Temporal binding and internal clocks

Thesis submitted by:

Richard Fereday

Supervised by:

Professor Marc J. Buehner

Professor Simon K. Rushton

In partial fulfilment for the degree of:

PhD (Psychology)

September 2016

School of Psychology

Acknowledgments

I have long maintained that we are nothing without the people that support us. I also believe that my journey throughout this PhD would have been a very different and rather less pleasant experience, without the support of these exceptional individuals:

Marc – you are knowledgeable, patient, and affable, qualities that have shone like a beacon during the past four years. It is in this light that I have learnt, observed, and matured. Indeed, I have learnt far more than is included in this thesis, far more than words can express. Allow me then to say, quite simply – cheers for everything.

Simon – I always enjoyed our meetings. You provided a refreshing and different perspective that enhanced the quality of the research. I can honestly say that at least 50% of this thesis would not have been possible without your input. Thank you for help, it has been invaluable.

Tom – Things said in passing often have more value than intended. It is an off-the-cuff comment of yours that led to two experiments and an entire chapter. Nice one.

Alison and Dan – It's been interesting, informative, unpredictable, and most of all – entertaining. Your beaming smiles and eccentricity have been wonderful. Never change.

Published work

Experiments 1 and 2 were presented at the Cognitive Science Conference 2015, in Pasadena, Los Angeles. The experiments appeared in: Fereday,R., & Buehner, M. J. (2015). Temporal binding and internal clocks: Is clock slowing general or specific? In Noelle, D. C., Dale, R., Warlaumont, A. S., Yoshimi, J., Matlock, T., Jennings,C. D., & Maglio, P. P. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 686-691). Austin,TX: Cognitive Science Society.

Experiments 1 – 4 are in a manuscript entitled *Temporal Binding and internal clocks: No evidence for general clock slowing* (currently invited for revision with Journal of Experimental Psychology: Human Perception and Performance).

Table of Contents

List of Figures

List of Tables

Abstract

Temporal binding refers to the perceptual attraction of causally related events, which are perceived as closer together in time than unrelated events. This effect is not only characterised by the perceived attraction of cause and effect, but also by a contraction of the interval separating the events. Since the original article on temporal binding (Haggard, Clarke, & Kalogeras, 2002), research has identified the conditions necessary for the effect to occur. While predictability and contiguity are both necessary, it is causality and not intentional action that is the root of the effect (Buehner, 2012). Despite this fruitful work, little is known about how temporal binding is realised. Event perception approaches suggest that binding arises as a realignment of sensory streams. Time perception approaches, in contrast, suggest that binding arises due to a changes in temporal perception during the interval. Given the precedence for the latter approach in the literature (Humphreys & Buehner, 2009; Wenke & Haggard, 2009), I therefore applied an internal clock model of time perception to temporal binding. In Experiments 1 – 4 (Chapter 3), I explored whether binding is effected by the general slowing of a rate of an internal clock. Participants made verbal estimates of either an interval (in causal and noncausal conditions), or of an unrelated event embedded either before or during the interval. I hypothesised that changes in a general clock rate would both affect intervals and embedded events, such that events embedded during causal intervals would be judged as shorter than those embedded during noncausal intervals. The results revealed that causal trial intervals were judged as shorter than noncausal intervals, while no effect was found for embedded events. These results suggested that binding is effected by clock processes specific to cause-effect intervals. Experiments 5 - 8 (Chapter 4) examined whether binding might arise either due to changes

in a specific clock rate or to differential timing latencies. Using a temporal discrimination procedure, participants judged whether a variable duration interval was shorter or longer than a reference interval. The point of subjective equality (PSE) was computed for each reference duration, and then modelled using regression. The results revealed a significant binding effect, but more importantly, significant differences in regression slopes between causal and noncausal conditions in three out of four experiments. These results supported the hypothesis of a slower clock rate in temporal binding. In Experiments 9 - 10 (Chapter 5) I verified the results of Chapter 4 by examining discrimination thresholds between two causal and two noncausal intervals. In both experiments (Chapter 5), higher just-noticeable-difference (JND) thresholds were found in causal conditions, supporting the notion of a slower clock rate in cause-effect intervals. Taken together, the present body of work supports the notion that temporal binding is effected by a slower internal clock rate. Future experiments might investigate whether clock slowing in binding is driven by causality or predictability.

1. Chapter 1: Time and causality

Unlike the processing of a picturesque scene or an acoustic experience, in which the information is available directly via respective visual and auditory systems, the information afforded to the processing of cause and effect is not direct. The human brain is not equipped with a 'causality' sense mode that can directly perceive the causal link between two events. To understand that event A caused event B is only possible by a process of inference, using various cues to form a causal impression. These cues, which largely hark back to Hume (1739/1888), strongly emphasise temporal components. For instance, to understand that A causes B is to first observe that the occurrence of A temporally precedes B (temporal priority). If there are sufficient experiences of this pairing, such that B almost always follows A, then we can be more certain that A *causes* B. This cue (contingency; Allan, 1993; Elsner & Hommel, 2004; Mutter, DeCaro & Plumlee, 2009) can only be conveyed by repeated exposure to specific cause-effect pairs; if B follows A in 75 out of 100, and if B occurs on its own without A in 0 out of 100 trials, then we can be confident in judging A as the cause of B. If, however, B follows A in only 25 out of 100 occurrences, and if B occurs on its own without A in 75 out of 100 occurrences, then we are more likely to ascribe another event as a cause of B. Finally, a strong determinant of whether a causal impression is formed is temporal contiguity (Buehner & May, 2003; Buehner & McGregor, 2006; Shanks, Pearson & Dickinson, 1989). Consider pressing a button on a computer mouse or pressing a light switch, both of which are usually immediate outcomes of the pressing actions. What if pressing the mouse button did not immediately yield a response? Compare the following two conditions: pressing a mouse button triggers a small window to appear on screen immediately, versus a condition in which the window appears 20 seconds after pressing the

mouse button. Would both conditions trigger an equal-strength cause-effect link? Hume argued that causal impressions result when there is strong contiguity between two events. Therefore, causality is more likely to be felt when the window appears on screen immediately following the mouse button press, compared to the delayed (20 second) reaction.

Of course, it is important to note that whilst the roles of temporal priority, contingency, and contiguity in causal induction are generally supported (Greville & Buehner, 2015), there are exceptions. Prior knowledge, or beliefs about the relationship between two events can affect causal judgments (Buehner & May, 2003; Buehner & McGregor, 2006). If, to continue with the mouse button example above, we have prior knowledge of the expected delay between the action and the appearance of a small on-screen window, then the principle of temporal contiguity does not hold. For example, when clicking a mouse an hourglass symbol often appears, signifying that the request is being processed but, for whatever reason, is taking longer than usual. In this case we know that the mouse click has caused the outcome, even after a considerable delay. Despite the weak contiguity then, evinced with a delay between clicking the mouse and the desired result, the impression of causality remains. Thus, the impression of causality is affected not only by contingency and contiguity, but also by cognitive top-down mechanisms.

Causal associations are formed then, when one event follows another regularly and with contiguity. Of course, contiguity is not an absolute constant, but is situation dependent. Experiencing a delay between a mouse button press and an on-screen outcome on a computer may be expected, particularly if one has prior knowledge that the computer is slow in processing data. Conversely, a delay between activating the light switch and the bulb illuminating is unexpected, given that the connection is usually instant. Prior

knowledge about the expected time of outcome thus affects causal judgments (Buehner & McGregor, 2006). However, predicting the time of an outcome is not an easy task, chiefly because temporal information is not served by a dedicated sense modality (such as vision or audition). Instead, temporal information is based on a variety of factors, such as the modality of the stimulus (e.g., timing the duration of a sound or light stimulus; Wearden, Edwards, Fakhri & Percival, 1998), whether an interval is a single stimulus or delineated by two markers (Wearden, Norton, Martin & Montford-Bebb, 2007), the modality of the markers (e.g., the interval may be delineated by two auditory stimuli, two visual stimuli, or cross-modal combinations; Grondin, 2014), arousal (Droit-Volet, 2011) and attention (Tse, Intriligator, Rivest & Cavanagh, 2004). Judgments of time perception are therefore gleaned from a number of factors, making the temporal judgments particularly noisy.

Time then, affects our perception of causality. One might now reverse this statement by asking if causality affects our perception of time. From a Bayesian perspective, causality exerts a bias on time perception: If contiguous event pairs are more likely to be judged as causally related, then from a simple reversal, it follows that causally related events might be more likely to be judged as contiguous (Eagleman & Holcombe, 2002; Parsons, Novich & Eagleman, 2013). Put simply, if events occurring closer together in time are more likely to be judged as causally related, then events with a causal relation are more likely to be judged as closer together in time than unrelated events. Empirical work finds evidence for such a bias, known generally as ‘temporal binding’. The remainder of this chapter introduces temporal binding and describes the necessary and modulating factors of this effect.

1.1.1. Temporal binding

In the first article to report the temporal binding effect, Haggard, Clarke and Kalogeras (2002) used the Libet clock method to study the subjective time of various events. Famously used by Libet and colleagues in their seminal work into human volition in the 1980s (Libet, Gleason, Wright & Pearl, 1983), this procedure involved participants observing a clock face, with marked intervals corresponding to a standard analogue clock. In contrast to a standard clock however, a spot would complete one rotation every 2560 ms. Participants judged the time of certain events by reporting the position of the spot when a particular event occurred. Haggard et al. used this method (in which a clock hand replaced the rotating spot of Libet et al.) to study the perceived time of actions and their outcomes.

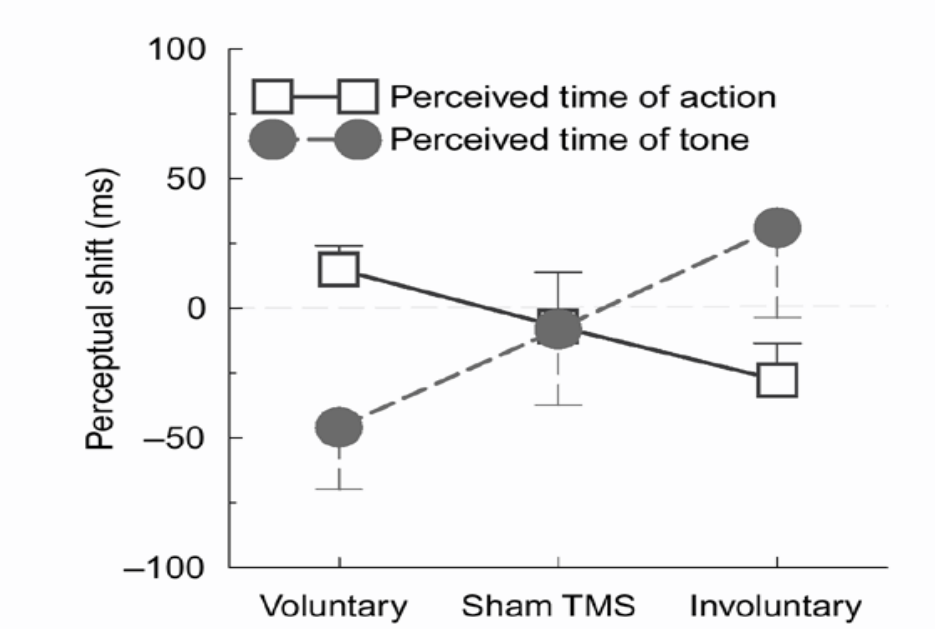

Figure 1.1. Results of Haggard, Clarke & Kalogeras (2002). In voluntary conditions, the key press was reported as occurring later, and the tone earlier, relative to the time events in the involuntary condition. Figure reproduced with permission.

There were two types of conditions in the experiment: baseline and operant. In baseline conditions, participants experienced single events. In some trials participants would

press a key, while in others they would experience a tone. In operant conditions, participants' key presses triggered a tone after a 250 ms delay. The experiment was blocked so that participants judged the position of the clock hand either when they pressed a key or when a tone sounded. Judgement errors (the difference between reported and actual event times) were used as a measure of event awareness. The authors found a systematic shift in awareness for events in operant conditions. Specifically, key presses that triggered a tone were reported later, and tones earlier, than in single event baseline conditions (see Figure 1.1). In other words, actions and outcomes were temporally attracted to each other.

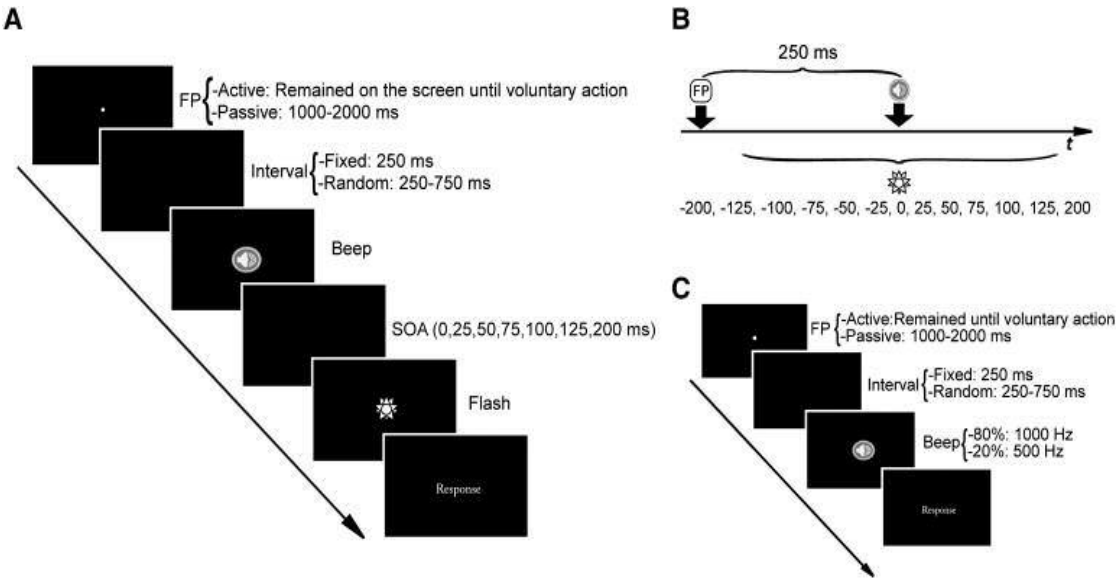

Figure 1.2. Schematic showing the trial structure of Cravo, Claessens and Baldo (2011). (A) In active or passive conditions, participants judged whether a flash occurred simultaneously with a beep. (B) The range of stimulus onset asynchronies (SOA) used in the Experiments. (C) Adaptation trials, designed to encourage participants' attention to the beep. Figure reproduced with permission.

Importantly, the opposite effect was found for conditions when a tone followed an involuntary TMS-induced finger twitch such that involuntary actions and tones were temporally repulsed (note that in this condition the finger twitch did not cause the outcome

tone). Thus, it may be inferred that the interval between voluntary actions and outcomes is temporally contracted, relative to the involuntary action-outcome interval.

Many studies have since replicated this binding effect (see Moore & Obhi, 2012, for a review) using the Libet clock method. However, this method is not without its critics. Engbert and Wohlschlaeger (2007) criticise the method on the grounds of attention, claiming that the continual monitoring of the rotating spot (or hand) may distract from the cognitive process underpinning action control. A further criticism is the large variability of temporal estimates. In the original Haggard study (Haggard et al, 2002) effects of 46 and 96 ms were found, using an identical procedure. According to Cravo, Claessens and Baldo (2011), this variability makes the quantitative comparisons of different conditions difficult. Although, it should be noted that high variability is commonplace in time perception and is not sufficient to abolish the binding effect, as Cravo et al. acknowledge. It simply means that saying the effect in one condition is larger than that in another by X milliseconds is difficult, due to the variability in judgments obtained across different experiments. Therefore, in order to investigate temporal binding without these confounds, Cravo et al. used a novel method. Drawing on psychophysical methods, the authors used simultaneity judgments as a proxy of event time, rather than the rotating spot method of the Libet clock. In their first experiment, participants completed two movement conditions; in active conditions, a voluntary finger lift triggered a tone after a 250 ms delay; in passive conditions, a tone sounded after a 250 ms delay following the disappearance of a visual stimulus (see Figure 1.2). An additional flash stimulus was scheduled to appear either simultaneously with the tone or at various temporal asynchronies. For example, the flash might appear 25 ms before or 25 ms after the tone, up to a maximum temporal asynchrony of 200 ms. Binary responses were collected, with participants judging whether the tone and flash stimuli were

simultaneous or not. Fitted functions allowed the authors to determine the Point of Subjective Simultaneity (PSS) for each movement condition, and revealed a lower PSS in the active relative to the passive condition. Specifically, the flash had to be presented approximately 30 ms before the tone in active condition, and 10 ms before the tone in the passive condition, in order for tone and flash to be perceived as simultaneous. In other words, because the beep was perceived as occurring earlier in the active, relative to the passive condition, the flash required a shorter PSS in order to be perceived as occurring simultaneously with the beep. Thus, the study replicated the temporal binding effect of Haggard et al. (2002), with a shorter perceived interval between voluntary actions and outcomes. In this case however, the binding effect manifests as the difference between voluntary action and computer-generated events, rather than between voluntary and involuntary actions.

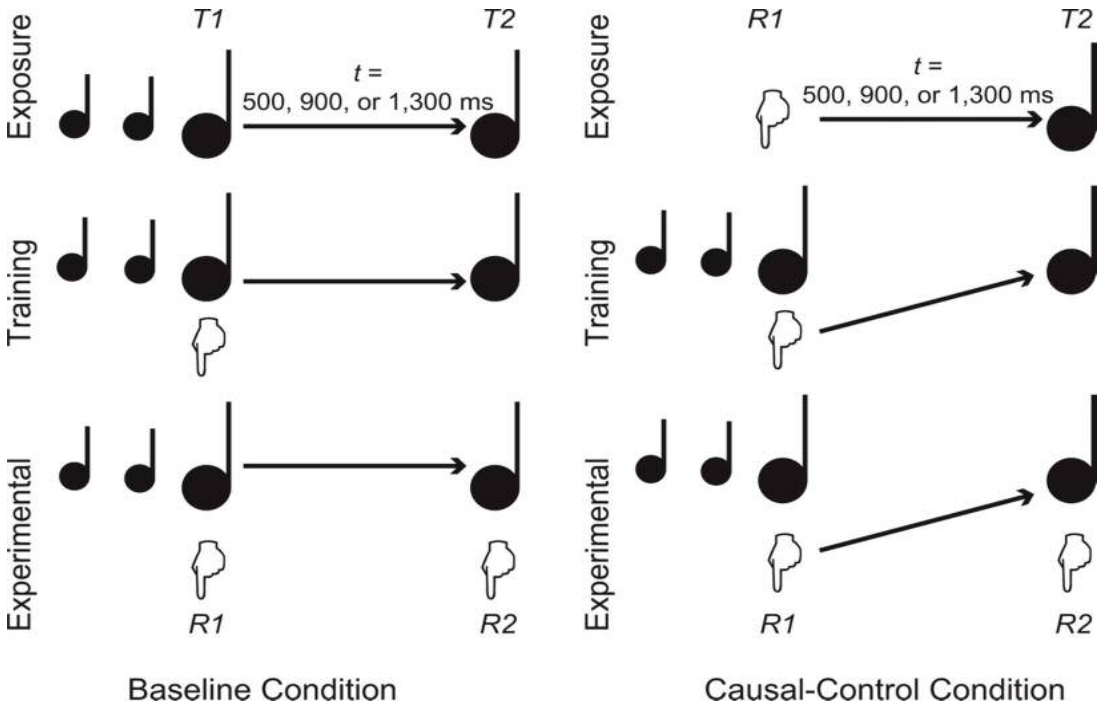

Figure 1.3. Schematic showing the trial structure employed in Buehner and Humphreys (2009). Figure reproduced with permission.

A method related to that of Cravo et al. (2011) uses event anticipation instead of simultaneity judgments. Buehner and colleagues, for example, developed the Stimulus Anticipation Method (SAM) to replicate and extend the binding effect (Buehner & Humphreys, 2009; Buehner, 2012). In the SAM, participants have to press a key in anticipation of a target event (see Figure 1.3). A series of studies using the SAM has repeatedly demonstrated early anticipation of target events triggered by an action, relative to target events that followed, but were not triggered, by a preceding event. In other words, participants perceived the target event (in these studies a tone stimulus) to occur earlier when triggered by a voluntary action than when the target event was computer-generated. Thus, binding manifests as a perceived shift in cause and effect events, whether using the Libet clock, simultaneity judgments, or the SAM procedure.

1.1.2. Is binding effected by shifts in events or by a contraction of time?

In addition to subjective shifts in events delineating intervals, many researchers have found temporal binding when directly investigating interval timing. This is a more direct and appropriate method of assessing temporal contraction in binding because event perception methods, such as the Libet clock, are based on inferred temporal contraction. For example, as Repp (2011) is careful to point out, there are two temporal binding effects: first, there is a shift of the action (or cause) *forward* towards the outcome; second, there is a shift of the outcome *backward* toward the action. Of course, the Libet clock method only measures one event (i.e., action or outcome) on any given trial. Typically, binding is defined as the judged error between the actual and experienced event time, so that the total judgment error is the sum of the (experienced - actual action time) + (actual - experienced outcome time). Thus, a larger judgment error in cause effect conditions reflects (implicitly) a shorter cause-effect interval. The downside with this method is that perceived interval

duration is based upon judgments of events computed on *different* trials. Therefore, making inferences about interval perception using event perception (i.e., Libet clock) methods is an indirect proxy at best. A better method would be to assess action and outcome events on the same trial, but this is too difficult a task. Therefore, direct judgments of the interval are used, which are usually interval estimation (often called verbal estimation) or reproduction (Wearden, 2008). In the former, participants must provide an estimate of a stimulus' duration in ms, either by typing a number into a computer or by using a continuous rating scale. In the latter, participants reproduce the experienced duration by holding down a key for a particular length of time. A number of studies have used these more direct measures of interval perception, which I now describe.

Using interval estimation, Humphreys and Buehner (2009) found estimates of intervals separating actions and outcomes to be lower than intervals separating two unrelated events. Other studies have also replicated the binding effect using this interval estimation procedure (Cravo, Haddad, Claessens & Baldo, 2013; Engbert, Wohlschlager & Haggard, 2008; Moore, Wegner & Haggard, 2009; Wenke & Haggard, 2009). Humphreys and Buehner (2009) also replicated the effect using temporal reproduction. They found shorter reproduced action-outcome intervals than intervals separated by two unrelated events. Using psychophysical methods, Nolden, Haering and Kiesel (2012) used a method of constant stimuli, in which participants judged comparison intervals as longer or shorter than a standard interval. Psychophysical functions were then fitted to the responses to determine the point of subjective equality (PSE). In line with other interval perception methods, the authors found lower PSEs for intervals separated by a key press and its outcome, than for intervals separated by two causally unrelated events. In other words, judging comparison and standard intervals to be of equal duration required a lower duration

threshold in action-outcome intervals, than in intervals marked by unrelated events. Thus, the results of Nolden et al. further support the finding that binding manifests as a contraction of the cause-effect interval, in addition to perceived event shifts.

