scene after hitting a child with his car in a crosswalk
Paul lost his job of 20 years due to corporate downsizing

DC2 [to SC4, John-John]

Mark kicked a homeless woman in the face because she smiled at him
Tony contracted a serious illness

SI1 [to DI3, James-Charlie]

Michael robbed a convenience store at gunpoint
Michael won $100,000 from a scratch-and-win lottery ticket

SI2 [to DI4, Graham-Alexander]

Bruce poisoned his dog because he didn't want to take care of it any longer

Bruce received word that he won a luxury cruise trip from a sweepstakes he entered

DI1 [to SI3, Michael-Michael]

James has been cheating on his wife with his younger intern at work
Charlie's stocks and shares skyrocketed

DI2 [to SI4, Bruce-Bruce]

Graham's been selling drugs to school kids
Alexander won a year's worth of free air travel from a travel agency's promotion

<u>Scenarios used in Chapter 5, Experiment 3</u>:

Set 1	[sentence started with same name during the task]	
Frank raped a co-worker	was in a freak car accident	b
Joe robbed a store a gunpoint	apartment was destroyed by fire	b
Neil assaulted a child	contracted a serious illness	b
John drowned a puppy	slipped and fractured his leg	b
Ben punched an elderly woman	won the lottery	g
Tim kidnapped a child	received a major pay-rise at work	g
Jeff beat his wife	won a luxury cruise trip	g
Jim poisoned his mistress	received a massive inheritance	g

Set 2		
Frank raped a co-worker	received a massive inheritance	g
Joe robbed a store a gunpoint	was in a freak car accident	b
Neil assaulted a child	apartment was destroyed by fire	b
John drowned a puppy	contracted a serious illness	b
Ben punched an elderly woman	slipped and fractured his leg	b
Tim kidnapped a child	won the lottery	g
Jeff beat his wife	received a major pay-rise at work	g
Jim poisoned his mistress	won a luxury cruise trip	g

Set 3		
Frank raped a co-worker	won a luxury cruise trip	g
Joe robbed a store a gunpoint	received a massive inheritance	g
Neil assaulted a child	was in a freak car accident	b
John drowned a puppy	apartment was destroyed by fire	b
Ben punched an elderly woman	contracted a serious illness	b
Tim kidnapped a child	slipped and fractured his leg	b
Jeff beat his wife	won the lottery	g
Jim poisoned his mistress	received a major pay-rise at work	g

Set 4		
Frank raped a co-worker	received a major pay-rise at work	g
Joe robbed a store a gunpoint	won a luxury cruise trip	g

Neil assaulted a child	received a massive inheritance	g
John drowned a puppy	was in a freak car accident	b
Ben punched an elderly woman	apartment was destroyed by fire	b
Tim kidnapped a child	contracted a serious illness	b
Jeff beat his wife	slipped and fractured his leg	b
Jim poisoned his mistress	won the lottery	g

Set 5

Frank raped a co-worker	won the lottery	g
Joe robbed a store a gunpoint	received a major pay-rise at work	g
Neil assaulted a child	won a luxury cruise trip	g
John drowned a puppy	received a massive inheritance	g
Ben punched an elderly woman	was in a freak car accident	b
Tim kidnapped a child	apartment was destroyed by fire	b
Jeff beat his wife	contracted a serious illness	b
Jim poisoned his mistress	slipped and fractured his leg	b

Set 6

Frank raped a co-worker	slipped and fractured his leg	b
Joe robbed a store a gunpoint	won the lottery	g
Neil assaulted a child	received a major pay-rise at work	g
John drowned a puppy	won a luxury cruise trip	g
Ben punched an elderly woman	received a massive inheritance	g
Tim kidnapped a child	was in a freak car accident	b
Jeff beat his wife	apartment was destroyed by fire	b
Jim poisoned his mistress	contracted a serious illness	b

Set 7

Frank raped a co-worker	contracted a serious illness	b
Joe robbed a store a gunpoint	slipped and fractured his leg	b
Neil assaulted a child	won the lottery	g
John drowned a puppy	received a major pay-rise at work	g
Ben punched an elderly woman	won a luxury cruise trip	g
Tim kidnapped a child	received a massive inheritance	g
Jeff beat his wife	was in a freak car accident	b
Jim poisoned his mistress	apartment was destroyed by fire	b

Set 8

Frank raped a co-worker	apartment was destroyed by fire	b
Joe robbed a store a gunpoint	contracted a serious illness	b
Neil assaulted a child	slipped and fractured his leg	b
John drowned a puppy	won the lottery	g
Ben punched an elderly woman	received a major pay-rise at work	g
Tim kidnapped a child	won a luxury cruise trip	g
Jeff beat his wife	received a massive inheritance	g
Jim poisoned his mistress	was in a freak car accident	b