1.1.3. Basic conditions of temporal binding: Predictability, contiguity, and contingency

In the studies mentioned so far, one thing is evident: temporal binding occurs only in conditions that involve intentional action. Are there other factors necessary for the binding effect, or is intentional action sufficient? Buehner and Humphreys (2009) found that intentional action per se, is not sufficient to elicit temporal binding. Participants in this study were told to synchronise their key presses with two tones (T1 and T2) in separate blocks. In a causal condition, T2 was contingent upon the key press, whereas T2 in the noncausal condition always followed by, but was causally unrelated to, T1. In other words, participants learnt that in noncausal trials, pressing the key in synchrony with T1 did not trigger T2, which followed T1 regardless of whether a key was pressed. However, in causal trials, participants learnt that T2 only followed a key press. Using the SAM, the authors found early awareness of T2 in causal, relative to noncausal trials. The authors argue that intentional action in itself is insufficient for temporal binding. Instead, it is the causal connection between action and outcome that is necessary for binding to occur. Thus, merely intending to perform an action is not sufficient to elicit temporal binding; it is the causal link between action and outcome that is essential.

The original Haggard et al. (2002) study considered the roles of predictability and temporal contiguity. In their second experiment the authors varied the action-outcome delay using durations of 250, 450 and 650 ms. By considering the subjective time of the outcome tone only, the authors found a decrease in the magnitude of the binding effect as interval length increased. Also, temporal binding was strongest when the action-outcome

delay was constant over trials; when the intervals randomly alternated, temporal binding diminished or disappeared altogether. Thus, temporal binding was strongest when the outcome was both predictable and contiguous.

The results of Cravo et al. (2011), in the study described above, support this finding. Cravo et al. used simultaneity judgments as a proxy of event awareness. In their study, the authors also included conditions in which the action-outcome interval was fixed (300 or 600 ms) or variable. When the interval was predictable but long (600 ms) the effect decreased, compared to conditions in which the interval was predictable and short (300 ms). Their results also revealed that binding was strongest at predictable (set), rather than random interval durations. Thus predictability and contiguity are both necessary for temporal binding.

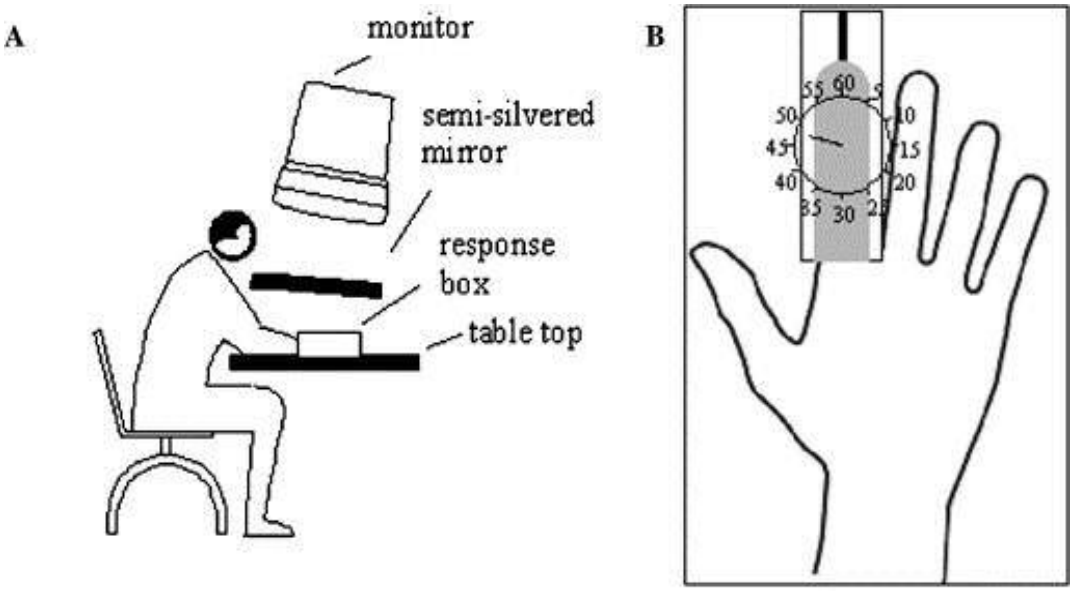

Figure 1.4. Schematic of the apparatus in Engbert and Wohlschlaeger (2007). Figure reproduced with permission.

Engbert and Wohlschlaeger (2007) also demonstrated the interaction of predictability and intentional action. In their study, participants viewed a miniature Libet clock projected onto a response box, upon which the participants' finger was placed (see Figure 1.4). Measuring only the time of action, the authors found that when an action triggered a tone on two-thirds of trials, the action subjectively shifted towards the tone to a greater extent than when the action triggered the tone with a one-third contingency. This difference did not hold for conditions where a tone merely followed a computer-generated action (see Figure 1.5).

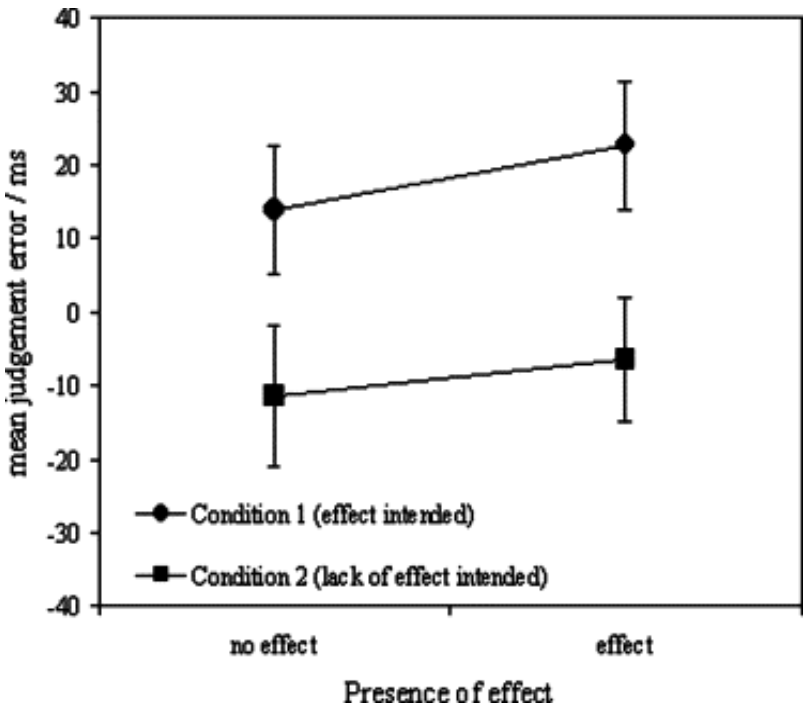

Figure 1.5. Results of Engbert and Wohlschlaeger (2007). Judgment errors (reported time – actual time) of action. The action was subjectively delayed in conditions where the effect was intended, whether the effect actually occurred or not. No delay was found in the unintended condition. Figure reproduced with permission.

Moore and Haggard (2008) further investigated predictability in temporal binding. The authors were interested in isolating the roles of predictive and retrospective processes. Specifically, the former refer to models of motor control, where voluntary actions involve a

prediction of the action's consequence. The latter, retrospective component, relies on inferential processes to establish the cause of an effect. Using the Libet clock method, participants' voluntary key presses produced tones after two different contingency levels: in a predictable condition, key presses triggered tones on 75% of trials, with the remaining 25% trials featuring key presses only. In an unpredictable condition, key presses triggered tones on 50% of trials with key presses only on the remaining 50%. Therefore, the authors were able to contrast trials in which key presses triggered outcomes, with trials in trials in which only key presses occurred. To isolate the predictive and retrospective components, Moore and Haggard analysed the reported action time in action-only trials (the authors did not measure the reported time of the outcome). These analyses revealed a larger subjective shift in predictable versus unpredictable conditions. This means that even when no outcome actually occurred, participants' reported actions were shifted forward in time, supporting the role of predictability in binding. Additionally, the results also a revealed a larger forward shift of reported action time in action-outcome trials, relative to action-only trials, in the unpredictable condition. Given that the outcome tone was unpredictable in action-outcome trials, the authors suggest that the presence of the outcome tone was responsible for the results, such that the tone was retrospectively shifted. This was later replicated and somewhat extended in Moore, Lagnado, Deal and Haggard (2009). Using a similar method, the authors found both that predictive and retrospective shifts were sensitive to contingency; higher contingency resulted in stronger binding. In sum then, predictability, contiguity, and contingency are not only necessary for causal inference, but also for temporal binding. However, none are sufficient for the binding effect; Instead, the literature is divided between intentional action and causality as the roof of the effect.

1.1.4. Is intentional action or causality at the root of temporal binding?

From the studies reviewed above it should be apparent that two key comparisons are used to evidence the temporal binding effect. One is between intentional and unintentional conditions (e.g., Haggard et al, 2002), while the other is between causal and noncausal conditions (e.g., Humphreys & Buehner, 2009; Zhao, Chen, Yan & Fu, 2013). The former comparison, used extensively by Haggard and colleagues, suggests that temporal binding is a product of intentional action; the absence of a binding effect in conditions where an outcome is generated by an unintended action, implicates the role of intentions, i.e., binding emerges as a consequence of intended actions. The latter comparison is broader in scope, implicating the more general role of causality in temporal binding. Does temporal binding then reflect intentional action or causality? For instance, consider observing another individual performing a temporal binding task, in which an action triggers an outcome after a short delay on every trial. Would there be a binding effect for participant and observer, or participant only? If binding reflects a process unique to self-generated motor commands, then no effect should be found for the observer. However, the evidence reviewed above finds that the basic conditions necessary for temporal binding (i.e., a meaningful causal connection between cause and effect) are also involved in causal perception. Accurately predicting the outcome of an action is improved when the interval between cause and effect is short (temporal contiguity), and when the outcome always follows the cause (contingency). Therefore, if causality is at the root of temporal binding (Buehner, 2012), there should an effect for the observer, in addition to the participant. I will explore the intentional account first.

Researchers (e.g., Engbert et al, 2008; Haggard & Clark, 2003; Tsakiris & Haggard, 2003) often explain binding in terms of a dedicated process of voluntary action. According to

this position, temporal binding is generated by processes dedicated to motor control (Moore & Obhi, 2012), which usually involve forward models of motor control (Wolpert & Miall, 1996; Wolpert & Ghahramani, 2000). These models express the action-outcome relation by generating predictions based on efferent copies of motor commands. At the time of movement an efferent copy is generated and processed by a dedicated motor control system, which in turn predicts the action's effect. When the prediction is matched by sensory feedback, the intensity of the signal is attenuated, leading to a less intense experience (Blakemore, Frith & Wolpert, 1999). Tsakiris and Haggard's (2003) results support this motor-predictive approach, with a binding effect for actions that triggered a somatic outcome (a TMS-induced twitch of an index finger). Also, somatic effects for voluntary-triggered actions were less intense than those for triggered by involuntary actions.

Evidence for 'intentional' binding appears in a study by Engbert and Wohlschlaeger (2007), who investigated whether intentions are necessary for temporal binding. Using the Libet clock method, participants were told that a section of the clock would be illuminated for a brief period. This period, they were told, was subliminal and would not be consciously perceived. Participants were told to press a key during this illumination period and were given feedback about their success (a tone either signalled a hit or miss during this target zone, depending on the condition). In reality, there was no such illumination period; it merely served as a cover story to provide meaning and intention to participants' actions. In the first experiment, participants either intended to produce an effect (in one conditions) or tried to avoid producing the effect (in another condition). Using the reported time of action as a measure of binding, the results (see Figure 1.5) revealed a delayed awareness of action when intending to produce the outcome (i.e., when a tone signalled a 'hit'). Importantly,

this result held even for trials in which no outcome actually occurred, although the delay was greater when the intended action actually resulted in an outcome. In the condition where participants avoided producing the outcome tone, no delayed awareness of action was found (regardless of whether an outcome actually occurred). Engbert and Wohlschlaeger concluded that intentional action is necessary for the binding effect.

Engbert et al (2008) examined binding within the context of self-, other- and machine-generated actions. The authors directly examined the temporal perception of the interval, rather than employ an event perception method. Their results revealed binding in only the self-generated action condition, leading the authors to conclude that binding is specifically a product of motor action, and is due to efferent motor commands.

The studies by Engbert et al. (2007; 2008) provide compelling evidence that binding is linked with forward models. However, Wohlschlaeger, Haggard, Gesierich and Prinz (2003) found binding for both self- and other-generated action conditions but no effect in machine-generated conditions. While forward models can explain the effect in self-generated conditions, they cannot account for the results in the other-generated action condition. Instead, the authors suggest that in the other-generated condition, participants adopted an intentional stance (Dennet, 1987), in which agency is attributed to other individuals. According to this view, individuals have access to their own private intentions, which generate the conscious awareness of action. By analogy with our own conscious intentions, individuals are able to infer the intentions of others based on their actions. This is similar to theory of mind (Baron-Cohen, 1995), in which the beliefs of other individuals are represented in our own minds. The authors suggest that representations of others' actions likely develop before the representations of others' beliefs, stressing that understanding the actions of other individuals is key to social understanding (Gallese & Goldman, 1998). The

authors further explain that the lack of binding in the machine-generated condition is due to the inability to represent the actions, and thus infer the intentions, of a mechanical object. Therefore, the results of this study suggest that binding emerges as a process of intentional attribution, both to ourselves and to others.

However, the results of Buehner (2012) challenge this intentional account of binding. Buehner found temporal binding in self-, other- and machine-generated conditions. Using the SAM, Buehner found early anticipation of the outcome in all three causal conditions, relative to a noncausal (baseline) condition. The author argues that causality, not intentional action, is the driving force of temporal binding. Indeed, Buehner suggests that it is causality that is responsible for the lack of binding in machine trials in Wohlschlaeger et al. (2003), not the inability to attribute intentions to mechanical objects; it might have been evident to participants that the computer was controlling the machine, thus obfuscating the causal link between machine action and outcome. In other words, outcomes might have been merely associated with their triggers in the machine-generated condition, analogous to a typical noncausal observational condition whereby an outcome follows, but is not caused by, a preceding event. The apparatus in Buehner's study likely ensured that the machine was perceived as autonomous mechanical agent, such that the machine caused (and was not merely associated with) the outcome. This situation might not have been sufficiently credible in Wohlschlaeger et al. Therefore, a binding of action and outcomes is indeed possible in machine generated conditions, providing there is a clearly perceived causal link. Thus, Buehner's results provide compelling evidence for the role of causality in temporal binding, over and above intentional action.

Whether temporal binding is rooted in causality or intentional action is still an on-going debate in the literature. More empirical work is necessary to fully lay claim to a causal

basis of binding. However, note that causal and intentional accounts of temporal binding are not mutually exclusive. The causal perspective is broader in scope and the intentional account is a special case of 'causal' binding. The key focus is the causal connection between action (or cause) and outcome, and the extent to which self-actions are involved. It may simply be the case that binding is stronger in voluntary conditions because there is less ambiguity about the cause-effect relation. In contrast, there is likely more ambiguity when observing cause-effect events generated by others human individuals, or possibly to a greater extent, by machines.

1.1.5. Learning, prior belief and knowledge

Related to the notion of causal ambiguity is that of learning. The evidence reviewed above finds that learning is an important factor in temporal binding, at least on an implicit level. In order to better predict the outcome of an action it is necessary to experience multiple cause-effect pairings, during which, causal information can be learnt; contingency and contiguity requires repeated exposure to cause-effect pairs, both of which increase the level of predictability. To what extent then, are prior knowledge and beliefs involved in temporal binding?

Cravo, Haddad, Claessens and Baldo (2013) investigated the effects of prior bias and learning in temporal binding. Using interval estimation, in which participants make direct temporal judgments of the interval, the authors replicated the binding effect with shorter estimates in operant conditions (relative to observational conditions, in which an interval was separated by two computer-generated events). More importantly, though, Cravo et al. found that interval estimates became shorter throughout the block in both conditions. That is, intervals subjectively decreased with repeated exposure. However, this trend was not sufficient to explain the binding effect, i.e., the difference in magnitude between conditions

did not increase as a function of trial number, it remained the same from trial one onwards. Given that this difference was evident during the first trial, the authors rule out learning effects. Instead, they suggest that participants had different prior biases for operant, relative to observational conditions. Thus, the results suggest a different bias for causal intervals than noncausal.

If temporal binding reflects a causal bias, as the results of Cravo et al. (2013) suggest, then to what extent are *beliefs* about the cause-effect relation involved? If participants in Cravo et al. were told that only half of all key presses produced outcomes but the actual contingency remained at 100%, then would the binding effect decrease in line with participants' beliefs? Desantis, Roussel and Waszak (2011) investigated the role of belief in the context of temporal binding by manipulating the apparent source of the outcome. Participants were led to believe that either their actions or a partner's actions produced an outcome, or the cause was ambiguous (both might have cause the outcome). In reality, the participant always caused the outcome; cues indicating the source were presented prior to action. Desantis et al. found a stronger binding effect when the participants believed they themselves caused the outcome, relative to the belief that their partner caused the outcome. The authors suggest that a priori causal beliefs can affect the predictive motoric processes of the forward model, such that motor prediction and actual feedback are closer matched when 'I' cause the outcome. Of course, these results also fit well with the more general causal account of temporal binding. Participants both were informed and could doubtlessly perceive a causal link between action and outcome, whether they believed their or their partner caused the outcome. The stronger binding in the self-belief action condition may simply reflect the close match between motor prediction and actual feedback, thus providing an additional boost above to observed causation. However, whether rooted in

causality or intentionality, the important point is that binding seems to be sensitive to higher-level contextual information.

To conclude this chapter, binding occurs not only in the perceived shifts in events delineating intervals, but also manifests as a contraction of the interval itself. The evidence reviewed above finds that the principles underlying causal inference (such as regularity, contingency and contiguity) are all necessary conditions for binding to occur. However, these conditions – whilst necessary – are not sufficient. The issue of whether binding is attributable to intentional action or to causality, has divided opinion in the literature. Whilst Haggard (e.g., Haggard & Clark, 2003) maintains that binding is a product of intentional, relative to unintentional actions, other researchers (e.g., Buehner, 2012; Buehner & Humphreys, 2009) evidence causality as the sufficient condition. Implicated in the latter perspective are the roles of prior knowledge and belief, both of which influence the magnitude of the binding effect. Regardless, though, of whether causality or intentional actions are the sufficient conditions for binding, the key question is *how* causality biases temporal judgments, i.e., by what processes do causally related events bind together in time? Eagleman argues that binding occurs due to a realignment of perceptual streams (Eagleman & Holcombe, 2002; Parsons, Novich & Eagleman, 2013). Another view is that causality directly affects time perception, such that temporal binding occurs due to actual changes to perceived inter-event intervals. In this thesis I investigate the latter idea, of whether actual changes in time perception may be realized in temporal binding. First, however, I will describe event and time perception approaches in the following chapter.

2. Chapter 2: Theories of time perception and their application to temporal binding

In this chapter I describe theories of temporal binding. Specifically, I describe theories of *how* binding is effected (i.e., by what process is binding brought about), rather than theories that describe *why* binding occurs (see Chapter 1 for a discussion of these). Theoretical explanations of binding can be split into two classes: event and time perception theories. Event perception approaches place emphasis on the events delineating cause-effect intervals, such that the cause (action) is shifted forward and its outcome backward in time (Haggard et al., 2002). This perceptual shift in events affects interval judgments, such that the cause-effect interval is perceived as shorter than an interval delineated by unrelated events. The key point here is that cause-effect intervals are perceived as shorter simply due to the shift in cause and outcome events. The time perception approach differs to the event perspective in that shorter judged cause-effect intervals are not merely a by-product of event shifts, but posit an actual change to our temporal perception of the interval. This is essentially the reverse of the event perspective, in that cause and outcome events are temporally shifted together because of a change in our perception of the interval (although, note that the theories are not mutually exclusive: it is equally possible for changes in time and event perception to occur simultaneously).

I will first describe theories of time perception, before briefly describing event perception approaches to temporal binding. I will conclude this chapter by explaining how a time perception model can explain temporal binding.

2.1.1. Time perception accounts

Temporal binding paradigms involve the judgments of upcoming events (or intervals). Participants know in advance that judgments of an interval, for example, are required; participants are usually presented with an interval, and make judgments of its duration after its presentation. This is known as prospective timing; although judgments are made after exposure to the interval, participants know in advance that temporal judgments are required. In retrospective paradigms, in contrast, an individual is unaware that timing judgments are required until after the event. The key difference between the two is that retrospective judgments are implicit and derived from episodic memory, while prospective timing is explicit and requires attention to temporal information (Block & Gruber, 2014; Zakay & Block, 2004). Therefore, judgments of duration depend on whether or not a person is aware that temporal judgments are required; theoretical models that account for prospective timing cannot be applied to remembered durations (i.e., retrospective timing). Since temporal binding involves the judgments of upcoming events, I will focus on prospective timing for the remainder of this thesis.

2.1.2. Internal clock models: Basic components

The most popular and indeed, most successful, models of timing represent perceived durations with a pacemaker-accumulator system (Creelman, 1962; Treisman, 1963). Also known as a clock-counter device, this timing system is based on the idea that humans use an internal clock on timing tasks. There are three sub-processes to this model (see Figure 2.1). First, the clock stage consists of a pacemaker-accumulator system that is responsible for the emission and counting of pulses. When timing a stimulus, a switch closes that allow the pulses to flow from the pacemaker into the accumulator. When timing is complete the switch opens, thereby stopping the pulses from flowing. Second, the pulses accrued during

the timing of a stimulus are stored in working memory, which are then compared to a long-term memory store for a particular duration. Third is the decision process, which might involve judging whether the duration in working memory is shorter or longer than a reference stimulus, or in other cases, providing an explicit duration judgment.

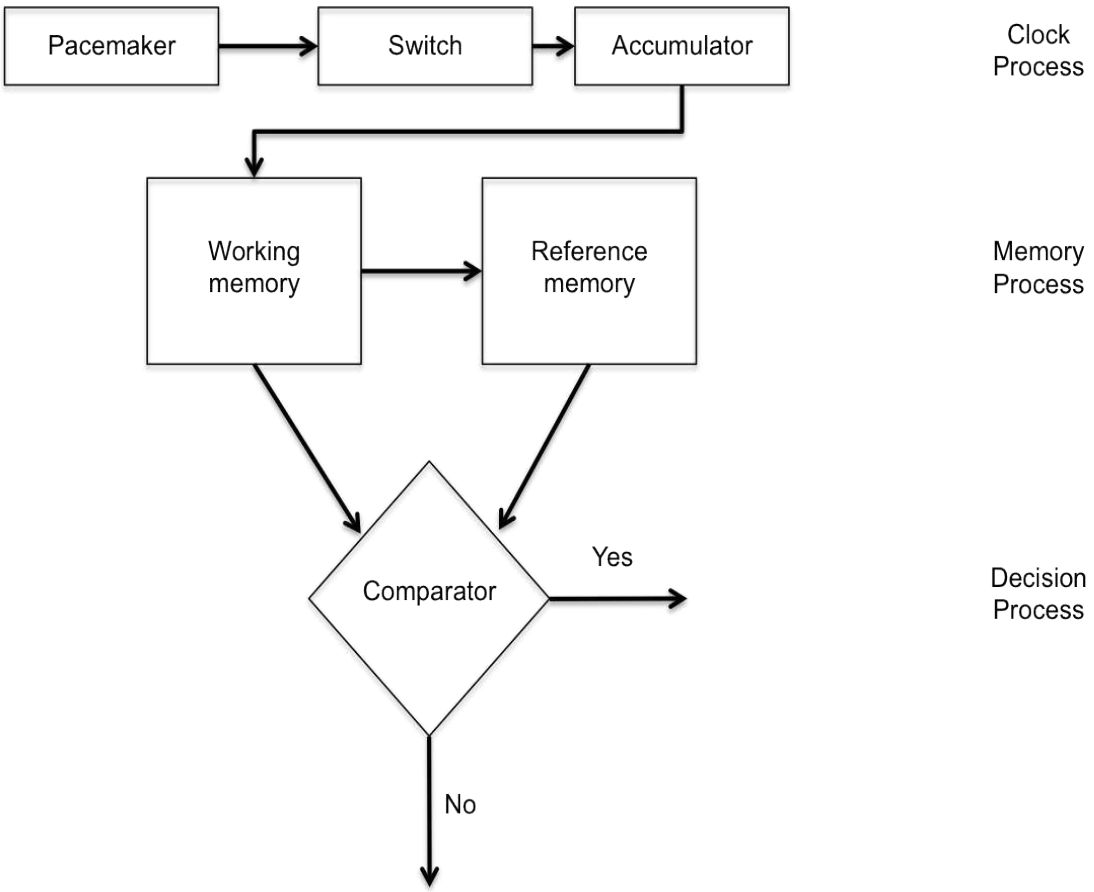

Figure 2.1. The information-processing model of Gibbon and Church (1984).

An example will suffice here. Consider a task where you are required to judge the duration of a tone, with an actual duration of 500 ms. When the tone begins the switch closes, allowing pulses to flow from the pacemaker to the accumulator. When the tone ends the switch opens, and no further pulses are accumulated. The number of pulses accumulated forms the basis of the temporal judgment; more pulses lead to longer perceived durations

and fewer pulses to shorter perceived durations. For this example, assume that the 5 pulses have accumulated, that each represent 100 ms (i.e., the judgment is veridical). This duration is then compared to a long-term store and evaluated - is the working memory duration longer, shorter or the same as the duration in reference memory? In the present example both durations are the same, and a judgment of 500 ms is given. However, judgments of time are rarely veridical, such that judgments of duration are either overestimated or underestimated. One factor that affects time perception, is arousal (Bar-Haim, Kerem, Lamy, & Zakay, 2010; Droit-Volet & Meck, 2007). For example, a 500 ms stimulus would be perceived as longer under emotionally charged, relative to neutral conditions. This fits well with a pacemaker-accumulator system, such that arousal increases pacemaker rate, which then results in a longer subjective interval. Thus, the internal clock model is parsimonious, in that the pulses emitted from the accumulator can account for a wide variety of temporal phenomena.

2.1.3. Scalar Expectancy Theory (SET)

A popular instantiation of the internal clock model is scalar expectancy theory (Gibbon, 1977; but see Allen, 1998; and Wearden & Lejeune, 2008, for reviews). The central idea behind SET is that timing sensitivity remains constant across a range of durations. In other words, the variability in temporal judgments can be rescaled as a function of stimulus duration (Grondin, 2010). A method of computing this scalar component is to divide the standard deviation by the mean for a range of durations. This is a type of Weber fraction, known as the coefficient of variation (CV). There are various sources of variability within SET, but in general, human timing conforms to this *scalar property of variance*, with the CV remaining constant for a range of durations (for a review of this scalar component in humans, see Wearden & Lejeune, 2008; and for animals, see Lejeune & Wearden, 2006).

Another feature of SET is the finding that, despite being inaccurate, human timing is nevertheless precise. That is, participants perform with reasonable precision across a range of durations, such that subjective judgments tend to increase linearly with real-time durations (Grondin, 2001, 2010; Wearden, 2004; Wearden & Lejeune, 2008). This - *mean accuracy* - is a further property of SET, which is that subjective times are average approximations of real durations.

Why is the scalar property important? Wearden (2003) provides an appropriate example. Imagine testing the time perception of two different populations - children and adults. Performance will undoubtedly differ between each group, as would the performance of adults compared to the elderly, and brain damaged patients compared to healthy controls. This merely demonstrates that different groups of people perform differently on timing tasks; it does not provide any insight into *why* performance is different. The application of SET allows researchers to provide more probable and meaningful surmises as to the underlying cause of performance differences. Chiefly, SET allows for the underlying variability to between populations to be computed, rather than simply compare mean scores.

2.1.4. Measuring perceived time: discrimination methods

Time perception researchers use a range of techniques to measure perceived time. Some take the form of discrimination experiments, which involve judging whether a particular duration is shorter or longer than another. Examples of discrimination methods involve temporal generalization (McCormack, Brown, Maylor, Darby & Green, 1999; Wearden, 1992), bisection (Allan & Gibbon, 1991; Wearden 1991), and adaptive psychophysical procedures that determine a temporal threshold (Rammsayer, 1999). Other 'classical' methods use direct estimates of intervals, such as providing verbal estimates (in

milliseconds [ms] or seconds [s] perhaps) or reproducing an interval by holding down a key for a certain duration. Examples of classical approaches involve interval estimation (Penton-Voak, Edwards, Percival & Wearden, 1996), reproduction (Wearden, 2003), and production (Wearden & McShane, 1998). I shall first describe discrimination methods then classical techniques.

The temporal generalization procedure is rather elegant in its simplicity. It involves asking participants if a duration is equal to another. In a learning phase, participants learn the duration of a reference stimulus. In a test phase a series of comparison durations shorter than, equal to, and longer than the reference are presented. Participants then judge whether each comparison duration is equal to the reference duration, yielding a YES/NO response. Results are obtained by computing the proportion of YES responses for all comparison durations tested. Correct identification of the reference stimulus is ascertained by examining the proportion of YES responses when the reference and comparison stimuli are equal. A higher proportion of YES responses indicate that intervals in one condition are perceived as longer than another. Moreover, timing sensitivity can be assessed by using a variety of reference durations; by plotting the response curves from different reference durations on the same scale and plot (a method known as superimposing), timing sensitivity can be ascertained. Aligned curves (i.e., superimposed) imply that timing sensitivity is constant. Thus, mean accuracy and timing sensitivity can both be assessed using temporal generalization.

Another procedure, similar to generalization, is the temporal bisection task. Participants are again exposed to test and comparison intervals but the procedure is slightly different. In a training phase, participants are exposed to two stimuli, of short and long durations. In a test phase, a series of comparison intervals between, and including, the test

durations are presented. For any comparison duration presented, participants judge whether the stimulus is closer to the short or the long duration. Results are obtained by estimating the PSE (point of subjective equality), defined as the point at which 50% of the responses are judged as closer to long; a lower PSE in one condition shows that the stimuli are perceived as longer (i.e., longer perceived stimuli would require a shorter duration to be judged as equal to a reference stimulus). Moreover, the scalar component can be assessed either by superimposing the curves, or by calculating the CV, which involves dividing the just noticeable difference (JND: the actual stimulus change necessary to be detectable, computed as 75 - 25% response probability/2) by the PSE. In the former case, if the deviation to the mean is constant (for different ranges of durations), then the curves ought to superimpose. In the latter case, the JND and PSE are analogous to the standard deviation and the mean, respectively. The strength of temporal bisection then, like the generalization procedure, is that participants are only required to judge whether a test stimulus is shorter or longer than a reference stimulus (this is in contrast to making explicit duration judgments).

There is another method that involves the discrimination of reference and comparison intervals, but in this case the duration of the comparison intervals vary from trial to trial. This task, known as a temporal discrimination procedure, use adaptive staircase methods procedures to vary the intensity of a test stimulus (Cornsweet, 1962; Johnson, 1992; Leek, 2001; Spielmann, 2013). In a typical experiment (e.g., Rammsayer, 1999), participants are presented with the reference and comparison intervals, the order of which might vary from trial to trial. Participants are asked to judge which interval is the longest, or whether one interval is longer or shorter than the other. The duration of the comparison interval is controlled by a staircase procedure; if participants respond ‘correct’ then the

comparison interval is increased in duration on the next trial. On the other hand, if participants responded 'incorrect' then the duration of the comparison interval is decreased on the next trial. Studies usually target the threshold at which the participant is correct on 75% of trials, but it is also possible to determine either the PSE or JND by fitting psychophysical functions. To calculate the CV using this technique, a ratio of the JND/PSE may be computed (Wearden & Lejeune, 2008). In sum, adaptive methods, like generalization and bisection techniques, provide an effective way of exploring temporal perception.

2.1.5. Measuring perceived time: 'classical' methods

Classical methods of time perception differ to discrimination methods in that no reference durations are presented. Verbal estimation, for example, involves making subjective estimates of an interval without prior exposure to a reference stimulus (e.g., Wearden, Norton, Martin & Montford-Bebb, 2007). Therefore, the judgment of an interval in the first trial is based on long-term internal reference memory. Participants are usually exposed to a range of durations, and are required provide explicit duration estimates. Mean accuracy is assessed by inspecting the linear trend in mean judgments; judgments should increase with actual durations. The CV (standard deviation/mean) should be constant if scalar variability is supported. Interval estimation, despite no external reference stimulus, is a popular and effective method in time perception research.

Other techniques are temporal reproduction and production. Reproduction involves holding a key for a duration corresponding to a perceived interval. Production, on the other hand, involves holding a key for a given duration; an experimenter might instruct the participant to hold a key for 300 ms, for example. In this case, the participant will have an internal representation of 300 ms and hold the key down for the perceived duration. The

basic difference is that in reproduction, participants are presented with an interval and then required to reproduce it, whereas no interval is presented in experiments involving the production technique. However, both techniques are commonplace in time perception research.

2.1.6. Disentangling internal clock processes

Research using discrimination and classical techniques show that human timing may be precise, but not accurate (Grondin, 2010; Wearden, 2003). That is, perceived durations tend to increase linearly with actual stimulus durations but are rarely veridical. What accounts for this inaccuracy? One source is pacemaker reliability (Grondin, 2010; Wearden & Lejeune, 2008), such that the rate of the pacemaker varies according to different conditions. Another is that the switch to start and stop timing is delayed. How can pacemaker reliability and switch latencies be disentangled? Consider a study of Wearden et al. (1998) on the temporal perception of auditory and visual stimuli. In their Experiment 2, participants estimated a range of durations (from 77 - 1181 ms) in both modalities. Wearden et al then conducted individual participant regressions, by regressing interval estimates onto the actual stimulus durations. The coefficients of these regressions represent different aspects of the internal clock model: the slope coefficient represents pacemaker rate, while the intercept represents switch latencies. Therefore, differences between slope coefficients would imply that timing in auditory and visual modalities are effected by different pacemaker rates. Meanwhile, differences between intercept coefficients would imply different switch latencies between the modalities, such that there is a greater stop/start delay in one modality. Their results revealed no difference between intercepts but a significant difference between slope coefficients. In particular, the slope of the visual condition was shallower than the auditory condition, implying a relatively slower

pacemaker for visual stimuli. Interestingly, evidence shows that the pacemaker also runs at different rate for filled versus unfilled stimuli (Wearden et al., 2007), for stimuli preceded versus not preceded by clicks (Penton-Voak et al., 1996), and for loud versus quiet tones (Matthews, Stewart, & Wearden, 2011). Overall, there is a substantial body of evidence demonstrating that variation in pacemaker rate can account for the inaccuracies found in time perception (Matthews, 2011; Matthews & Meck, 2014).

Can switch latencies account for differences in judgments of timing? There is scant evidence for this. The majority of temporal illusions evidence variable pacemaker rates. However, buoyed by these findings, Matthews (2011) investigated whether switch latencies *ever* occur. In the first experiment, participants estimated a range of intervals marked by different visual events. In both conditions the stimuli marking the start of intervals were different in size than those marking the end, such that the end-of-interval marker was either substantially or slightly different in size from the first marker. Trials began with either the smallest or largest square first; in small first trials, the second marker was either substantially or slightly larger (deemed large or small jump conditions, respectively). In large first trials, analogously, the second marker was smaller in size by either a small or large jump. The intention was to determine if intercept differences can be found, when pacemaker rate differences are impossible, i.e., because the first marker and the interval are identical for small and large jump conditions, any difference in estimates can only be due to switch latency at the end of the interval. Matthews' results revealed differences in intercepts only, between conditions where the marker size difference was either large or small, providing evidence that intercept differences *can* occur. However, in his Experiment 3, there were conditions in which the stimuli marking the start of the intervals were the same in both in conditions, but crucially, the end of interval markers differed in size

between conditions. The logic of this approach was that the pacemaker cannot change rate at the start of the interval because the same sequence of events are common to both conditions. Nevertheless, Matthews found evidence of slope differences, implying different pacemaker rates. It is not clear why this was the case, but nevertheless, the key point is that Matthews' study provides rare evidence that intercept differences can occur.

In sum, internal clock models explain differences in subjective time under various conditions as arising either from different pacemaker rates or switch latencies.

2.1.7. What modulates subjective time? Top-down factors

In the section above I described how pacemaker rate and switch latencies can account for subjective distortions of time, and how regression analyses can disentangle these processes. The question now is, what are the factors that distort subjective time?

A fundamental factor in time perception research is attention (Tse, Intriligator, Rivest & Cavanagh, 2004). This is commonly used to explain perceived durations (Grondin, 2010), such that attending to time increases subjective duration. Within the SET framework, attention is purported to operate via switch latencies, such that the attentional switch is not synchronous with veridical timing. The switch may operate at different latencies, depending on attentional demands in the particular context. Is the increase in perceived duration merely due to the amount of information processed or to increased attention? Tse et al. addressed this question. They argued that because attention requires ~120 ms to allocate to a new stimulus, that any subjective expansion effect should not be found under these durations if only attention is involved. Additionally, they argued that because attention is a central mechanism, then its effects should not be modality specific. The authors used the oddball paradigm, in which an expanding black circle (the oddball) was embedded within a sequence of static black circles (standard stimuli). They found that the oddball was

perceived as longer than the standard stimuli, an effect that also occurred with a static oddball, and with auditory stimuli. Importantly, they found no subjective expansion at the lowest duration tested (75 ms), but the effect was evident at durations above 135 ms. Tse et al. therefore attribute the effect to increased attention for oddball stimuli. Thus, attentional capture distorts subjective time, an effect that has been replicated by Eagleman and colleagues (Eagleman, 2008; Pariyadath & Eagleman, 2007).

An ubiquitous idea in time perception is that attention is resource limited. In dual-task paradigms, a concurrent (non-temporal) task interferes with timing, such that durations are perceived as shorter (Brown, 2006, 2008; Brown & Merchant, 2007) relative to single-task conditions. Interestingly, research also shows that time perception is affected by a preceding non-temporal task. Wearden, O'Rourke, Matchwick, Min & Maeers (2010) found that switching between an addition task and a timing task resulted in decreased duration judgments. Furthermore, time perception is also affected when timing multiple, partially overlapping stimuli; temporal accuracy is shown to decrease as the number of events to be timed increases (Klapproth, 2011). In general then, time perception suffers as attentional demand increases, supporting the notion that attention is resource limited.

When reviewing the factors that distort subjective time, one must necessarily include arousal. To be clear, arousal is an umbrella term for a range of specific factors. For example, emotional (Droit-Volet & Meck, 2007) and physiological factors (Tamm, Jakobson, Havik, Burk, Timpmann, & Allik et al, 2013) affect subjective time. Droit-Volet, Brunot, and Niedenthal, 2004, for example, show that intervals are perceived as longer when angry faces are presented during the intervals, compared to neutral. Likewise, emotional sounds are perceived as longer than neutral sounds (Noulhiane, Mella, Samson, Ragot, & Pouthas, 2007). Additionally, Stetson, Fiesta and Eagleman (2007) conducted a novel experiment in

which participants experienced a free fall of 31 metres. Stetson et al. found that participants' estimates of their own free fall were longer than the estimates of other participants' free falls. Interpreted via SET, the common explanation for these results is that arousal accelerates the rate of the pacemaker.

2.1.8. What modulates subjective time? Interval structure

In addition to attention and arousal, the structure of an interval also affects time perception. Whether an interval is filled or unfilled (Wearden, Norton, Martin, & Montford-Bebb, 2007), the duration of the interval markers, and the modality of the markers all affect temporal judgments (Grondin, Roussel, Gamache, Roy & Ouellet, 2005). Grondin et al. investigated how marker duration and modality affect temporal discrimination. Intervals were marked by either auditory or visual stimuli, and of either short or long durations. The results revealed that duration judgments were unaffected by marker duration, but were affected by modality; intervals marked with an auditory and visual stimulus (auditory-visual) were perceived as longer than visual-visual marked intervals. The authors also examined variability, and found that intervals marked by 500 ms stimuli were more variable than intervals marked by 10 ms stimuli. Additionally, variability was reduced when intramodal (visual-visual) markers were used, compared to intermodal (visual-auditory). However, time perception also differs depending on whether an interval is marked by two stimuli or is a single stimulus. Wearden et al. (2007) found that intervals consisting of a single stimulus (e.g., a tone) were perceived as longer than an interval delineated by two markers (e.g., two tones). Therefore, the perception of time is dictated not only by top-down processes, but also by the structure of events that define an interval. Thus, interval structure is another source of variability to consider, when comparing conditions in time perception studies (in addition to top-down processes).

2.1.9. Is timing general?

An interesting question in time perception is whether timing is a general process or is underpinned by separate mechanisms (Eagleman, 2008). In other words, when subjective time changes, does time in general change or is the change specific to the stimulus? In one study that addressed this question, Pariyadath & Eagleman (2007) presented sequences of repeated stimuli, one of which was unexpected, i.e., an image of a clock embedded during a series of shoe images. Participants judged whether the oddball (the clock) was shorter or longer than the images of shoes preceding and succeeding it. The results revealed a perceived expansion of time for the oddball, compatible with the results of Tse et al. (2004). Importantly, to investigate whether this perceived dilation reflected a general clock process, additional experiments presented simultaneous auditory stimuli (tones) or flickering images. The authors reasoned that if the perceived expansion of time reflects a general increase in temporal acuity, then (1) the pitch of tones should appear at lower frequencies and (2) flickering images should be perceived at higher frequencies (relative to preceding and succeeding stimuli). The authors found that, while the duration of the oddball was perceived as longer than the repeated stimuli, there was no perceived difference in frequency, either for the images or the tones. The authors replicated this finding in their 2007 study (Stetson et al, 2007). Motivated by the anecdote that time slows down during frightening events, the authors tested whether time in general slows down. Consequently, if time slowing is general then there should be an ability to perceive events at a higher temporal resolution, such as perceiving a higher flicker rate when falling. Accordingly, participants wore a specially constructed watch that presented flickering digits. The threshold for correct digit identification was determined before the fall, and subsequently increased prior to the fall, during which participants had to identify the digits. In accordance with the Pariyadath &

Eagleman, the authors found that participants judged their own fall as lasting longer than others', yet there was no increase in temporal resolution during the fall. Thus, the authors conclude that time "is not one thing, but is instead composed of separate neural mechanisms that usually work together but can be teased apart in the laboratory" (Eagleman, 2008, p.133).

2.1.10. Internal clocks and temporal binding

Thus far, I have described the how changes in internal clock processes can account for changes in subjective duration, and the various factors that affect duration judgments. Most studies find that timing differences between various conditions are underpinned by a difference in pacemaker rate. Therefore, given that temporal binding refers to a contraction of time during cause-effect intervals, would this be effected by a slower pacemaker?

To answer this question, Wenke and Haggard (2009) combined the temporal binding procedure with a temporal discrimination paradigm: In a causal condition, participants pressed a key to generate a tone after 600, 800 or 1000 ms; in a corresponding non-causal condition, their finger was passively pulled down by a motor, which was followed by a tone – the computer scheduled both events, making it obvious that there was no causal connection between the passive movement and the tone. In addition, participants experienced two electric shocks administered via electrodes on their fingers. On a given trial, participants were prompted to either report whether the shocks were successive or simultaneous, or to estimate the duration of the action/movement – tone interval. In addition to replicating the binding effect, Wenke and Haggard found that simultaneity detection on causal trials was poorer (i.e. higher thresholds) compared to noncausal trials. This is what would be predicted if temporal binding is effected via a slowing of the internal clock: a slower pacemaker lengthens the period between pulses, increasing the likelihood that two sequential shocks

fall into the same period, which in turn leads to higher discrimination thresholds. Thus, the authors concluded that temporal binding is effected via a slower pacemaker.

However, there are a few concerns that urge caution, rather than accepting the conclusions at face value. It may be that the higher threshold for discrimination judgments is not due to a slower clock, but instead simply reflects a tactile specific contraction of time. For example, Tomassini, Gori, Baud-Bovy, Sandini, and Morrone (2014) measured time perception surrounding the onset of voluntary movements. Tactile pads were attached to a static hand and a moving hand, and delivered two pairs of taps. Participants judged the duration of the second variable-spaced pair of taps (presented toward the end of each trial) in relation to the first fixed-spaced pair (presented prior to hand movement). The authors found a contraction in conditions where the taps were delivered to the moving hand, relative to the when taps were delivered to the static hand, i.e., an effector-specific contraction of time just prior to and immediately following hand movement. Therefore, Wenke and Haggard's (2009) results might simply reflect a similar, tactile-specific process, rather than a contraction of time specific to temporal binding. However, one glaring issue with this study is the relation between actual and estimated intervals. A regression of the latter onto the former would yield slope differences between causal and noncausal conditions. This was not the case with Wenke and Haggard, who found no slope differences.

A study that provides better evidence of clock slowing in temporal binding is Buehner and Humphreys (2009). They investigated temporal binding across a range of durations, notably finding a binding effect for durations up to 4 s. A regression of interval estimates onto actual interval durations found shallower slopes for causal, compared to noncausal trial intervals, indicative of a slower pacemaker rate in causal intervals. However, one issue with this study is the use of verbal estimation, which, as Matthews (2011) has

shown, often produces slope differences when none should occur. Therefore, it is not possible to conclusively say that a slower clock operates in temporal binding on the strength of current, purported evidence. Furthermore, it is unknown whether potential clock slowing in binding is specific to the cause-effect interval, or is a general timing process. Therefore, more empirical work is necessary to discover whether temporal binding is effected by a slower pacemaker.

2.1.11. Event perception accounts

Although this thesis is concerned with a time perception account of temporal binding, it would not be permissible to exclude event perception approaches. Therefore, I briefly describe how an event perception approach can account for the perceptual attraction of events that is temporal binding.

The term 'binding' refers, in the domain of perception, to the integration of sensory information. For example, light travels faster than sound - 300,000,000 compared to 330 metres per second to be exact (Spence & Squire, 2003). Yet, when observing an individual clicking his or her fingers, the click sound and the movement of the fingers are perceived at the same time. However, when experiencing thunder and lightning, we see the lightening *before* hearing the thunder. Why are these auditory and visual streams not integrated into a single event? According to the 'horizon of simultaneity' (Poppel, 1988; Sugita & Suzuki, 2003) the brain temporally binds auditory and visual information into a unitary event, up to a distance of around 10 - 15 metres. Sugita and Suzuki, for example, measured the effect of distance on perceived simultaneity. Participants were exposed to auditory asynchronies via headphones, and watched LEDs at a range of distances. The participants were told to imagine the LEDs as the locus of the sound, and make simultaneity judgments about the auditory stimuli. The results showed that the point of subjective simultaneity (PSS: the point

at which the auditory and visual stimuli were perceived to co-occur) increased with distance, up to 15 metres. In other words, the window of integration increased with distance, allowing for the fact that sound will lag behind visual signals as distance increases, such that a wider window of integration is necessary at longer distances. Thus, there is no absolute (temporal) lag threshold for unification but rather the window moves as audiovisual stimuli become more distant (Spence & Squire, 2003). This extends work by Dixon and Spitz (1980), who found that when auditory speech and visual lip movements are separated by an interval of up to 250 ms, they are perceived as unitary. The window of integration then, proposes (1) that temporal asynchronies are perceived as simultaneous if the events occur within the same window; and (2) the limits of the window temporally expand with distance.

A 'window of integration' is one account of sensory integration. Another is 'temporal ventriloquism', which refers to the perceptual shift of sensory events into a unified percept; input from one sensory modality is shifted in time into alignment with another sensory input (Chen & Vroomen, 2013). The most striking aspect of temporal ventriloquism is that visual stimuli are shifted into alignment with auditory stimuli. In an early study, Gebhard and Mowbray (1959) presented participants with a flickering light and a corresponding fluttering sound. They found that a change in the flutter rate of the auditory stimuli resulted in a corresponding change in the perceived flicker rate, but not vice versa (i.e., visual flicker rate did not alter the perceived rate of the fluttering sound). Similarly, in more recent studies using tapping tasks, participants were able to tap in synchrony with an auditory click while ignoring a visual stimulus (a temporally misaligned flash), but were not able to ignore an auditory distractor while tapping to a visual flash (Repp, 2005). In a study with single events, Morein-Zamir, Soto-Faraco, and Kingstone (2003) found that when an auditory stimulus

follows a visual stimulus after an interval, the visual event is perceived as shifted forward in time towards the auditory event. This auditory dominance occurs because audition is the dominant mode for temporal information, according to the theory of 'modality appropriateness' (Welch, DuttonHurt, & Warren, 1986). A spatial dominance occurs when sound is ventriloquized towards the spatial location of the stimuli (Bartelson & de Gelder, 2003; see also Chen & Vroomen, 2013, for a review of sensory binding). However, the main point is that temporal ventriloquism accounts for the binding of temporal asynchronies by bringing events into alignment.

While most studies on temporal ventriloquism have focused on audio and visual modes, there are a small number of studies that have investigated the effect using tactile and sensory stimuli. Keetels and Vroomen (2008) investigated whether motor-sensory asynchronies would lead to a long-term recalibration. In other words, rather than probing for *immediate* temporal shifts in sensory signals, such as the attraction of a visual towards an auditory stimulus, the authors measured if exposure to motor-sensory asynchronies resulted in after-effects. Their experiment consisted of two phases - an exposure phase and a test phase. During the exposure phase participants experienced temporal asynchronies with lag times of -100, 0, and 100 ms between a visual stimulus and tactile vibration. During the test phase participants were briefly re-exposed to the same time lag as in the exposure phase, then given a test pair: a tactile vibration and visual stimulus were presented with a range of temporal asynchronies, and participants had to judge which stimulus was presented first. The authors found that the point of subjective simultaneity (PSS) was shifted in the direction of the exposure lag; when the tactile stimulus was presented first participants reported more visual-first pairings, while presenting the visual stimulus first resulted in reports of more tactile-first pairings. Put simply, the brain attempts to

compensate for motor-sensory lags by realigning events, such that when the tactile event occurred first, the subsequent visual stimulus was shifted backward in time towards the tactile. Interestingly, the results also found that when both stimulus pairs were actually presented simultaneously, the visual had to precede the tactile by 8.6 ms to be perceived as simultaneous. This is largely in line with studies that employ audiovisual stimuli, in that visual information is processed slower than auditory and (possibly) tactile stimuli (Keetels & Vroomen, 2012).

So what does this evidence suggest so far? At a minimum, it implies that the alignment of temporal asynchronies are not limited to audiovisual domains but also involve the tactile modality. One question that naturally poses itself is would the brain adapt to temporal asynchronies when motor action is involved? Keetels and Vroomen (2008) employed tactile stimulation, rather than motor action (i.e., key presses). How then, would temporal asynchronies be perceived when motor action (not stimulation) is involved? One study that addresses this question is Stetson, Cui, Montague, and Eagleman (2006). Participants in their study completed control (baseline) blocks and experimental (injected delay blocks). Control and experimental blocks shared a similar set-up: on 60% of trials a flash appeared after a fixed delay following a key press (35 ms in baseline, 135 ms in injected delay). In the other 40% of trials unexpected flashes appeared randomly 150 ms before or after the key press. Participants reported whether the flash occurred before or after the key press. The key finding was that in the injected delay blocks a flash appearing 44 ms, for example, after the key press would have been perceived as occurring *before* the key press, while in baseline trials it would be reported as occurring *after* the key press. This is because when participants adapt to a relatively long key press-flash delay, an unexpected flash after a short delay is perceived sooner. Indeed, the striking feature is that the flash is

perceived before the key press that *caused* it. However, on 40% of trials the flash was scheduled to randomly appear either before or after the key press, which might have reduced the causal link between the key press and flash; because causes precede outcomes, the presentation of flashes *before* the key presses might have led participants to believe that key press did not cause the flash to appear. To what extent then is causal belief a factor in determining the direction of the sensory shift? Given that causes precede their outcomes, would a causality-induced stimulus be more likely to be perceived after its cause, relative to an unrelated key press-stimulus pairing?

Desantis, Waszak, Moutsopoulou and Haggard (2016) investigated the effect of causal belief on the temporal order of actions and their outcome. Their study involved learning action-outcome associations and then judging the temporal order of events when the outcomes were either congruent or incongruent with the associations. In an association phase, participants learned the relation between key presses and the direction of random dot kinematograms (RDKs), such that pressing a certain key triggered upward motion and another key triggered downward motion. In a test phase, RDKs were presented for a range of lag times before or after the key press, and participants were asked to judge whether the coherent motion of the RDKs occurred before or after their key press. Importantly, the participants were explicitly told about the causal relationship, such that a coherent dot motion that occurred before the key press was computer generated, while coherent motion after pressing the key was caused by the participant (in reality the motion was random, regardless of whether it occurred before or after. The authors simply used a rolling average of participants' key press times and scheduled dot motion to occur sometime before or after). Furthermore, participants also believed that the key press-motion associations still applied, such that pressing a certain key triggered an upward direction. This was so that any

effect of congruence can be attributed to the key press-motion mappings learned in the association phase (again, the motion of the RDKs was entirely random and did not respect the learned associations). The results can be summarised thus: when the coherent motion was congruent with the previously learned association, the motion was reported to occur more often *after* the key press. However, when participants were told about the true causal relationship between key presses and RDK motion, specifically that motion was not contingent on the previously learned association, then no congruency effect was obtained. It is important to note that coherent motion was reported to occur after the actual key press time regardless of causal belief (in line with classical causal principles of cause preceding outcome). Yet, when motion was in line with a causal belief then the subjective shift was greater, such that *more* coherent motions were reported after the key press, relative to incongruent motions. Thus, although this Desantis et al. study is not a study of recalibration *per se*, it does evidence the role of causality in the perception of events.

Thus far, an adequate summary might be as follows. Input from different sensory streams is processed by the brain with different latencies. One way in which the brain unites the different inputs into a single event is by sensory realignment. With audiovisual stimuli there is an auditory dominance, such that the arrival times of visual events are shifted towards the time of audition. Recalibration also occurs with tactile stimuli, particularly if there is a causal connection between tactile and sensory events; outcomes are more likely to be perceived after their purported cause, relative to events not causally triggered.

How can sensory realignment explain temporal binding? According to David Eagleman (Eagleman & Holcombe, 2002; Parsons, Novich & Eagleman, 2013), binding can be explained by a process of adaptation and recalibration. Individuals usually have a prior belief that outcomes follow causes without delay, such as clicking one's fingers causes an

immediate 'clicking' sound. If, however, there is a delay between action and outcome, then the expectation of an immediate outcome is violated. The timing of events are then recalibrated to compensate for this delay. The Stetson et al., (2006) study is a notable example of this sensory recalibration hypothesis. When participants had adapted to a delay between an action and outcome, the presentation of the outcome *immediately* following the action resulted in a perceived reversal of action-outcome events; the outcome was perceived to occur *before* the action. Sensory recalibration then accounts for temporal binding with a process of adapting to motor-sensory delays, bringing action and outcome closer together in time.

2.1.12. Sensory recalibration, time perception, and the current experiments

Proponents of sensory recalibration (Eagleman & Holcombe, 2002; Parsons et al., 2013) suggest that temporal binding cannot be effected by a slower pacemaker. One of the reasons for this is that after adapting to an action-outcome delay, the presentation of an outcome immediately following the action results in a perceived reversal: outcomes are perceived before their actions. Parsons et al. note that the internal clock model cannot account for this subjective interval that is negative in duration. Another reason is that in Parsons et al., when two sensory outcomes were presented (a beep and flash), binding only occurred between action and one of the sensory outcomes. In other words, the beep and flash were able to subjectively shift independently of each other. If an internal clock model effected binding, then both sensory outcomes, argues Parsons et al., would have shifted by an equal amount. Binding then, occurs due to sensory realignment, and not to changes in the rate of an internal clock (according to this perspective).

There are, however, two reasons not to discount internal clock models. First, motor-sensory recalibration occurs only in conditions that involve motor action; it cannot account

for binding between mechanical devices that trigger outcomes (e.g., Buehner, 2012). Thus, this finding is difficult to explain without implicating internal clock models. Second, there is precedence in the literature for multiple clocks (Buhusi & Meck, 2009; Klapproth, 2011), such that multiple overlapping stimuli are timed by separate clocks rather than a single timing system. Pariyadath and Eagleman (2007; see also Eagleman, 2008) find that increases in subjective duration for one stimulus, do not results in orthogonal increases in an additional, simultaneously presented stimulus. Therefore, the results of Parsons et al. (2013) can be explained by the involvement of multiple clocks, each timing a specific pair of action-outcome stimuli.

Internal clock models then, are not only difficult to discount (as in point two, above), but are indispensable in certain instances (as in point one). However, given the paucity and inconsistency of evidence for pacemaker slowing in temporal binding, further empirical work is necessary. The current experiments are therefore, an application of the internal clock model to temporal binding. Specifically, I ask one key question: is temporal binding effected by a slower pacemaker?

2.1.13. Scope

Given that previous research has shown that temporal binding is rooted in causality (Buehner, 2015, 2012; Buehner & Humphreys, 2009; Moore, Lagnado, Deal, & Haggard, 2009), occurs across various stimulus modalities and tasks (Humphreys & Buehner, 2009,2010; Haering & Kiesel, 2014,2015), and does not merely reflect processes related to motor preparation and execution, I focused my efforts on a simple causal versus non-causal interval distinction: In the experiments reported in this thesis, causal intervals were delineated by a participant's key press, which always caused a sensory outcome after a delay, while noncausal intervals were delineated by two sequential sensory outcomes. This

manner of operationalizing causal and noncausal intervals presented the most efficient method of manipulating causality. The key difference is that in causal intervals, only a key press triggers an outcome, hence the action is causally related to the outcome stimulus. In noncausal trials, the interval markers (the visual change and the outcome stimulus) are not causally related but are merely associated - the outcome *always* follows the visual change. Thus, causal intervals were always delineated by a key press and sensory outcome, while noncausal intervals were delineated by two sensory stimuli (for information this causal versus noncausal distinction, see the introduction to Chapter 3, on the following page).

I began by investigating whether the contraction of time in temporal binding is served by the slowing of a general pacemaker, or instead, by the slowing of clock processes specific to the cause (action)-effect interval (Experiments 1 - 4 in Chapter 3). The results were in accordance with Pariyadath and Eagleman (2007), with the finding that binding is served by a specific clock process (admittedly, the results also fit with a sensory realignment process, but I defer this discussion until Chapter 6). In order to understand the nature of this specific process, I conducted experiments with a regression procedure to disentangle switch latencies and changes in pacemaker rate (Experiments 5 - 8 in Chapter 4). In accordance with the majority of temporal illusions, I found that binding is effected by differences in pacemaker rate. Therefore, as a final and stronger measure of pacemaker slowing, I conducted temporal discrimination tasks, in which participants discriminate causal and noncausal intervals (Experiments 9 - 10 in Chapter 5). The results of these latter experiments found higher thresholds when discriminating causal, versus noncausal intervals, as expected from a slower clock. The present body of work then, finds empirical support for pacemaker rate changes in temporal binding.

3. Chapter 3: Investigating a general clock slowing account of temporal binding using the embedded interval estimation procedure

I developed a new procedure - the *embedded interval estimation procedure* - to address whether temporal binding is effected by a slower clock rate. Given that previous research has shown that temporal binding is rooted in causality (Buehner, 2015, 2012; Buehner & Humphreys, 2009; Moore, Lagnado, Deal & Haggard, 2009), occurs across various stimulus modalities and tasks (Humphreys & Buehner, 2009,2010; Haering & Kiesel, 2014,2015), and does not merely reflect processes related to motor preparation and execution, I focused my efforts on a simple causal versus non-causal interval distinction: Causal intervals were delineated by a participant's key press, which always caused a sensory outcome after a delay, while noncausal intervals were delineated by two sequential sensory outcomes. This manner of operationalizing causal and noncausal intervals presented the most efficient method of manipulating causality. Whilst the events of both conditions appear to be associated, the key difference is that in causal intervals, only a key press triggers an outcome, i.e., the outcome occurs only when a key is pressed and does not occur in the absence of a key press. In noncausal trials, by contrast, the interval markers (the visual change and the outcome stimulus) are not causally related but are merely associated - the outcome *always* follows the visual change[1]. Critically, I embedded an additional event into

[1] Although the physical difference between conditions is whether or not intentional actions occur, I prefer to use the terms 'causal' and 'noncausal', given the evidence for a causal root to temporal binding: intentional actions (i.e., key presses) that are associated with outcomes, do not yield a binding effect if there is an a priori belief that such outcomes will occur even in the absence of such actions (Buehner & Humphreys, 2009). In other words, association is not sufficient for binding to occur; a more meaningful causal link is required, such that the trigger (e.g., the key press) is causally linked to the outcome, and not merely a stimulus in a sequence of associated stimuli.

certain trials at different points; sometimes this event occurred before the interval, in other trials it occurred during the interval, and in others not at all.

One may at this point ask how causality could possibly influence pacemaker function, given that I can only be sure of causality *after* the outcome has occurred and not at the time of action (there is a rich literature in in cognitive science on this, based on David Hume's (1888/1739) empiricism; for an overview see Cheng & Buehner, 2012). In other words, the motor system only obtains information about the successful completion of a causal action once the outcome occurs, at which time the switch closes, thereby stopping pulses to accumulate. However, Moore and Haggard (2008), using the Libet clock method described earlier (Libet et al, 1983), demonstrated that temporal binding occurs not only due to retrospective awareness of causality, but also due to the prediction of causal relations. They contrasted a predictable condition (a key press produced a tone on 75% of trials), with an unpredictable condition (a key press produced a tone on 50% of trials). This allowed them to study the extent of perceptual shifts when a causal action was followed by its outcome as well as when it failed to do so. On trials where the action did not produce an outcome, Moore and Haggard found a larger subjective shift of actions towards (expected, but absent) tones in the 75% compared to 50% conditions. In other words, when participants expected an outcome, they experienced binding even when in fact no outcome occurred. Furthermore, Moore, Lagnado, Deal, and Haggard (2009) showed that the extent of binding varied as a function of the contingency, in line with a causal theory of binding. In the experiments reported in this chapter the action-outcome contingency was set to 100% (i.e., perfect predictability). It is therefore reasonable to assume that participants would show strong binding effects based on cause-effect predictability. This, in turn, means that pacemaker rate could, at least in principle, be affected by this manipulation of causality.

In order to provide a comprehensive test of my hypotheses and to increase the generalizability of the results, I conducted four experiments that factorially combined stimulus modality (auditory, visual) with event type (outcome, embedded event) as follows: In Experiments 1 and 2, the outcome event marking the end of the overall interval was visual (a flash), while in Experiments 3 and 4, the outcome was an auditory click. Furthermore, the embedded event was auditory (a constant tone of variable duration) in Experiments 1 and 3, and visual (a flash of a polygon of variable duration) in Experiments 2 and 4. In other words, in Experiments 1 and 4, the embedded event and the outcome event marking the end of the interval were of different modalities, while they were of the same modality in Experiments 2 and 3 (see Table 3.1).

Table 3.1

Modality of embedded events and outcomes for each experiment.

Experiment	Embedded Event	Outcome
1		
2		
3		‘Click’
4		‘Click’

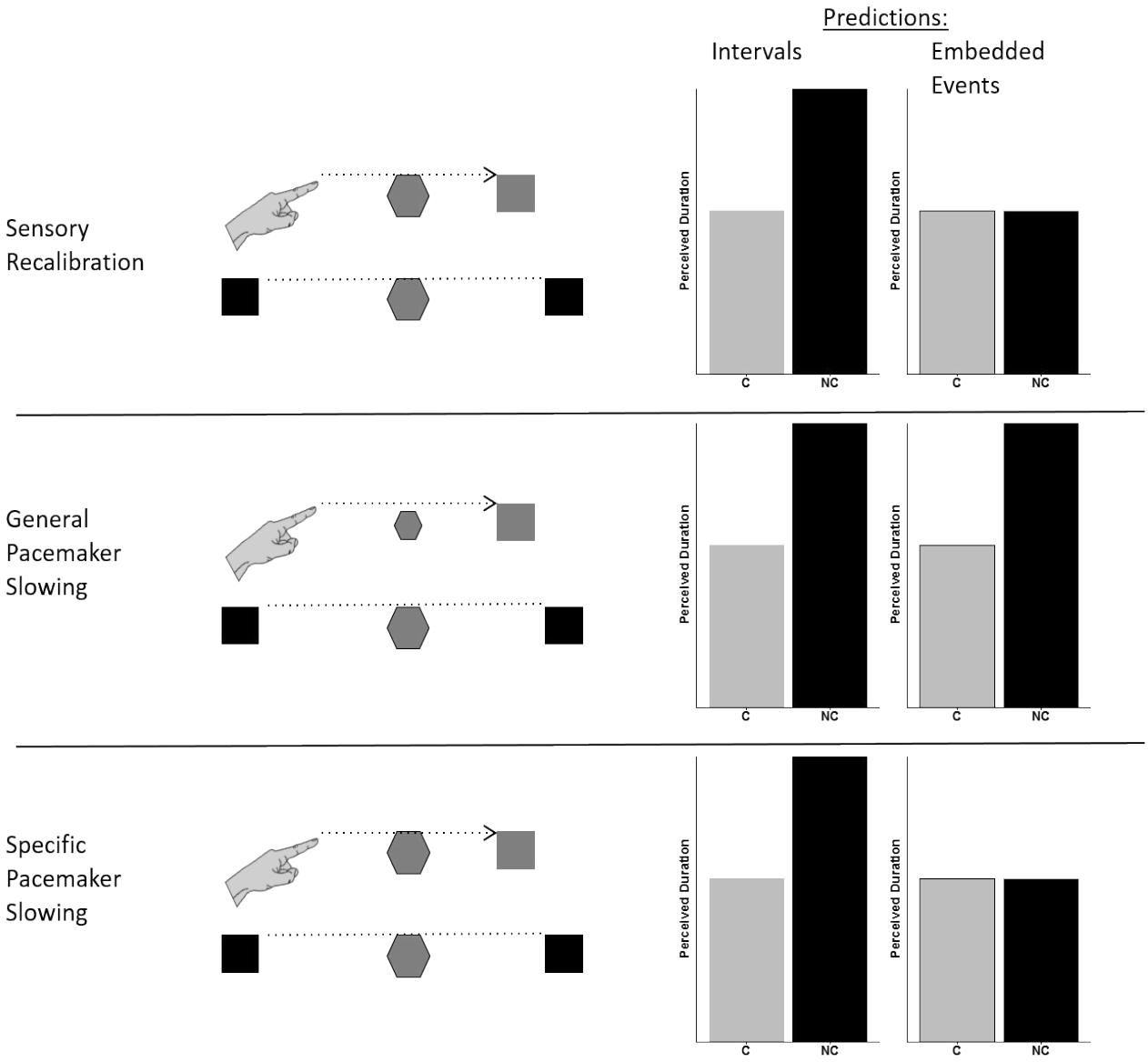

Figure 3.1. Predictions for trials with an embedded event *during* the interval. Interval judgments: I predicted a temporal binding effect regardless of theoretical approach. Embedded event judgments: Only a general pacemaker slowing will result in shorter perceived embedded events during causal intervals. C: Causal; NC: Non-Causal.

In all experiments, participants had to estimate the duration of either the overall interval, or the embedded event. If causality-induced clock slowing is general, then embedded event estimates should be shorter for events embedded into causal intervals, compared to events embedded into noncausal intervals. Alternatively, if clock slowing is a binding-specific process, then estimates for embedded events should not differ between

causal and noncausal conditions. This would necessitate multiple clocks operating in parallel, rather than a single clock system: One clock might time the interval and another the embedded event simultaneously. Likewise, if temporal binding does not implicate any changes to clock speed, then there should also be no differences in the estimates for embedded events (see Figure 3.1 for a schematic of these predictions). Because there are modality-specific differences in the timing of auditory versus visual events (Wearden et al. 1998), I decided to factorially combine stimulus modality with event type in a 2 x 2 design to ensure that whatever effects I might find are not confined to a particular modality.

A key concern in developing the *embedded interval estimation procedure* was to ensure that the embedded stimulus was independent of the main interval. More specifically, embedded events had to be perceived as causally unrelated to key presses. This was to ensure that no temporal binding between key presses and embedded events occurred; if this were the case, then embedded events and outcomes might be perceived as a sequence of action-related outcomes. This, in turn, would make any effects of clock-slowing during the key press-interval, difficult to ascertain. To make it clear that the embedded event was independent of the action, I scheduled one-third of trials to contain an embedded event before the key press, one-third after the key press, and one-third to contain no embedded event. To achieve this, I used an algorithm that predicted a participant's key press time for each causal trial, and scheduled delivery of the embedded event either before or after this predicted time. In noncausal trials I employed stimulus delivery times recorded from participants in a pilot experiment using the same algorithm. The pilot experiment consisted of 24 participants using a procedure similar to that in the present experiments. The only difference was that causal blocks were experienced first, with the values for different stimuli duration (interval duration, time of key press, time of embedded event, and so on) replayed

in noncausal blocks. This pilot experiment used the same prediction algorithm as the current experiments. For the current experiment, this procedure ensured noncausal trials mirrored causal trials as closely as possible, whilst also allowing for the counterbalancing of causal and noncausal blocks.

Participants were asked to estimate the duration of either the interval or the embedded event. To make sure that they focus on all aspects of the task, participants were not told which event they had to estimate until the end of each trial.

3.1. Experiments 1 and 2

3.1.1. Method

3.1.1.1. Participants

Thirty-five Students (32 female, 3 male, M_{age} = 19.9 years, age range: 17-40) of Cardiff University participated in Experiment 1, and 34 (32 female, 2 male, M_{age} = 18.9 years, age range: 18-22) participants in Experiments 2. Students were provided with course credits or £5 payment.

3.1.1.2. Apparatus and stimuli

The experiment was implemented in Psychopy (Peirce, 2007) on PCs connected to 19" monitor with resolution of 1280 x 1024 pixels. The embedded event was a tone (523.25 Hz) in Experiment 1 and a yellow polygon (vertices bounded by a 270 x 210 pixel rectangle) in Experiment 2, of either 300 or 500 ms duration. In both experiments a 1000 ms presentation of a red square (400 pixels^2) served as the outcome on causal trials and marked the end of noncausal intervals. A black square (400 pixels^2) was presented at the beginning of each trial in all experiments. All visual stimuli were presented centrally on the screen.

3.1.1.3. Design and procedure

Three factors were employed in the study: Trial Type consisted of two levels (causal, noncausal), Embedded Event Location of three levels (before interval, during interval, no event) and Embedded Event Duration of two levels (300, 500 ms). Dependent variables (measured on separate trials) were estimates of overall interval and embedded event duration.

Figure 3.2 shows the trial structure of Experiment 1. Causal trials began with the black square on the screen. Participants pressed a key at a time of their choice. This led to the immediate disappearance of the black square, and triggered the red square after a random interval (range 700-1300 ms). Noncausal trials also began with the black square, which remained on the screen according to a predetermined time derived from the causal trials of previous pilot participants as explained below; following the disappearance of the black square, the red square appeared after a random interval of the same 700-1300ms range. Both types of trials were presented in blocks of 30 trials, 10 of which were scheduled to contain an embedded stimulus *during* the relevant interval, 10 to contain an embedded stimulus *before* the interval (i.e. before the participant pressed the key or before the black square disappeared on its own), and 10 trials in which no embedded stimulus was scheduled. At the end of each trial, participants were prompted to estimate *either* the duration of the embedded stimulus, *or* the duration of the key press – outcome interval (on causal trials) or the duration between the disappearance of the black square and the appearance of the red square (on noncausal trials).

To schedule delivery of the embedded stimulus, an algorithm was used to predict participants' key press time, with the embedded event scheduled at a random time (range 50 - 400 ms) before or after the predicted key press. Based on pilot data, I set the algorithm

to begin with a prediction of 800 ms on the first trial, and implemented a cumulative average based on key press times up to the first five trials. Thereafter, a rolling average calculated over the last five trials was employed. Values shorter than 400 ms or longer than 3000 ms were not considered for the averages.

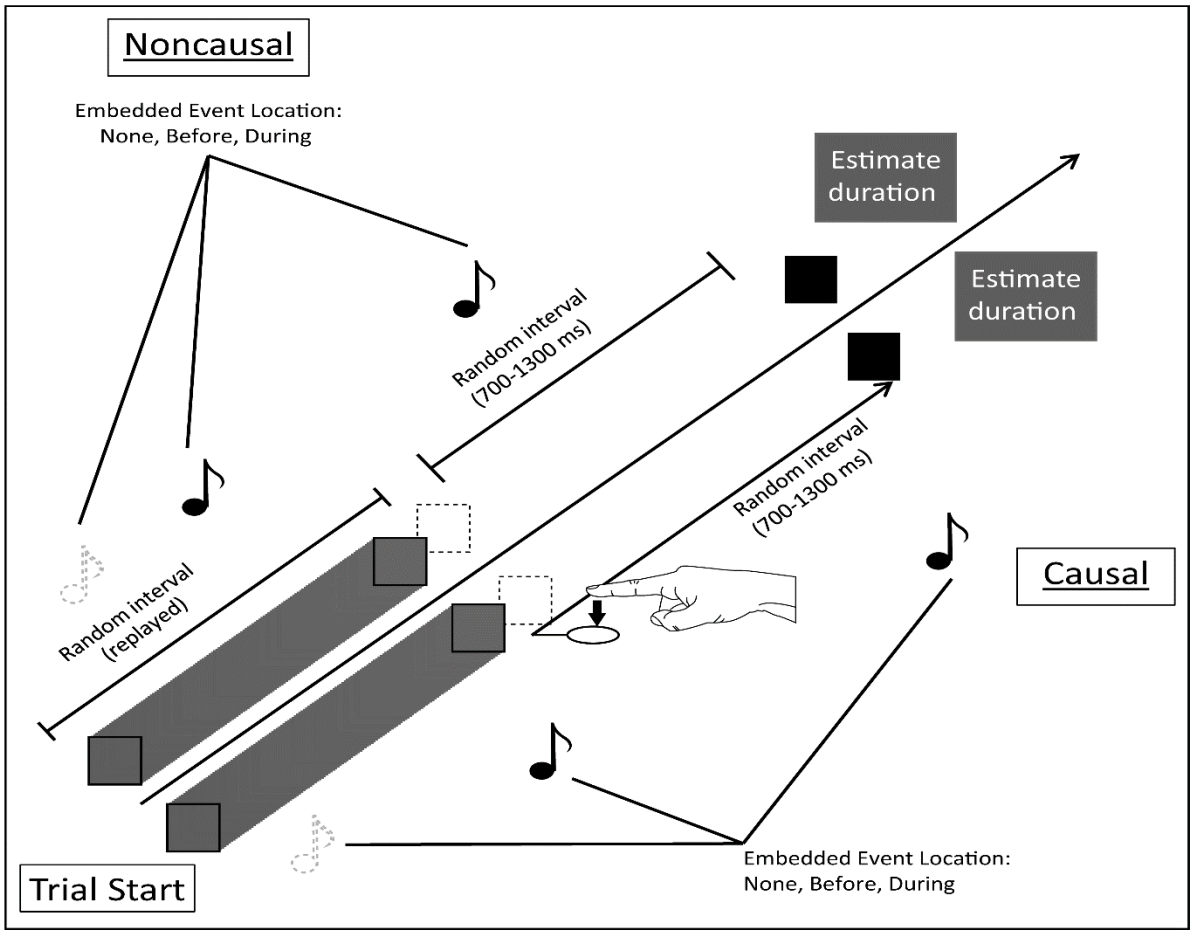

Figure 3.2. Trial structure of Experiment 1. In causal trials a black square was displayed until participants pressed a key, which led to the immediate disappearance of the square and triggered an outcome stimulus after a random interval (700 - 1300 ms). In noncausal trials a black square was displayed for a random time (replayed from a pilot version of the experiment) before disappearing, followed by an outcome stimulus after a random interval of the same 700 - 1300 ms range. Additionally, in both trial types an event was embedded before or during the interval, or in some trials no embedded event was scheduled. Participants estimated the duration of either the interval *or* embedded event. In all experiments a black square always began the display, while the modality of outcomes and embedded events varied factorially between experiments. In Experiments 1 and 2 the outcome was a red square, and in Experiments 3 and 4 an auditory click sound. Meanwhile, the embedded event was a tone in Experiments 1 and 3, and a polygon in Experiments 2 and 4.

At the beginning of each causal trial, the computer thus determined the length of the to-be-experienced key press – outcome interval (from a random range of 700-1300 ms), as well as whether and when it was to contain an embedded event. If an embedded event was scheduled, its duration could be either 300 or 500 ms (see design specifics below). For the participant, the trial began with the display of the black square, which remained on the screen until he or she pressed the Z key, which led to the immediate disappearing of the black square and triggered the appearance of the red square after the scheduled interval. Depending on the schedule, the embedded event occurred between 50-400 ms before or after the *predicted* time of the key press, or not at all. Following the display of the outcome (red square), participants were prompted to estimate either the duration of the action-outcome interval (on a scale from 0 – 2000 ms) or the duration of the embedded event (on a scale from 0 – 1000 ms). Participants estimated the duration by entering a numerical estimate in milliseconds. The screen then blanked for a random duration (1200 - 2000 ms) before the next trial.

For noncausal trials I adopted an analogous procedure. I replayed values from causal trials of a pilot version of the experiment, where I recorded the time of participants' key presses, the durations of the intervals, as well as the positions and durations of the embedded events (based on the same prediction algorithm as described above). In the current experiment, for each participant, I randomly selected a pilot participant's data file and replayed its values in noncausal trials, using the recorded key press time to schedule the disappearance of the black square. For example, if a pilot participant pressed the key 900 ms into the trial, experienced an interval of 1100 ms, with a 300 ms embedded event presented 800 ms after his or her key press, a corresponding noncausal trial in the current experiment would display the black square for 900 ms at the start of the trial, followed by the red square

after 1100 ms; in addition, an embedded event of 300 ms duration would be presented 800 ms after the disappearance of the black square. Note that the algorithm cannot perfectly predict a participant's key press, and that consequently the number of trials where the embedded event was experienced before or after the action will fluctuate between participants. Table 3.2 shows that the prediction algorithm achieved approximately equal distribution of embedded events before and during intervals. For example, Table 1 shows that for interval judgments in Experiment 1, 50.16% of trials featured an embedded event *before* the interval, out of a target of 50%. Likewise, 44.35% of trials featured an embedded event *during* the interval (also out of an intended target of 50%). Thus, there were no systematic deviations from the intended 50-50 balance, and no deviation was larger than 9%.

Table 3.2

Mean percentage of trials on which an embedded event was experienced before and during the overall interval, for interval and embedded event judgment trials in Experiments 1 and 2 (standard deviation in parenthesis).

	Experiment 1				Experiment 2			
	Causal		Noncausal		Causal		Noncausal	
Judgment	BI	DI	BI	DI	BI	DI	BI	DI
Intervals	50.16	44.35	41.94	53.63	47.66	45.55	44.38	51.72
	(11.09)	(12.63)	(6.28)	(6.35)	(9.71)	(11.55)	(7.04)	(6.64)
Embedded	50.40	45.16	46.05	49.27	51.41	41.95	46.48	48.67
Event	(10.02)	(11.76)	(5.62)	(6.23)	(10.34)	(12.28)	(5.53)	(6.57)

Note: BI, DI = Embedded Event occurred before/during the interval. Percentages do not sum to 100 because a mean range of 3.91 - 6.80% of trials contained embedded events that began during or after the outcome, and are thus not included. 2 x 2 (Trial Type [causal, noncausal] x Embedded Event Location [before, during]) ANOVAs in which the dependent variable was the proportion of trials with an embedded event, found no significant effects (*p*s>.05), with one exception: Interval judgment trials in Experiment 2 had more events in noncausal than causal trials (means of 48.05 and 46.60 %, collapsed across BI and DI trials, respectively), $F(1, 31) = 7.79$, $p < .01$. I attribute this to random fluctuation.

To optimize the experience on noncausal trials, I screened previous participants' stimulus patterns (from the pilot version) and excluded those where embedded event timings deviated by more then 30% from the schedule (e.g. where the balance of embedded events occurring before and during the relevant interval deviated from the scheduled 50-50% balance; fewer than 35% or greater than 65% resulted in exclusion. Each causal and noncausal block consisted of 30 trials prompting for an interval and 30 requesting an embedded event duration judgment, presented in random order. For each judgment type, there were 10 trials with the embedded event scheduled before, 10 with it scheduled during the interval, and 10 trials with no event. The duration of embedded events was either 300 or 500 ms, with both durations occurring equally often across trial and judgment types. Ten trials in each block prompted for an embedded event judgment when in fact no embedded event had occurred. These served as catch trials, and participants were instructed beforehand to enter an X on such trials.

Each participant worked through two causal and noncausal blocks in an alternating sequence, with the beginning of the sequence (causal, noncausal) counterbalanced. Participants were tested in groups of 10 - 15 and the experiments took about 45 minutes.

3.1.2. Results

3.1.2.1. Data analysis

Data were screened based on catch trials. Participants who failed to correctly recognize more than 30% of catch trials in at least one condition (Four participants in Experiment 1, two in Experiment 2) were not considered for analysis. In addition, while Table 3.2 and ANOVAs show that generally, the distribution of before and during interval embedded events did not differ between causal and noncausal trials, inspection of individual data revealed that certain participants experienced a distribution of causal before and

during trials with greater than 30% deviance, i.e., some participants experienced fewer than 35% or greater than 65% of before and during trials, rather than the intended 50 - 50 balance. Twenty percent (Experiment 1) and 35% (Experiment 2) of participants fell into this category. Whilst I considered removing these participants from analyses, doing so does not change the pattern of results in any of the four experiments (all interval judgment *P*s < .05, all embedded event judgment *P*s > .05), so I decided to report analyses based on the entire valid sample (see Appendix C, however, for analysis with these participants removed).

Data for interval judgments were classified into three categories: trials where an embedded event occurred before the interval, during interval, or not at all. Data for embedded event judgments were classified into two categories: trials where the embedded event occurred before the interval and trials where it occurred during the interval. Note that this classification is based on *actual* rather than *scheduled* embedded event location, and classification is based on the *start* of the embedded event. This meant that some events began *before* but overlapped *into* the interval. Likewise, certain events began *during* the interval but overlapped *into* the outcome stimulus (in both cases this occurred on an average of 14.8% and 13.7% of trials in Experiments 1 and 2, respectively). The analyses reported below included such overlapping trials; excluding them does not change the pattern of results (see Appendix A). All analyses were computed with respect to judgment errors[2], defined as the difference between the estimated and the actual value, where *under*estimation is conveyed by negative values and *over*estimation by positive values. Additionally, analyses with Embedded Event Duration included as a factor yielded the same results as those reported below (all interval judgment *p*s < .05; embedded event judgment

[2] See Appendix B for scatter plots showing actual versus estimated durations, for interval and embedded event judgments.

[Experiments 1,2 and 4] *p*s > .05 and Experiment 4 *p* = .05). I therefore conducted the analyses by collapsing across Embedded Event Duration.

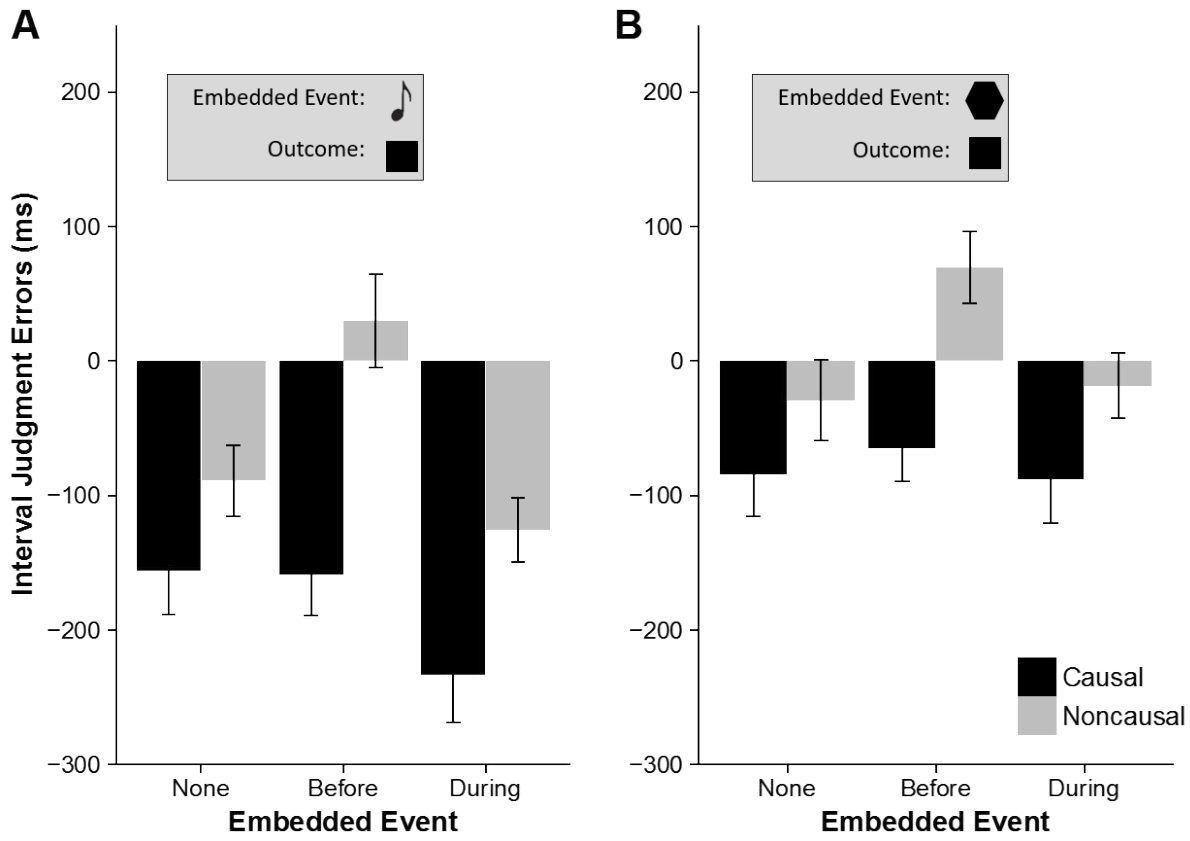

Figure 3.3. (A) Experiment 1. (B) Experiment 2. Mean interval judgment errors from causal and noncausal trials broken down by embedded event location. Error bars represent within-subjects confidence intervals (Cousineau, 2005; Morey, 2008).

3.1.2.2. Interval Judgments: Experiment 1

Figure 3.3 (A) shows that intervals were generally underestimated, with causal intervals underestimated to a greater extent, replicating the typical binding effect. An exception to this pattern of general underestimation is the result from noncausal intervals with an embedded event before the interval, which shows overestimation. Figure 3.3 also shows that intervals with an event presented during the interval were underestimated more than intervals with no event and an event before. I conducted an analysis of variance

(ANOVA) with Trial Type (causal, noncausal) and Embedded Event Location (none, before, during) as factors. This analysis found a significant effect of Trial Type, $F(1, 30) = 24.97$, $p < .001$, partial $\eta^2 = .45$, and a significant effect of Embedded Event Location, $F(2, 60) = 17.58$, $p < .001$, partial $\eta^2 = .37$. The Trial Type x Embedded Event Location interaction was also significant, $F(2, 60) = 6.33$, $p < .01$, partial $\eta^2 = .17$. Simple effects analysis found significant differences between Trial Type with no embedded events ($p < .05$), for events before ($p < .001$) and for events during the interval ($p < .01$).

3.1.2.3. Interval Judgments: Experiment 2

Figure 3.3 (B) shows that intervals again were generally underestimated, with greater underestimation for causal intervals. A deviation from this pattern, analogous to Experiment 1, concerns interval judgments from noncausal trials where an embedded event occurred before the interval, which seemed to reflect overestimation. Above all, Figure 3.3 (B) shows a clear binding effect regardless of embedded event location. ANOVA supports these findings, with a significant effect of Trial Type, $F(1, 31) = 13.03$, $p < .01$, partial $\eta^2 = .30$. The effect of Embedded Event Location, $F(2, 62) = 7.14$, $p < .01$, partial $\eta^2 = .19$, as well as the Trial Type x Embedded Event Location interaction, $F(2, 62) = 3.75$, $p < .05$, partial $\eta^2 = .11$, were also significant. Simple effects analysis found no difference between Trial Type with no embedded events ($p = .10$), but did for events embedded before ($p < .001$), and during the interval ($p < .05$).

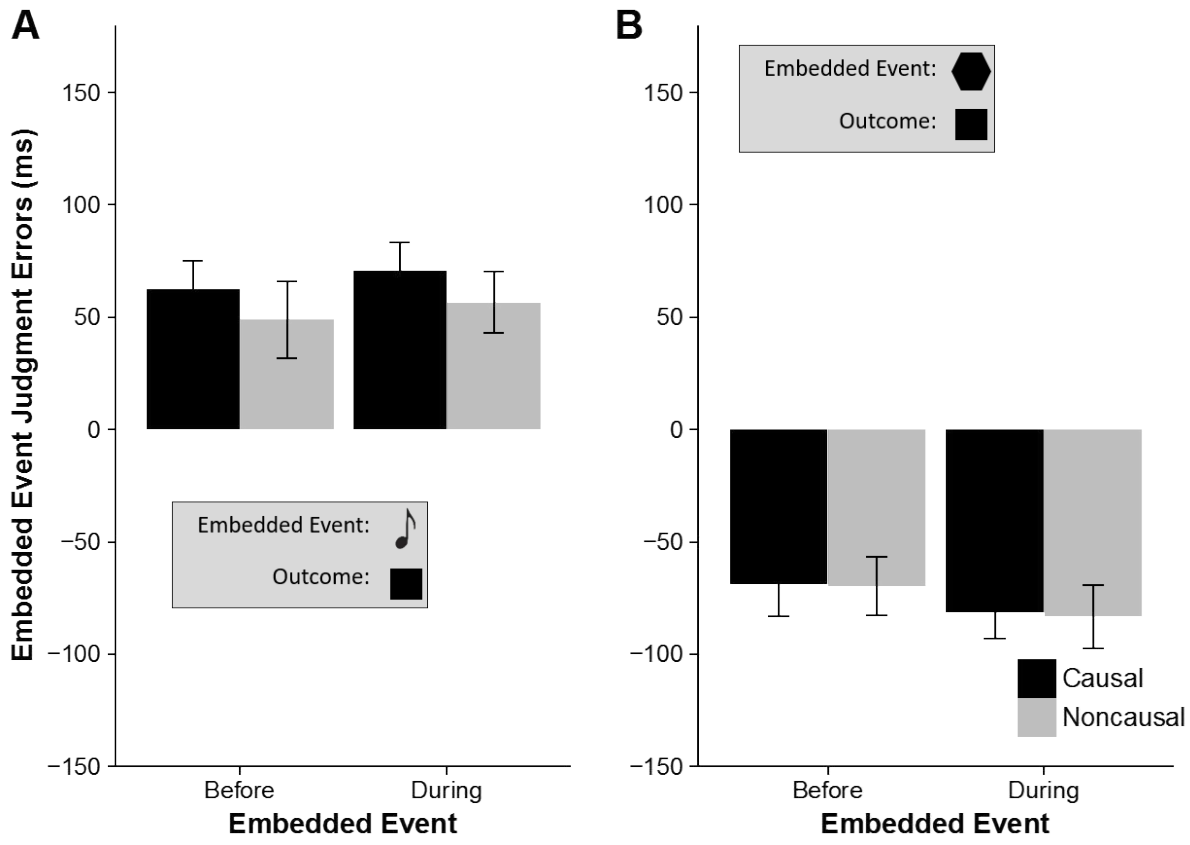

Figure 3.4. (A) Experiment 1. (B) Experiment 2. Mean embedded event judgment errors for events occurring before and during the interval, broken down by trial type. Error bars represent within-subjects confidence intervals.

3.1.2.4. Embedded Event Judgments: Experiment 1

Inspection of Figure 3.4 (A) finds a general overestimation of embedded events, with somewhat greater overestimation for events embedded in causal intervals, contrary to what is predicted from a general clock slowing hypothesis. ANOVA (with Trial Type [causal, noncausal] and Embedded Event Location [before, during]) found no effect of Trial Type, $F(1, 30) = 1.29$, $p = .27$, partial $\eta^2 = .04$, or Embedded Event Location, $F(1, 30) = .83$, $p = .37$, partial $\eta^2 = .03$, nor an interaction, $F(1,30) < .01$, $p = .97$, partial $\eta^2 < .001$. In addition, I conducted a Bayesian analysis to compare the fit of the data under the null against the alternative hypothesis, using the Bayes Factor package in R (Morey & Rouder, 2015). Using the same 2 x 2 (Trial Type [causal, noncausal] x Embedded Event Location [before, during])

design, I compared a restricted model against the null (see Figure 3.5). The restricted model captured my original prediction, in which the means of events embedded during causal intervals would be subjectively shorter than those embedded during noncausal intervals. I did not specify any other order restrictions (i.e., I made no predictions about embedded events that occurred before the interval, either in causal or noncausal trials). This analysis yielded a Bayes factor of 22 in favour of the null versus the alternative restricted model, suggesting that the data are around 22 times more likely to occur under the null than the alternative hypothesis.

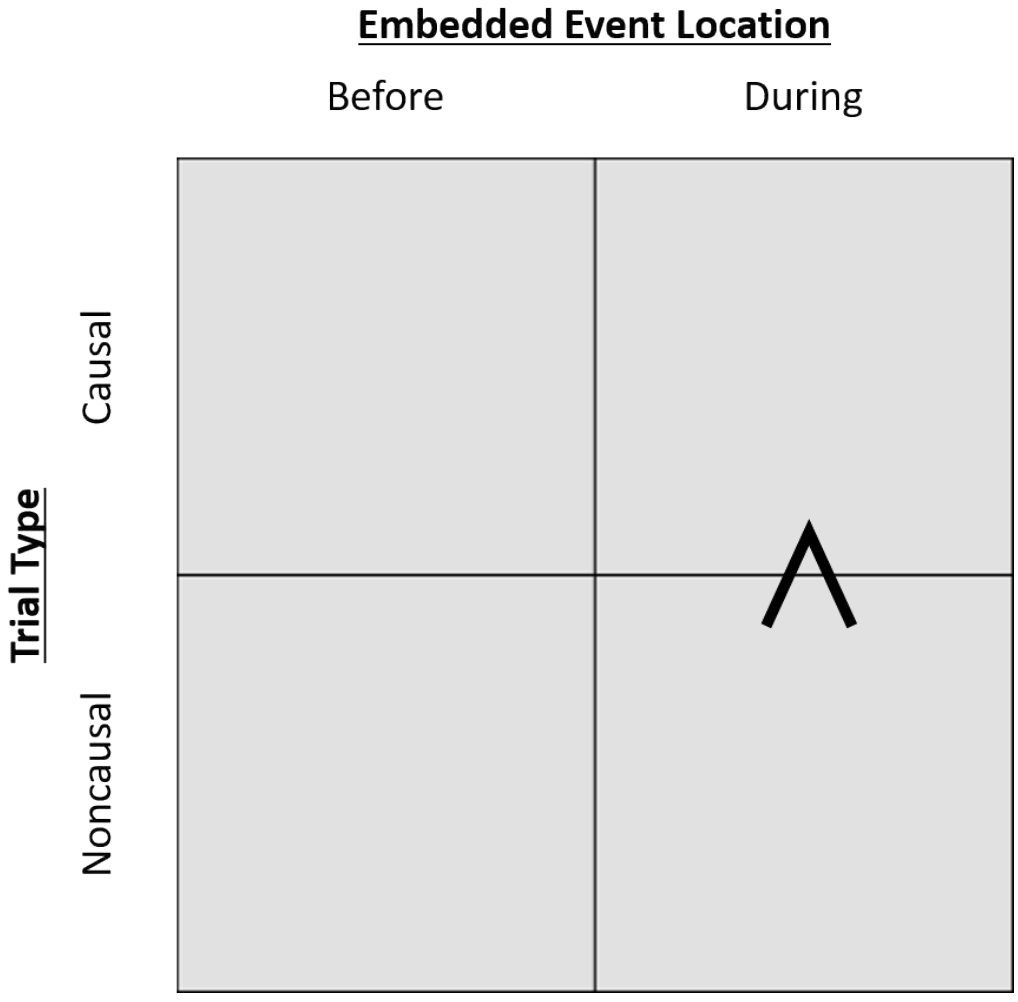

Figure 3.5. The restricted model entered into my Bayesian analysis. I specified a single constraint: events embedded during causal intervals would be judged shorter than events embedded into noncausal intervals. I made no assumptions about any other differences.

3.1.2.5. Embedded Event Judgments: Experiment 2

Figure 3.4 (B) shows that embedded visual events were underestimated, in contrast to embedded auditory events in Experiment 2, which were overestimated. This most likely reflects the typical finding that auditory events are judged longer than visual events (Wearden et al, 1998). More important for my purposes here, though, is whether embedded event judgments varied as a function of trial type. As in Experiment 1, this was not the case. There were no significant effects of Trial Type, $F(1, 31) = .03$, $p = .87$, partial $\eta^2 < .01$, Embedded Event Location, $F(1, 31) = 2.36$, $p = .14$, partial $\eta^2 = .07$, nor a Trial Type x Embedded Event Location interaction, $F(1, 31) = .00$, $p = .96$, partial $\eta^2 < .001$. I also conducted a Bayesian analysis using the same procedure I applied to embedded event judgments in Experiment 1. This yielded a Bayes factor of 12 in favour of the null versus the alternative restricted model.

3.1.3. Discussion

In both experiments I replicated the temporal binding effect. Interval judgments in Experiments 1 and 2 show that causal intervals were perceived as shorter than noncausal intervals, regardless of embedded event location and modality. Interestingly, I also found an interaction, whereby noncausal intervals with embedded events *before* the interval were perceived as longer than intervals with events embedded in other locations. While this is interesting, it is tangential to my main purpose and therefore, I return to this in Chapter 3: Discussion section.

My main concern, however, was with temporal judgments for events embedded into causal and noncausal intervals. I found no difference between causal and noncausal embedded event judgments, both when the event began *before* or *during* the interval, and

irrespective of whether the event was auditory or visual. This is also the case for embedded events that are clearly delineated as before or during, i.e., no overlaps (see Appendix A).

What do these results say about clock slowing in temporal binding? If an internal clock operates at a slower speed during action-outcome intervals relative to noncausal intervals, then embedded events presented during the interval should be judged as shorter in causal than noncausal trials. My results do not bear this out, with no difference in embedded event judgments between trial type or location. Thus, Experiments 1 and 2 provide no evidence for general clock slowing during temporal binding. Instead, it could be that binding-induced clock slowing, rather than being a general cognitive process, selectively affects action-outcome intervals only. However, Experiments 1 and 2 only considered intervals marked by visual events. To glean a better understanding, and in an attempt to test the reliability and robustness of these findings, Experiments 3 and 4 replicate Experiments 1 and 2, but used an auditory event to mark the end of the action-outcome interval.

3.2. Experiments 3 and 4

In Experiments 3 and 4 I replaced the visual outcome with an auditory stimulus. In other respects, the procedure remained the same as in Experiments 1 and 2 (see design specifics below).

3.2.1. Method

3.2.1.1. Participants

Thirty-four students of Cardiff University participated both in Experiment 3 (30 female, 4 male, M_{age} = 19.5 years, age range: 18-22) and 34 another in Experiment 4 (28 female, 6 male, M_{age} = 21.9 years, age range: 17-46), in exchange for course credits or £5 payment.

3.2.1.2. Apparatus and stimuli

The embedded event was a tone (523.25 Hz) in Experiment 3 and a yellow polygon (vertices bounded by a 270 x 210 pixel rectangle) in Experiment 4. In both experiments, a 130 ms click sound served as the outcome in causal trials and marked the end of noncausal intervals. A black square (400 pixels^2) was presented at the beginning of each trial in both experiments. The apparatus was the same as in Experiment 1.

3.2.1.3. Design and procedure

The procedure remained the same as Experiments 1 and 2.

3.2.2. Results

3.2.2.1. Data analysis

As in Experiments 1 and 2, data were screened based on catch trials. This resulted in five participants in Experiment 3, and two in Experiment 4 not considered for analysis. The distributions of before and during interval embedded events for causal and noncausal trials are shown in Table 3.3. Twenty-four percent (Experiment 3) and 26% (Experiment 4) of participants deviated from the intended 50-50 balance. Removing these participants from analyses did not change the pattern of results, and therefore I report the following results with these participants included. However, one participant was removed from analyses in Experiment 3 and another participant in Experiment 4, for not experiencing any causal trials with an event embedded during the interval. Finally, embedded events that began *before* but overlapped *into* the interval were removed and the data subjected to reanalysis. Likewise for events that began *during* but overlapped into the *outcome* (in both cases this occurred on average on 14.8% and 15.5% of trials per participant in Experiments 3 and 4, respectively). The analyses reported below included such overlapping trials; excluding them does not change the pattern of results (see Appendix A).

Table 3.3

Mean percentage of trials on which an embedded event was experienced before and during the overall interval, for interval and embedded event judgment trials in Experiments 3 and 4 (standard deviation in parenthesis).

	Experiment 3				Experiment 4			
	Causal		Noncausal		Causal		Noncausal	
Judgment	BI	DI	BI	DI	BI	DI	BI	DI
Intervals	48.93 (11.37)	46.61 (12.55)	43.93 (4.69)	52.50 (4.56)	50.73 (17.04)	45.40 (17.37)	45.32 (8.34)	51.21 (8.01)
Embedded Event	46.34 (10.33)	49.20 (11.55)	45.00 (5.93)	50.63 (7.03)	53.63 (15.86)	43.06 (15.35)	46.61 (5.11)	48.87 (5.73)

Note: BI, DI = Embedded Event occurred before/during the interval. Percentages do not sum to 100 because a mean range of 3.33 - 4.58% of trials contained embedded events that began during or after the outcome, and are thus not included. 2 x 2 (Trial Type [causal, noncausal] x Embedded Event Location [before, during]) ANOVAs in which the dependent variable was the proportion of trials with an embedded event, found no significant effects (*p*s>.05).

3.2.2.2. Interval Judgments: Experiment 3

Figure 3.6 (A) shows a similar pattern to Experiments 1 and 2. On the whole, intervals were underestimated, with greater underestimation for causal trials. Two exceptions are noncausal trials where an embedded event occurred before the interval, and trials where no event occurred. The former replicates a pattern shown in Experiments 1 and 2, suggesting that the presence of an event presented before the interval subjectively lengthens its duration (I explore reasons for this in Chapter 3: General Discussion). In general, Figure 3.6 shows a clear binding effect, regardless of embedded event location. ANOVA supports these findings, with a significant effect of Trial Type, $F(1, 27) = 44.77$, $p < .001$, partial $\eta^2 = .62$. The effect of Embedded Event Location, $F(2, 54) = 12.56$, $p < .001$, partial $\eta^2 = .32$, and the Trial Type x Embedded Event Location interaction was also significant, $F(2, 54) = 6.70$, $p < .05$, partial $\eta^2 = .20$. Simple effects analysis found significant differences between Trial Type for all levels of Embedded Event Location (all *ps* < .001).

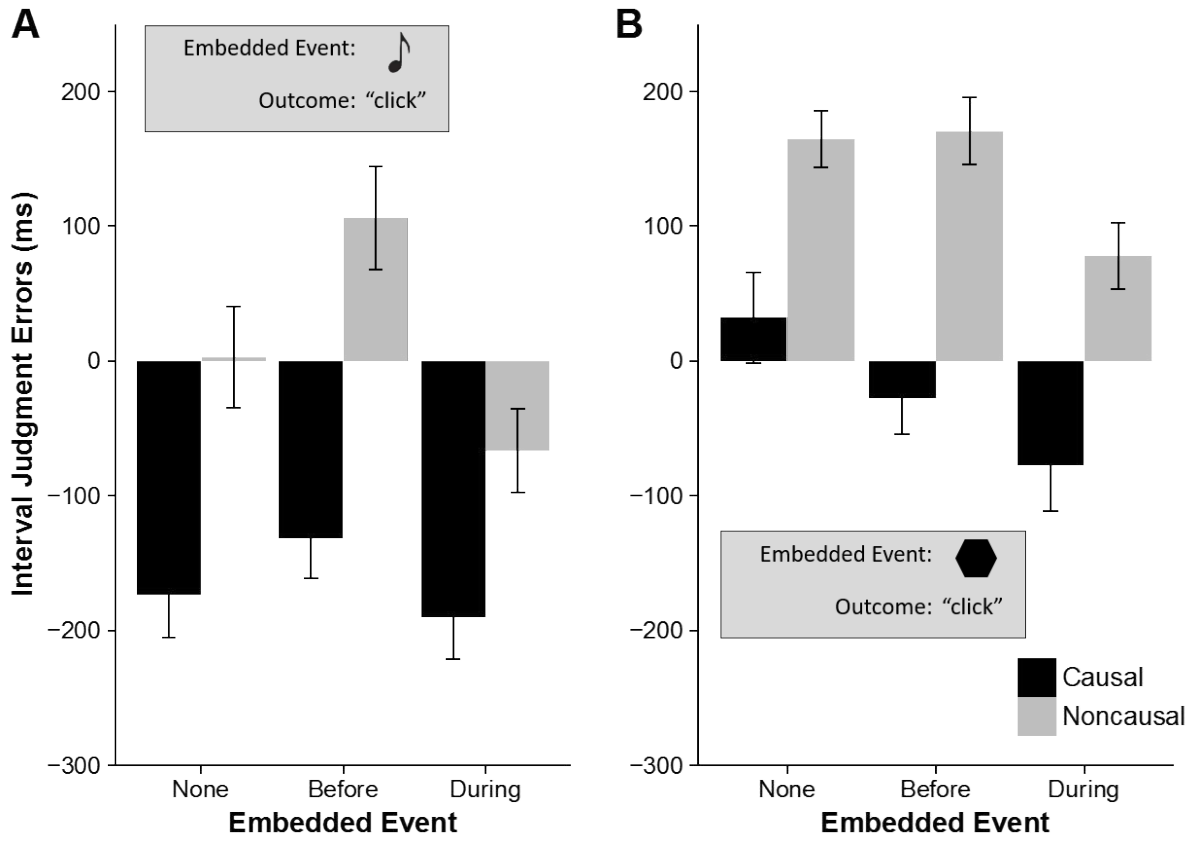

Figure 3.6. (A) Experiment 3. (B) Experiment 4. Mean interval judgment errors from causal and noncausal trials broken down by embedded event location. Error bars represent within-subjects confidence intervals.

3.2.2.3. Interval Judgments: Experiment 4

Figure 3.6 (B) shows a different pattern to Experiments 1 - 3, with a general overestimation of noncausal trial intervals. However, a clear binding effect is evident, regardless of embedded event location. ANOVA supports this finding, with a main effect of Trial Type, $F(1, 30) = 80.31$, $p < .001$, partial $\eta^2 = .73$ and Embedded Event Location, $F(2, 60) = 15.37$, $p < .001$, partial $\eta^2 = .34$. No Trial Type x Embedded Event Location interaction was found, $F(2, 60) = 1.57$, $p = .22$, partial $\eta^2 = .05$. I conducted simple effects analysis to explore the main effect of Embedded Event Location. These revealed significant differences between *before* and *during* levels, and between *during* and *none* levels (all *ps* < .001). No difference between *before* versus *none* levels was found ($p = .13$).

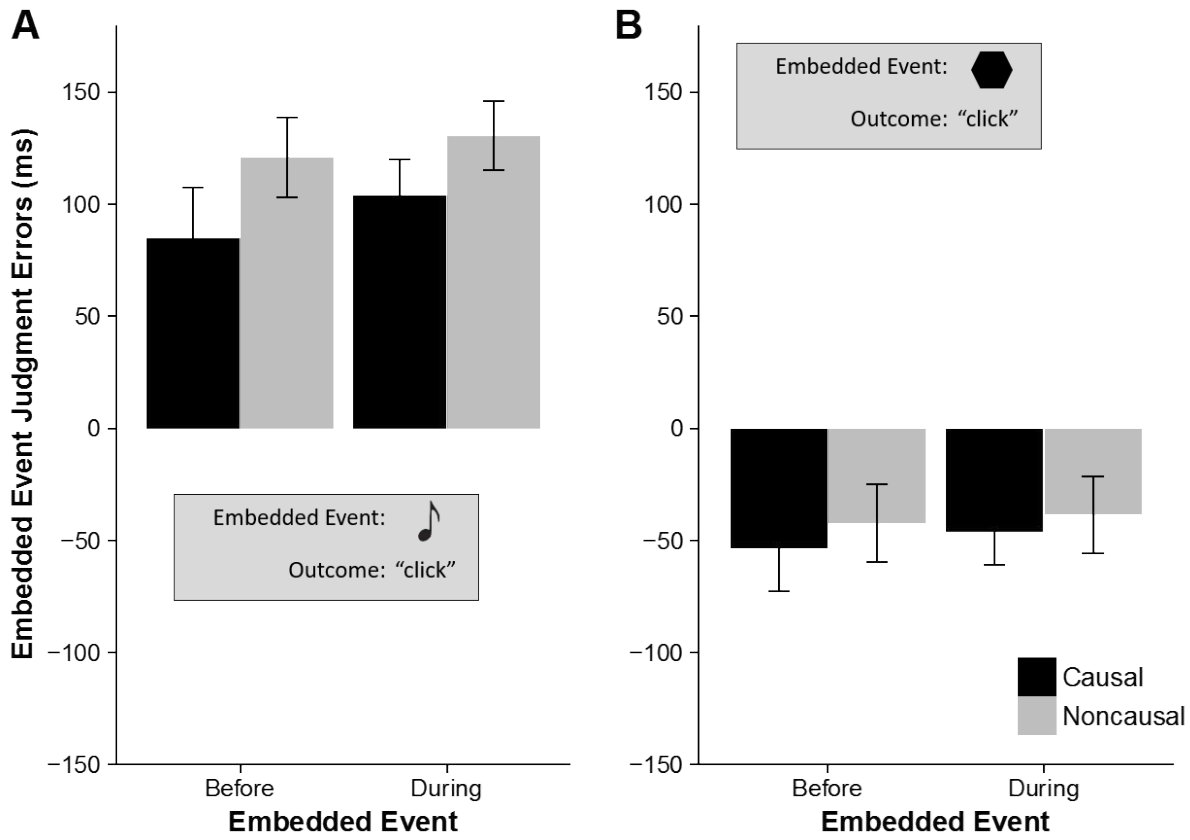

Figure 3.7. (A) Experiment 3. (B) Experiment 4. Mean embedded event judgment errors for events occurring before and during the interval, broken down by trial type. Error bars represent within-subjects confidence intervals.

3.2.2.4. Embedded Event Judgments: Experiment 3

Figure 3.7 (A) shows that embedded events are overestimated, corroborating the finding of Experiment 1. Inspection of Figure 3.7 (A) also suggests that, numerically, at least the pattern found in Experiment 1 might be reversed: events presented during causal trials might have been perceived as shorter than those presented during noncausal trials. However, ANOVA does not corroborate this impression: The effect of Trial Type failed to reach significance, $F(1, 27) = 3.94$, $p = .06$, partial $\eta^2 = .13$, and neither the effect of Embedded Event Location, $F(1, 27) = 1.71$, $p = .20$, partial $\eta^2 = .06$, or the Trial Type x Embedded Event Location interaction, $F(1, 27) = .25$, $p = .62$, partial $\eta^2 = .01$ were significant. A Bayesian analysis yielded a Bayes factor of 1.5 in favour of the *alternative* versus the null

model (see Embedded Event Judgments: Experiment 1 and Figure 3.5 for specifics). Thus, the evidence that events embedded during causal intervals are not perceived as shorter than those embedded during noncausal intervals, is inconclusive here.

3.2.2.5. Embedded Event Judgments: Experiment 4

Figure 3.7 (B) shows a general underestimation of embedded events, replicating the results of Experiment 2's embedded events (i.e., that auditory stimuli are judged longer than visual stimuli). Similarly to Experiment 3, Figure 3.7 (B) also suggests relative underestimation of embedded events in causal trials. However, just as in Experiment 3, statistical analysis does not bear this out, with no effect of Trial Type, $F(1, 30) = .42$, $p = .52$, partial $\eta^2 = .01$, Embedded Event Location, $F(1, 30) = .22$, $p = .64$, partial $\eta^2 < .01$, or Trial Type x Embedded Event Location interaction, $F(1, 30) = .06$, $p = .82$, partial $\eta^2 < .01$. A Bayesian analysis found a factor of 13 in favour of the null versus the alternative restricted model.

3.2.3. Discussion

The results of Experiments 3 and 4 replicate those of Experiments 1 and 2: I found a significant binding effect, regardless of embedded event location and modality. In Experiment 3 I also replicated the finding that noncausal intervals with an embedded event occurring *before* it were judged as longer than other Trial Type x Embedded Event Location combinations. As stated in the previous section, I will comment on this in Chapter 3: General Discussion. Regarding embedded event judgments, I again found no difference between temporal estimates of events embedded into causal and noncausal intervals, nor any differences depending on whether the event occurred *before* or *during* the interval. While the pattern of embedded event judgments is numerically reversed relative to Experiments 1 and 2, with the shortening of perceived durations for events embedded into causal intervals

approaching statistical significance in Experiment 3, I attribute this fluctuation to random noise, and perhaps changes to the outcome modality. Crucially, however, Experiments 3 and 4, like Experiments 1 and 2 do not support the hypothesis of general clock slowing, which would be evidenced by reliable subjective shortening of events embedded into causal intervals. Furthermore, neither Experiment 3 nor Experiment 4 showed any evidence of embedded event judgments being affected by Event Location. If traces of relative underestimation of events embedded into causal intervals were driven by general clock slowing during causal intervals, then such underestimation should only occur for events presented *during* causal intervals, and not for those presented before them. Results from both Experiments 3 and 4 clearly show no evidence for this, neither numerically, nor via a Trial Type x Embedded Event Location interaction. This result also holds when embedded events are clearly delineated, i.e., when embedded events *before* do not overlap *into* the interval (see Appendix A).

3.3. Experiments 1 – 4: Discussion

I developed a new procedure - the *embedded interval estimation procedure* - to study the potential implication of internal clock models in temporal binding. Using this method, I replicated the binding effect across four experiments, with causal intervals being consistently judged as shorter than noncausal intervals, irrespective of whether or when an additional event was embedded in the overall interval, and irrespective of the modality of the embedded event, and that marking the end of the interval. Interestingly, I found that when tones were deployed as embedded events (in Experiments 1 and 3), they were generally overestimated, while visual stimuli serving as embedded events (Experiments 2 and 4) were generally underestimated. This replicates a well-established finding that

auditory stimuli are judged as relatively longer than visual stimuli (Wearden et al, 1998) and reassures us of the overall validity of the embedded interval estimation method.

Contrary to what would be predicted if temporal binding were effected via a general slowing of subjective time, I found no difference in perceived duration of events embedded within causal and noncausal intervals. This pattern of results held across four experiments and regardless of whether the interval was marked by a visual or auditory event. Likewise, it was unaffected by whether the embedded event was visual or auditory. Interestingly, removing overlapping embedded events does not change the pattern of results. One anonymous reviewer expressed the concern that my failure to find an effect of causality on embedded event judgments, might simply be due to the (relatively) short durations I employed; differences in *r* might be more likely to manifest at longer intervals. However, I rule out this argument because Humphreys & Buehner (2009) found that participants gave shorter estimates for causal, compared to noncausal intervals, at durations as short as 150 ms, As the shortest interval in the present study was twice this duration (i.e., 300 ms), it is unlikely that my results can be explained be the mere durations of the embedded event.

Naturally, one always has to be cautious when interpreting null results. However, a Bayesian analysis of temporal judgments found Bayes factors greater than 10, in favor of the null hypothesis in three of the four experiments reported here (Experiments 1, 2, and 4; the exception is Experiment 3, which yielded a factor of less than 1). Furthermore, I combined the data from all Experiments and conducted an ANOVA with Embedded Event Modality and Outcome Modality as between-subjects factors. The results of this pooled analysis find significant effects of Embedded Event Modality ($p < .001$) and an Outcome Modality x Trial Type interaction ($p < .05$), reflecting the findings discussed above, namely that (1) auditory stimuli are judged as longer than visual (Wearden at el, 1998), and (2) the numerical

difference between judgment errors for events embedded in causal versus noncausal intervals reversed when the modality of the outcome stimulus was auditory compared to when it was visual. More importantly, I found no significant effects of Trial Type (p = .34) nor a Trial Type x Embedded Event Location interaction (p = .67). This corroborates the main observation that events embedded in causal intervals are not judged as shorter than events embedded in noncausal intervals, regardless of stimuli (outcome and embedded event) modality. Finally, I compared the restricted model against the null (see Embedded Event Judgments: Experiment 1) to yield a Bayes factor of 99 in favour of the null versus the alternative restricted model, pooling the data from all four Experiments (this pattern also holds for data with no overlapping embedded events; see Pooled Analysis in the Appendix A). In general then, this pattern represents positive evidence against a difference in perceived duration of events embedded in causal versus noncausal intervals.

An unanticipated result concerns noncausal intervals on trials with an embedded event *before* the interval, which were consistently judged as longer than any other interval category. One explanation implicates the trial structure I employed: Because on a given trial at most one embedded event occurred, the presentation of an event *before* the interval meant that participants knew that no further event will occur during the remainder of the trial. Participants would then have been able to attend exclusively to the interval, whereas otherwise they would still have to divide their attention between tracking overall interval duration and monitoring the potential occurrence of an embedded event. A common assumption in internal clock models is that subjective time is modulated by attention, with greater attention paid to time passing resulting in more clock pulses accrued, which in turn leads to the experience of a subjectively longer interval (cf. the common adage “A watched pot never boils” or “Time flies when you are having fun”, Avni-Babad & Ritov, 2003). This

attentional modulation would of course have affected causal and noncausal trials equally, and indeed this is reflected by analogously less negative judgment errors on causal trials with an interval embedded before the trial, relative to causal trials with an interval embedded during the trial. However, the binding effect evidently was robust enough to prevent causal intervals preceded by an embedded event from being overestimated. Another explanation is that interval judgments are biased by some property of the embedded event. For example, if an embedded event occurs *before* the marker delineating the start of the interval (disappearance of the black square or key press), perception of the marker could be biased towards the earlier embedded event. Consequently, the interval would then be perceived as longer. In contrast, if the embedded event occurs *after* the marker, such biases would lead to relatively shorter perception of the interval (I thank an anonymous reviewer for this suggestion).

Importantly, finding that temporal binding is robust regardless of the presence, location or modality of embedded events demonstrates the reliability of the procedure. What then, can my results say about how temporal binding relates to time perception itself? At a minimum, my results show that temporal binding does not occur due to general slowing of the pacemaker. These results are in accordance with work by Eagleman and colleagues (Pariyadath & Eagleman, 2007; Stetson et al., 2007; see also Eagleman, 2008, for a brief review). Pariyadath and Eagleman found that a perceived temporal expansion of an unexpected event did not result in the expansion of an additional, simultaneously presented event. They found that when the unexpected event (i.e., the oddball: a clock image presented during a series of shoe images), was simultaneously presented with tones, then no subjective temporal expansion occurred for the tones. Additionally, they found that flickering the visual images did not result in increased temporal acuity for the oddball

stimulus, which would be expected if the expansion is effected by an increase in the rate of an internal clock. This last effect was also demonstrated in a novel study in which participants made a 50 metre free-fall into a net, whilst looking at flickering digits on a wrist watch (Stetson et al., 2007). The authors reasoned that if an increased pacemaker is responsible for time 'slowing down' during a frightening event, then temporal resolution should also increase; participants should perceive the digits at a faster flicker rate whilst falling, than when stationary. Their results showed that this was not the case, as flicker thresholds during the fall did not differ to those measured before the fall. However, duration judgments did significantly differ, with participants overestimating their own fall in relation to the fall of other participants. Eagleman (2008) posits that distinct neural populations underlie different temporal judgments, such that duration judgments are independent of flicker rate judgments and simultaneity judgments, and so forth. One solution then, despite interval estimation being used for all duration judgments in the current experiments, is that different neural populations were responsible for the timing of intervals and embedded events (I discuss how this might be in Chapter 6: General Discussion). The key point is that temporal binding might be effected by a dedicated clock process, one allocated to keeping track of cause-effect intervals.

How might the timing of multiple events be achieved? Computationally, this would necessitate multiple clocks, each capable of independent and simultaneous timing. Klapproth (2011) provides evidence for such a notion. Using the temporal production technique, in which a key press terminates a temporally extended stimulus, participants were required to simultaneously produce two partially overlapping durations. More specifically, one (longer) duration served as the target interval, and the participant always began timing (i.e. producing) this interval first. Sometime after the onset of the stimulus that

was used to demarcate the target interval, a second temporally extended stimulus was presented, which served to demarcate a distracter interval. The temporally extended stimuli that defined the two to-be-produced target and distracter durations were visual or auditory (factorially combined). Klapproth's aim was to determine whether simultaneous timing (of separate durations) is effected via a single internal clock, or multiple clocks operating in parallel. The rationale was that multiple clocks would result in modality specific effects on clock speed: if auditory clocks run faster than visual ones (as found by Wearden et al, 1998) then auditory stimuli should be produced in less time than visual stimuli (the faster auditory clock would mean that pulses representing a given duration are accumulated in less objective time than pulses emitted from a visual clock). A key prediction made by Klapproth therefore was that auditory distractor events should result in shorter production times than visual distractors, regardless of whether they were embedded within auditory or visual target intervals. The results supported this prediction, in line with a multiple clock account of simultaneous timing. Given that there is precedence in the literature for multiple clock processes operating simultaneously, and at different clock speeds, it might well be possible that temporal binding reflects a change in time-keeping specific to tracking the elapsed interval between cause and effect. One possibility, for example, would be that *r* for clocks that time cause-effect intervals is lower relative to clocks dedicated to other timing processes.

Is there evidence for specific clock slowing in temporal binding? There is one study by Humphreys and Buehner (2009), who found evidence for changes in *r* during temporal binding. Using a verbal judgment paradigm, they found a linear relationship between perceived and actual durations for a range of causal and non-causal intervals. Importantly, they reported different slopes for causal and noncausal conditions, suggesting a stable

pacemaker rate *within* but variable *between* condition(s). Specifically, a shallower slope for subjective durations of causal relative to non-causal intervals suggested that clock processes dedicated to timing causal intervals ran slower than those used for tracking other events.

How do my findings relate to Wenke and Haggard (2009), who obtained evidence of clock slowing using simultaneity judgments? As noted in section 2.1.10 ('Internal clocks and temporal binding'), it is likely that Wenke and Haggard's results do not reflect clock slowing but instead simply reflect a motor-specific contraction of discrimination thresholds at the effector, which is independent of binding (Tomassini et al, 2014). Also, internal clock models are concerned with interval perception, and it is difficult to see how they may be applied to simultaneity judgments. Finally, a particularly glaring issue is that if pacemaker rate *r* is lower in the voluntary condition, then the slope of Wenke and Haggard's interval estimates should be shallower than that for the involuntary condition. This is not the case, which is another reason to believe that their temporal discrimination task does not measure pacemaker slowing. Given that the procedure I deployed in my studies is a more apposite measure of pacemaker rate than Wenke and Haggard's, I argue that the higher temporal discrimination thresholds they reported in causal versus noncausal trials reflect a temporal contraction related to motor planning (Tomassini et al, 2014).

3.3.1. Is binding effected by differences in timing latencies?

Until now I have discussed my results in light of the possibility that temporal binding might reflect a slowing of pacemaker speed *r*, with the important qualification that such changes would need to be specific to timekeeping of the cause-effect interval, and would not affect the timing of other events. However, my results are equally consistent with the possibility of causality-specific changes to timing latencies L_c and L_o: A delay between the actual onset of a causal interval and subjective timing, such that subjective timing begins

relatively later for causal compared to non-causal intervals would also result in shorter perceived durations. Similarly, if subjective timing terminates earlier in a causal condition relative to a noncausal one, then this would also contribute to shorter perceived durations. In other words, temporal binding could be due to causality-specific slowing of *r*, or to causality-specific changes in L_c and L_o. Importantly, whether binding is effected by causality-specific changes either to *r* or to switch latencies, the embedded event judgments in the current experiments would not change.

In sum, my results have conclusively ruled out the possibility that temporal binding is effected by a general slowing of an internal clock. Instead, they suggest that causality-induced changes in time perception necessitate the operation of multiple specific clock processes, such that causality leads to changes in *r* for the clock that tracks the cause-effect interval. Alternatively, causality could lead to changes in switch latencies L_c and L_o, which determine when time-keeping for cause-effect intervals begins and ends. One way to disentangle switch latencies versus specific pacemaker slowing is to conduct a temporal discrimination experiment, such that participants discriminate two causal intervals (in one condition) and two noncausal intervals (in another). If temporal binding is effected by a slower pacemaker rate, then the resultant poorer resolution would lead to higher discrimination thresholds for causal, relative to noncausal intervals.

Another way to disentangle these accounts is by systematically investigating clock processes in causal and noncausal intervals with regression models. Regression coefficients for the relationship between actual and perceived intervals indicate whether differences in time perception reflect changes in clock rate or differences in timing latencies (Wearden et al 1998; Matthews, 2011); differences in causal and noncausal intercept coefficients imply

different switch latencies, while different slope coefficients indicate different pacemaker rates.

At present, however, I conclude that temporal binding is not effected by a general pacemaker slowing. Instead, I assert that the change in perceived time, associated with temporal binding, occurs due to changes in time-keeping processes specific to the causal interval. Whether these changes reflect causality-specific pacemaker rate slowing, or changes to switch latencies, will be investigated in the following experiments.

4. Chapter 4: Investigating a specific account of clock slowing using psychophysical techniques

The results of Experiments 1 - 4 showed that binding might be effected by a specific clock process. Buehner and Humphreys (2009) provide tentative evidence for this hypothesis, by demonstrating a slower pacemaker in causal than noncausal trials using the regression technique (e.g., see Wearden et al., 1998; and 'Disentangling internal clock processes' in Chapter 1; the shallower slopes obtained from this procedure imply a slower pacemaker). Likewise, Wenke and Haggard (2009) also claim their evidence supports a slower clock in binding. Both studies are compatible, in that both use interval estimation as a measure of duration perception. However, whereas Buehner and Humphreys find slope differences between causal and noncausal conditions, Wenke and Haggard do not. This inconsistency means we cannot be certain whether temporal binding is effected by the slowing of a specific clock. Furthermore, the use of interval estimation is another reason to exercise caution; there is evidence that using interval estimation is not a reliable technique with which to disentangle internal clock processes (Matthews, 2011). In Matthews' study, slope differences were found when no differences should, theoretically, have occurred. This casts doubt on interval estimation as a technique to disentangle intercept and slope differences. Given Matthews' results, it would be wise to adopt a different technique in order to disentangle differences between slope and intercept coefficients.

Another reason to be wary of interval estimation as a means analysing regression coefficients, is that the assumptions of internal clock models are often violated (Wearden & Lejeune, 2008). One of the assumptions of internal clock models is that perceived durations should increase linearly with actual durations. Interval estimates, however, often display

Vierordt's law (Gu & Meck, 2011). This refers to the finding that shorter intervals are overestimated, while longer intervals are underestimated, which is a violation of the linear assumption of SET. One reason why interval estimation violates this linear assumption is because duration judgments are based on an internal, and not external, reference stimulus. The difference is that studies using an external reference stimulus present this particular (fixed) duration at some point(s) during the experiment. A test stimulus, of variable duration, is then compared to the reference. Thus, both stimuli that require comparison are presented on during the experiment.

In contrast, in interval estimation based-studies, there is no external reference stimulus. Instead, the reference is an internal memory of a duration, which might be prone to distortion (memory is not infallible, after all). A better option is to use discrimination methods (alluded to above), which present the reference stimulus during the experiment. Nolden et al. (2012) used a discrimination task, specifically, the method of constant stimuli, to replicate the binding effect. Participants judged whether a series of comparison durations (in causal and noncausal conditions), were shorter or longer than a fixed reference interval. For causal and noncausal conditions, the authors estimated the point of subjective equality (PSE), which refers to the duration at which the comparison is perceived the same as the reference interval. Their results revealed smaller PSEs in causal versus noncausal conditions, for both reference intervals tested (250 and 600 ms). However, there were too few reference durations employed to conduct regression analyses, and thus, determine whether different slopes were evident between causal and noncausal conditions.

Therefore, more evidence is required to establish whether temporal binding is actually effected by a slower clock, by employing psychophysical procedures and then conducting regression analyses. The experiments in the current chapter use a variant of that

used in other studies, whereby interval estimates are regressed onto actual durations and the slopes analysed (e.g., Wearden et al, 1998). In the present experiments (5 - 8) however, I regressed PSEs onto actual durations, before inspecting the slope and intercept coefficients. In these experiments, participants experienced two intervals both in causal and noncausal conditions: in causal trials, one of the intervals was between a key press and visual flash, while in noncausal trials, one of the intervals was delineated by two visual flashes. In both conditions an additional single stimulus served as the comparison interval. Participants judged whether the single stimulus interval was shorter or longer than the key press-flash interval (causal trials) or the flash-flash interval (noncausal trials). I computed the point of subjective equality (PSE) as a measure of perceived interval duration, for a range of durations. I expected shorter causal versus noncausal PSEs, due to temporal binding. Importantly, a regression of PSEs onto actual durations would reveal slope differences if pacemaker rates vary between causal and noncausal intervals.

4.1. Experiments 5 and 6

In causal and noncausal conditions participants experienced reference and comparison intervals. The reference interval was delineated by two events: a key press and visual flash (causal trials) or two visual flashes (noncausal trials). The comparison interval was a single visual stimulus in both conditions. The sequence of reference and comparison intervals differed between experiments, with either the reference interval first and comparison second (Experiment 5), or the reverse in Experiment 6. In both experiments participants compared the duration of the comparison to the reference interval. The PSEs were then estimated and regressed onto reference durations to analyse the coefficients. I predicted lower PSEs in addition to a shallower slope, for causal versus noncausal conditions.

4.1.1. Method

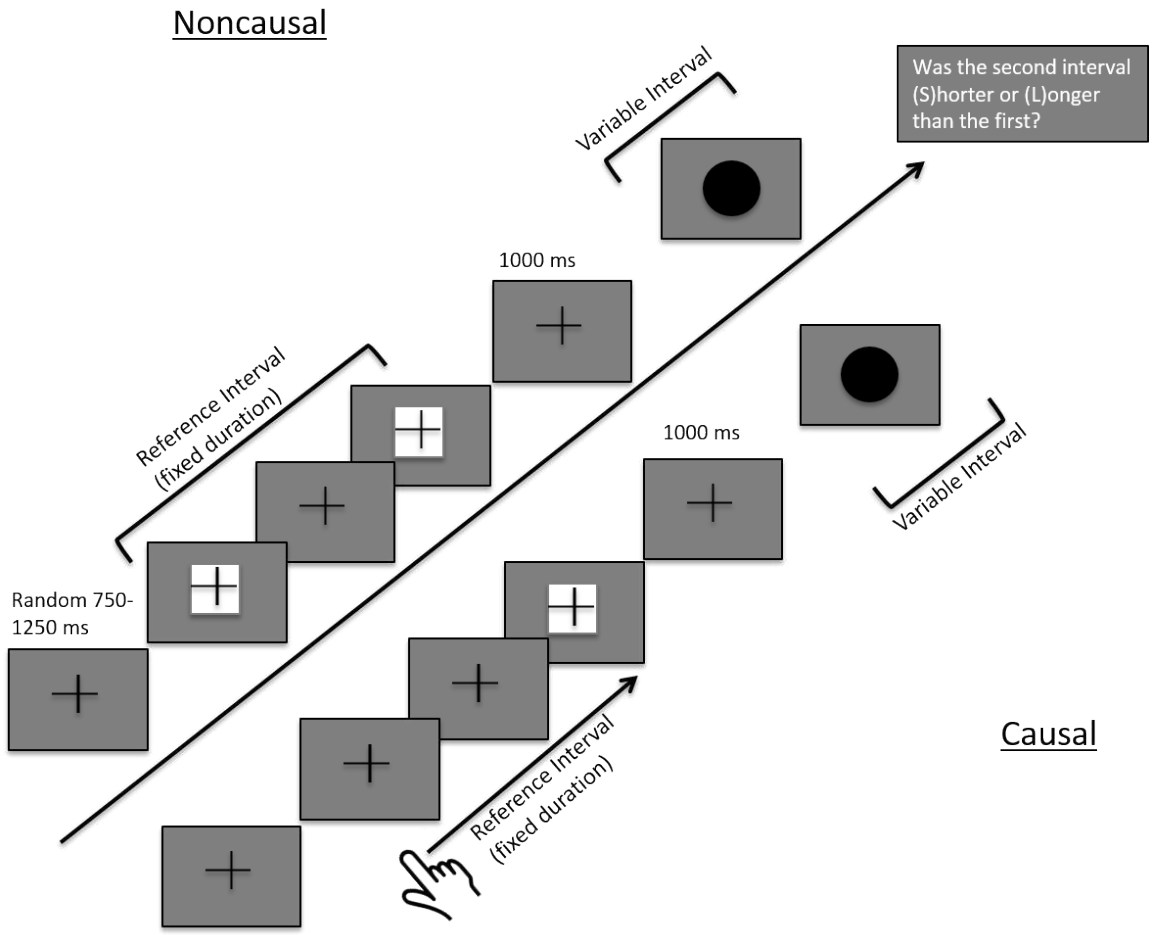

Figure 4.1. Trial structure of Experiment 5. Participants were presented with two intervals in each condition. The duration of the first interval was fixed for a block, while the second was variable. In causal trials, a key press triggered a visual flash after an delay (interval 1), followed by an extended temporal event (black circle: interval 2). In noncausal trials, two visual flashes delineated interval 1, followed by the black circle (interval 2). Participants were asked whether the second interval was shorter or longer than the first.

4.1.1.1. Participants

Sixteen students (14 female, 2 male, M_{age} = 20.4 years, age range: 18-25) of Cardiff University participated in Experiment 5 and 18 (15 female, 3 male, M_{age} = 19.6 years, age range: 18-23) in Experiment 6. Prior to analysis, the data were screened for normality. Participants whose PSEs deviated from the mean by more than 3 standard deviations were not entered into the analysis. This led to four participants being removed in Experiment 5

and six in Experiment 6. However, I replaced certain (but not all) participants in Experiment 6 by retesting, to enable a between-subjects comparison of each experiment. The reason was to afford a suitable comparison of Experiments 5 and 6 using equal sample sizes; whilst the pattern of results does change with different participant numbers, I nevertheless wished to avoid incurring any explanation of the results that might be attributable to unequal samples. Thus, the final sample entered into the analyses were 12. Participants received course credit or £10 payment.

4.1.1.2. Apparatus and stimuli

The experiment was implemented in Psychopy (Peirce, 2007) on cathode ray tube (CRT) monitors connected to Apple Mac Minis, with resolution of 1280 x 1024 and refresh rate of 120 Hz. Both conditions (causal, noncausal) featured two intervals. A reference interval was marked by two visual events: a key press and white square in causal trials, and two white squares in noncausal trials (all white squares were 200 pixels^2 and displayed for 50 ms). The second was a single visual event (a black circle; radius 70 pixels), which served as the comparison interval. The duration of reference intervals was fixed for a block, while the duration of the comparison interval varied from trial to trial. A white fixation cross (60 pixels^2) was displayed throughout each trial. All stimuli were presented centrally on the screen.

4.1.1.3. Design and procedure

I employed two factors: Trial Type (causal and noncausal) and Reference Duration (10 durations: 100, 200, 300, . . . 1000 ms). The dependent variable was participants' discrimination judgments. Participants completed 10 blocks of causal and noncausal trials (one block per Reference Duration x Trial Type combination), with each block comprising around 30 trials. This resulted in around 600 trials total, with an additional 12 practice trials

for causal and noncausal blocks (actual trial numbers varied due to the staircase procedure employed – each block ended when minimum reversal and trial number criteria were satisfied; see below for specifics).

Figure 4.1 shows the trial structure of Experiment 5. Causal trials began with the display of a fixation cross on screen. Participants were told to press a key at a time of their choosing, which triggered a visual flash after a fixed duration (100, 200, 300, . . . 1000 ms). After an ISI of 1000 ms, the black circle was displayed for a duration that varied from trial to trial (trials began with durations of 50% and 150% of the reference duration, for two interleaved staircases, respectively. See below for specifics). Participants then indicated whether the black circle was shorter or longer in duration than the key press - flash interval by pressing the S or L keys, respectively. After responding, the screen blanked for a random duration (1500 - 2300 ms) before the display of the fixation cross signalled the next trial.

Noncausal trials followed a similar procedure. The trial began with the display of a fixation cross, presented for a random duration (750 - 1250 ms). A visual flash marked the start of the first interval, followed by another flash after a fixed interval (100, 200, 300, . . . 1000 ms). After an ISI of 1000 ms, the second interval (black circle) was displayed for a duration (variable from trial to trial). Participants were then asked whether the second (variable) interval was shorter or longer than the first (fixed) interval. The screen then blanked for a random duration (1500 - 2300 ms) before the fixation cross signalled the beginning of the next trial.

The experiment began with either a causal or noncausal block, thereafter alternating between them. The duration of the fixed interval was selected at random from one of the ten Reference Duration intervals, and remained fixed for each block. The variable interval was controlled by a staircase procedure, using a Kesten stochastic approximation algorithm

(Kesten, 1958; Treutwein, 1995). This allows for fast and reliable convergence onto any probability threshold. I used 2 randomly interleaved staircases, each with target convergence thresholds of 0.5, with each block ending after a minimum of 4 reversals and 15 trials per staircase. The first trials of each staircase began with durations of 0.5 and 1.5 times the fixed duration, i.e., for a fixed interval of 400 ms, the variable interval would begin at durations of 200 and 600 ms for each respective staircase. Additionally, the first trial of each staircase was classed as a control trial, with the first and second trials of each staircase set to the same duration, regardless of response. This was to ensure that incorrect or mistaken responses early in each block did not adversely affect the direction of each staircase, e.g., a mistaken response would alter the direction of the staircase and reduce the likelihood of convergence within the scheduled number of trials.

Participants first completed practice blocks (causal and noncausal), which contained a fixed reference duration of 650 ms. These were shorter than main experimental blocks and were programmed to end after a minimum of 2 reversals and 6 trials per staircase. Participants then alternated between causal and noncausal blocks of the main experiment. Throughout the experiment, participants were given a short 2-minute break between blocks. The study was conducted across 2 sessions, with one session per day. The sessions were scheduled on two successive days. Participants began with a causal block in session one and noncausal in session two.

Experiment 6 followed the same general procedure except that I reversed the order of fixed and variable intervals. The comparison interval was presented first, and the reference interval second. The fixation cross disappeared during the ISI period, so that participants would know when to press the key to trigger the flash. Participants were still asked whether the comparison interval was shorter or longer than the reference interval.

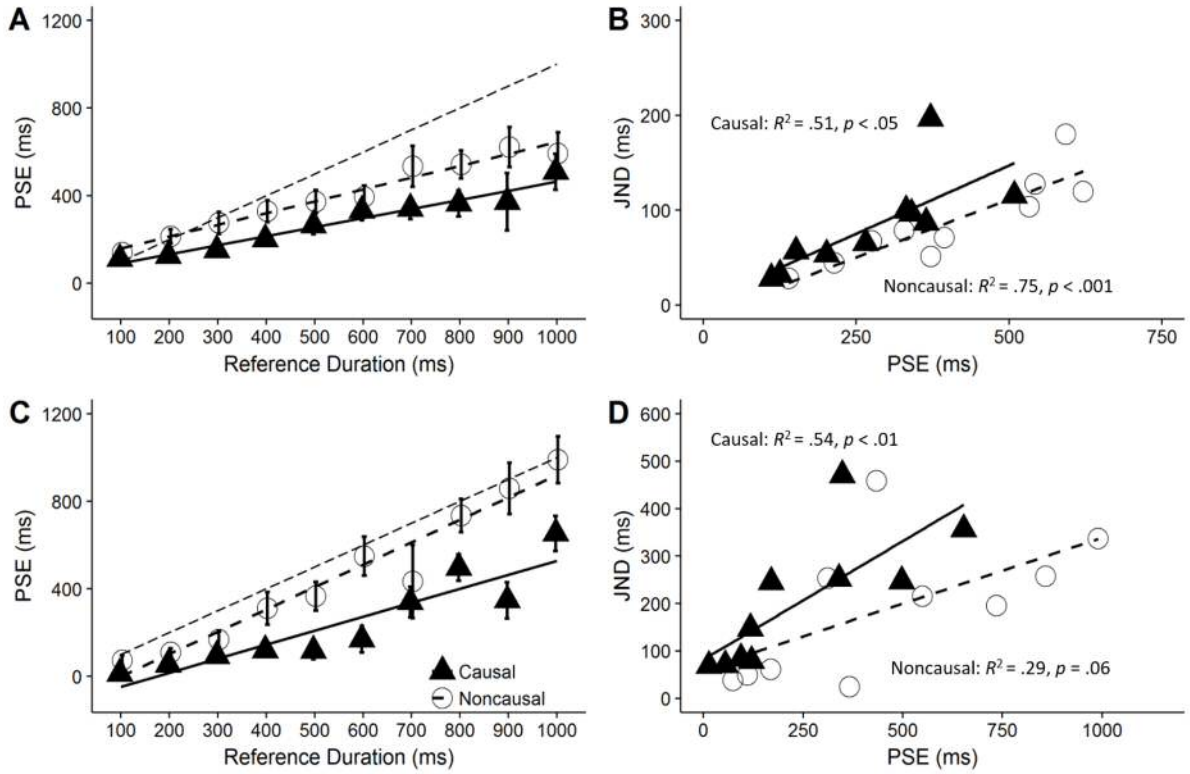

Figure 4.2. Results of Experiments 5 (upper row: A & B) and 6 (lower row: C & D). (Left panels A & C): Causal PSEs are smaller in both experiments, replicating the temporal binding effect. The long and bold dashed line represents the best linear fit; this shows that durations are generally underestimated in both conditions, particularly Experiment 5. Error bars show standard error. (Right panels B & D): Scatterplots show JNDs increase linearly with PSEs.

4.1.2. Results

4.1.2.1. Experiment 5

Prior to analysis, the control (first) trial of each staircase was removed and then cumulative Gaussian curves were fitted, following the generalized linear model (GLM) procedure in R (Knoblauch & Maloney, 2012). I then estimated the PSEs and JNDs for each participant.

PSE analysis. Figure 4.2 (A) shows that, in general, all combination of Trial Type x Reference Duration PSEs are underestimated, relative to objective durations. More importantly, causal PSEs were shorter than noncausal for most levels of Reference Duration. Figure 4.2 (A) also shows a linear relationship between PSEs and actual durations, indicating

that participants distinguished between reference durations. Statistical analyses support these observations, with ANOVA showing a significant effect of Trial Type, $F(1,11) = 6.34$, $p < .05$, partial $\eta^2 = .36$, Reference Duration, $F(9,99) = 20.56$, $p < .001$, partial $\eta^2 = .65$, but no Trial Type x Reference Duration interaction, $F(9,99) = 1.25$, $p = .28$, partial $\eta^2 = .10$.

Regression analyses. To investigate pacemaker and switch latency differences between causal and noncausal conditions, I regressed PSEs onto reference durations for each participant. I then conducted separate t-tests on the mean intercept and slope coefficients. While Figure 4.2 (A) shows that, numerically, the intercepts differ between causal versus noncausal conditions, the results are not significant, $t(11) = -1.93$, $p = .08$. Neither did I find a significant difference between slopes, $t(11) = -1.07$, $p = .31$. I conducted Bayesian t-tests to explore the likelihood both of the difference between intercept and between slope coefficients. A Bayes factor of 1 for the intercept difference indicates that the alternative and null hypotheses provide and equally likely fit of the data. Meanwhile, I found a Bayes factor of 0.5 in favour of difference between slopes (or, put differently, a Bayes factor of 2 in favour of *no* difference between slope coefficients). Thus, a difference in intercepts is twice as likely to explain the data than a difference between slopes.

Scalar variability. Finally, I was interested in the variability of timing across the range of durations employed. One quality of internal clock models is that variability remains constant across a range of to-be-timed durations. This refers to the scalar property of timing[3] (known as the coefficient of variation: CV, essentially an index of timing sensitivity),

[3] Wearden and Lejeune (2008) make a distinction between empirical and theoretical scalar timing. The latter refers to instances where the underlying representations are considered to have scalar properties, but additional factors (e.g., task difficulty) obscure this property. Wearden and Lejeune use the example of Ferrara, Lejeune and Wearden (1997), who found that the CV varied with task difficulty. However, because perceived durations increased with actual durations, the results were said to conform to *theoretical* timing.

and is a form of Weber's law (Grondin, 2010; Wearden & Lejeune, 2008). Because pacemaker reliability is usually explained as the cause of timing variability, any variability within the CV suggests that the pacemaker rate varies depending on the duration of the stimulus. One way to assess this component is to plot standard deviations against mean estimates; standard deviations should increase linearly with mean estimates, such that the ratio of the former to the latter is constant across durations. In the present case, this would involve plotting JNDs against PSEs for all reference durations, which can be seen in Figure 4.2 (B). If pacemaker rate is stable within a particular condition, then JNDs should increase linearly with PSEs. Inspection of this figure shows that JNDs do indeed increase linearly, both in causal and noncausal conditions. This is what would be expected if timing sensitivity were constant across a range of durations. To determine whether this observed linear trend is significant, I conducted separate regression analyses for causal and noncausal conditions. A regression of JNDs onto PSEs found significant models for causal, $p < .05$, adjusted $R^2 = .51$, and noncausal conditions, $p < .001$, adjusted $R^2 = .75$. Additionally, I conducted an ANOVA to determine whether JNDs differed between causal and noncausal conditions. In line with the regression results, ANOVA found a significant main effect of Reference Duration, $F(9,99) = 5.88$, $p < .01$, partial $\eta^2 = .35$, but no effect of Trial Type, $F(1,11) = .22$, $p = .65$, partial $\eta^2 = .02$, nor a Reference Duration x Trial Type interaction, $F(9,99) = .90$, $p = .41$, partial $\eta^2 = .08$. Thus, I conclude that timing sensitivity remained constant across the range of durations I employed, and was not affected by causality.

4.1.2.2. Experiment 6

PSE analysis. As in Experiment 5, I removed the first trial of each staircase before fitting cumulative Gaussian functions and estimating PSEs and JNDs. Figure 4.2 (C) shows that all PSEs are lower than objective durations. Also, PSEs increase linearly with reference

durations, implying that participants distinguished between reference durations. Importantly, causal PSEs are perceived as shorter than noncausal, in line with the results of the previous experiment. This causal versus noncausal difference is greater than in Experiment 5, particularly at reference durations of 400 ms and higher. A 2 x 10 ANOVA with Trial Type (causal, noncausal) and Standard Duration (100 - 1000 ms) as factors support these observations, with a main effect of Trial Type, $F(1,11) = 24.10$, $p < .001$, partial $\eta^2 = .69$, Reference Duration, $F(9,99) = 31.30$, $p < .001$, partial $\eta^2 = .74$, and a Trial Type x Reference Duration interaction, $F(9,99) = 3.29$, $p < .01$, partial $\eta^2 = .23$. Simple effects analysis found significant effects of Trial Type for all levels of Standard Duration except for 700 and 800 ms (*ps* < .05).

Regression analyses. I conducted individual regression analyses to investigate differences in slope and intercept coefficients between Trial Type levels. T-tests on these results did not find a significant difference between intercepts, $t(11) = -.23$, $p = .82$, but did yield a significant difference between slopes, $t(11) = -3.49$, $p < .01$. Bayesian t-tests found factors of 10, in favour of a slope difference, and 0.3 in favour of the intercept difference (or, put differently, a Bayes factor of 3 in favour of the null hypothesis for a difference between causal and noncausal intercept coefficients).

Scalar variability. I also investigated whether the data exhibited the scalar property found in Experiment 5. This is shown in Figure 4.2 (D), which shows JNDs plotted by PSEs. Generally, the JNDs increase linearly with PSEs, corroborating the findings of Experiment 5. However, JNDs are dispersed more widely around the line-of-best-fit in the current experiment, relative to Experiment 5. One interpretation suggests that changing the trial structure for the current experiment affected the memory of each interval. The duration of the reference interval was fixed throughout each block, which may have strengthened its

representation in memory. In contrast, the comparison interval varied from trial to trial, which may have resulted in a relatively weaker representation. Because the comparison interval was presented first in the trial sequence (and thus more distant in memory) in the current experiment, it may be that its memory trace faded somewhat; in Experiment 5, the comparison interval was presented second in the trial (and thus more recent in memory), and did not have sufficient time to decay. Nevertheless, reversing the trial structure might have increased task difficulty but was not sufficient to entirely abolish the linear relation between JNDs and PSEs. Therefore, to determine whether this linear trend is significant, I regressed JNDs onto PSEs separately for causal and noncausal conditions. I found a significant model for causal JNDs, $p < .01$, adjusted $R^2 = .54$, but the noncausal model was outside significance, $p = .06$, adjusted $R^2 = .29$. Additionally, ANOVA found that the effect of Reference Duration was also just outside significance, $F(9,99) = 3.29$, $p = .06$, partial $\eta^2 = .23$, while no significant effects were found for Trial Type, $F(1,11) = .13$, $p = .73$, partial $\eta^2 = .01$, nor the Reference Duration x Trial Type interaction, $F(9,99) = .96$, $p = .40$, partial $\eta^2 = .08$. These results are in line with those of Experiment 5, in that JNDs did not differ between causal and noncausal conditions.

4.1.3. Discussion

In two experiments I replicated the temporal binding effect using a temporal discrimination procedure. I found that PSEs increased linearly with Reference Duration in both conditions, indicating that participants successfully distinguished between durations. Also, PSEs were lower than objective reference durations in both conditions. I also found a positive linear association between discrimination thresholds (JNDs) and PSEs. Although, this differed between experiments, with more deviation from the trend line in Experiment 6. Clearly then, reversing the order of reference and comparison intervals affected the

difficulty of the task; the linear trend provides a better fit of the data when the comparison interval is presented second in the trial sequence (Experiment 5), compared to the reverse (Experiment 6). In other words, my data exhibit a recency effect, which has been documented in the time perception literature previously (Matthews, 2013). Importantly, while this recency effect may have affected timing sensitivity, it was not sufficient to abolish or mar, the temporal binding effect.

The main purpose of Experiments was to analyse the regression coefficients. The results of Experiment 6 revealed a significant difference between the slopes of causal and noncausal PSEs, while Experiment 5 found no significant differences either between slopes or intercepts (although I am among the first to examine slopes in temporal binding, the lack of an intercept difference is a common finding in time perception experiments, see e.g., Matthews, 2013; but see Matthews, 2011, for an exception). Given that the results of Experiments 5 and 6 do not lead to unequivocal conclusions, further experiments are necessary. To further strengthen my claim of a slower clock in temporal binding, and to attempt to overcome the conflict between the present set of results, I conducted further experiments that replicate 5 and 6. However, for practical reasons it was necessary to reduce the length of the experiment and therefore, fewer reference durations (5) were used. Given that slope differences are often found with fewer than 10 data points (e.g., Humphreys & Buehner, 2009; Matthews, 2011), I did not anticipate any reason why this reduction would adversely impact the validity of the results.

4.2. Experiments 7 and 8

The current experiments replicate those of 5 and 6 in order to gain a clearer understanding of the internal clock processes in operation in temporal binding. I replicated

Experiments 5 and 6 using fewer reference durations. Otherwise the procedure was identical.

4.2.1. Method

4.2.1.1. Participants

Sixteen students (16 female, M_{age} = 18.9 years, age range: 18-20) of Cardiff University participated in Experiment 7 and 19 (17 female, 2 male, M_{age} = 20.6 years, age range: 18-29) in Experiment 8. I applied the same exclusionary criteria as in Experiment 5, leading to the removal of 3 participants in Experiment 7, and 6 in Experiment 8. As before, to enable a comparison between experiments, I replaced participants in Experiment 8. This resulted in 13 participants entered in to analyses in both experiments. Participants received course credit or £5 payment.

4.2.1.2. Apparatus and stimuli

This remained the same as Experiment 5.

4.2.1.3. Design and procedure

The procedure was analogous to Experiments 5 and 6 with the exception that I used fewer reference durations (200, 400, 600, 800 and 1000 ms). In Experiment 7 the reference interval was presented first the comparison interval second, reversed in Experiment 8.

4.2.2. Results

4.2.2.1. Experiment 7

PSE analysis. I removed the first trial of each staircase and fitted cumulative Gaussian functions, before estimating PSEs and JNDs. Inspection of Figure 4.3 (A) shows that causal PSEs are perceived as shorter than noncausal, in line with the results from Experiments 5 and 6. Also, causal PSEs increase linearly with Reference Duration, indicating that

participants discriminated reference durations. However, this increase is of a lesser magnitude in causal than noncausal conditions (i.e., shallower slope), corroborating the results of Experiment 6. Visually then, the results support a slower clock rate in causal conditions. A 2 x 5 (Trial Type [causal, noncausal] x Reference Duration [200, 400, 600, 800, 1000]) ANOVA found a significant effect of Trial Type, $F(1, 12) = 8.01$, $p < .05$, partial $\eta^2 = .40$, Reference Duration, $F(4, 48) = 22.80$, $p < .001$, partial $\eta^2 = .66$, and a Trial Type x Reference Duration interaction, $F(4, 48) = 4.56$, $p < .01$, partial $\eta^2 = .28$. Simple effects analysis found significant effects of Trial Type at 600, 800, and 1000 ms (*ps* < .05).

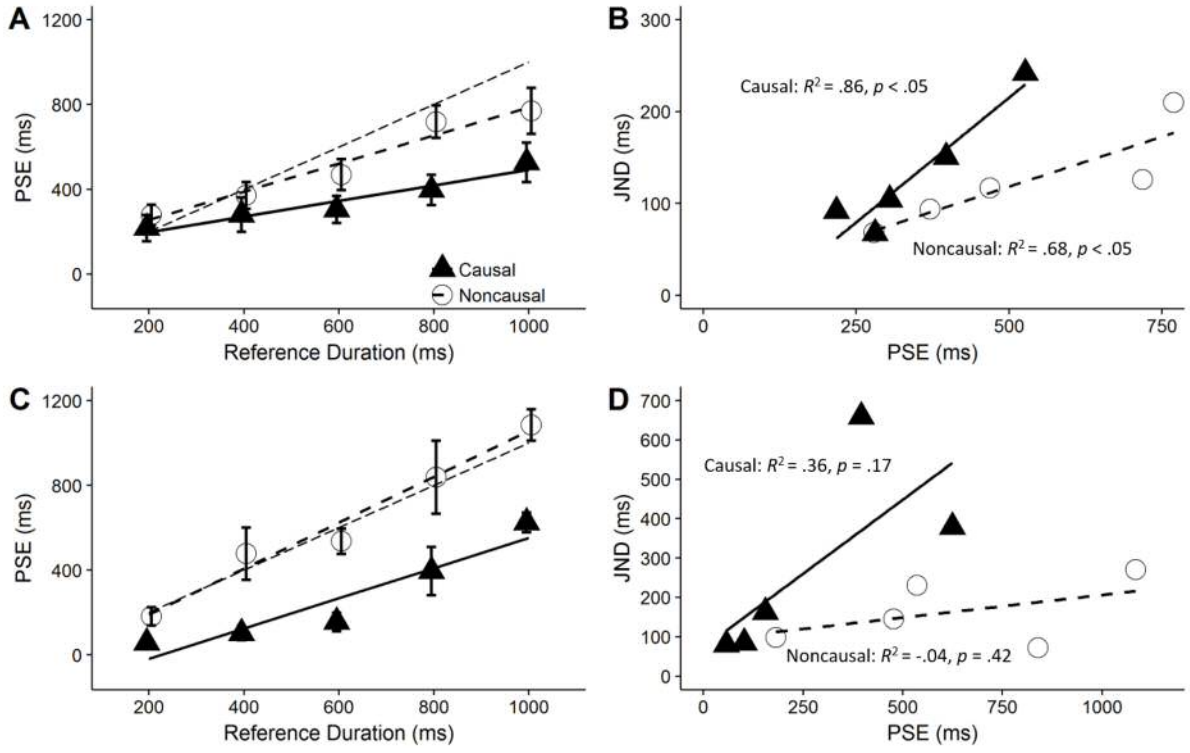

Figure 4.3. Results of Experiments 7 (upper row: A & B) and 8 (lower row: C & D). (Left panels A & C): Causal PSEs are smaller in both experiments, replicating the temporal binding effect. The long and bold dashed line represents the best linear fit; this shows that durations are generally underestimated in both conditions. Error bars represent standard error. (Right panels B & D) Scatterplots show JNDs increase linearly with PSEs.

Regression analyses. As with Experiments 5 and 6, I regressed PSEs onto reference durations for each individual participant, and conducted t-tests on the mean intercept and slope coefficients for causal and noncausal conditions. In line with the results of Experiment 6, these reveal no difference between intercepts, $t(12) = .03$, $p = .97$, but a significant difference between slopes, $t(12) = -3.87$, $p < .01$. Bayes t-tests found factors of 20 and 1 in favour of the alternative hypothesis, for slope and intercept differences, respectively.

Scalar variability. Finally, JNDs were regressed onto PSEs to determine whether the scalar assumption holds. This can be seen in Figure 4.4 (B). I found significant models both for causal, $p < .05$, adjusted $R^2 = .86$, and noncausal conditions, $p = .05$, adjusted $R^2 = .68$, supporting my observation that JNDs increase linearly with perceived duration (PSEs). ANOVA supports these results, with a significant effect of Reference Duration, $F(4,48) = 4.77$, $p < .05$, partial $\eta^2 = .28$, but neither Trial Type nor the Reference Duration x Trial Type interaction were significant, $F(1,12) = .06$, $p = .82$, partial $\eta^2 < .01$, and $F(4,48) = .23$, $p = .82$, partial $\eta^2 = .02$, respectively.

4.2.2.2. Experiment 8

PSE analysis. I again fitted cumulative Gaussian functions and estimated PSEs and JNDs. Figure 4.3 (C) displays the PSEs for causal and noncausal conditions. The results show that causal PSEs are perceived as shorter than noncausal, replicating findings from Experiments 5 - 7. Also noticeable in Figure 4.3 (C) is a shallower slope in the causal condition, indicating a slower clock rate. ANOVA supports these observations, with a significant effect of Trial Type, $F(1, 12) = 15.32$, $p < .01$, partial $\eta^2 = .56$, and Reference Duration, $F(4, 48) = 31.25$, $p < .001$, partial $\eta^2 = .72$, but no significant Trial Type x Reference Duration interaction, $F(4, 48) = 1.74$, $p = .16$, partial $\eta^2 = .13$.

Regression analyses. The mean intercept and slope coefficients for each level of Trial Type were also analysed, revealing no significant difference between intercepts, $t(12) = -1.58$, $p = .14$, but a significant difference between slopes, $t(12) = -2.78$, $p < .05$. Bayes t-tests found factors of 4 and 1 in favour of the alternative hypothesis, for slope and intercept differences, respectively.

Scalar variability. Lastly, to investigate the scalar assumption of timing I regressed JNDs onto PSEs (see Figure 4.4 [D]). Interestingly, I found no significant models for causal, $p = .17$, adjusted $R^2 = .36$, or noncausal conditions, $p = .42$, adjusted $R^2 = -0.04$. ANOVA confirms the results of these regressions, with no main effect of Reference Duration, $F(4,48) = .99$, $p = .36$, partial $\eta^2 = .08$. However, this can be explained by the wide dispersion of the JNDs, evidencing the increased difficulty of the task; when the reference duration is the second interval in a trial, participants require a higher JND to discriminate durations. An important finding is that JNDs were not affected by causality, with no effect of Trial Type, $F(1,12) = .61$, $p = .45$, partial $\eta^2 = .05$, nor the Reference Duration x Trial Type interaction, $F(4,48) = 1.38$, $p = .27$, partial $\eta^2 = .10$.

4.2.3. Discussion

The results of Experiments 7 and 8 largely replicate those of 5 and 6. I found a significant temporal binding effect in both experiments, with lower PSEs in causal, versus noncausal trials. The PSEs increased linearly with Reference Duration, suggesting that participants correctly distinguished the intervals. Additionally, PSEs in causal and noncausal conditions were lower than objective reference durations, again corroborating the results of Experiments 5 and 6. I also found that my data generally conformed to the scalar component of timing, as JNDs regressed onto PSEs exhibit significant models in both conditions in Experiment 7; in Experiment 8 though, the models were not significant.

However, this is likely due to the increased task difficulty brought about via changes in the trial structure from Experiment 7 to 8, corroborating the results of Experiment 6. Additionally, the regressions in Experiments 5 and 6 were performed using 10 data points, while those in Experiments 7 and 8 use half that number. This suggests that the results are robust as to be unaffected by the number of reference durations employed.

A further important finding is that, contrary to Experiments 5 and 6, I found a significant difference between causal and noncausal slopes in each experiment, but no difference between intercept coefficients. This occurred regardless of the differences in trial structure between Experiments 7 and 8, which, together with Experiment 6, supports a slower pacemaker in temporal binding.

4.3. Experiments 5 – 8: Discussion

The aim of these experiments was to determine whether temporal binding is effected by the slowing of a specific clock. Previous experiments found that clock slowing in binding is not a general effect but rather, is likely a process unique to cause-effect intervals (see Chapter 3). I therefore adopted a different paradigm to the experiments in the previous chapter, one more suited to investigating a specific cause-effect clock slowing.

I first estimated the PSEs for a range of durations, before conducting regression analyses to determine to slope and intercept coefficients. In three Experiments (6 – 8) I found significantly shallower slopes in causal, than noncausal intervals, which signifies a slower clock in cause-effect intervals. This replicates the results of Humphreys and Buehner (2009), who also found shallower slopes for cause-effect intervals using interval estimation. Importantly, I found that, generally speaking, the ability to discriminate intervals increased linearly with Reference Duration in both conditions. This indicates that the pacemaker is stable within conditions but different between. My results are comparable with Nolden et

al. (2012), who found lower PSEs in causal, relative to noncausal conditions, using the method of constant stimuli. Interestingly, their Weber fractions did not differ either between reference durations or key press condition (causal and noncausal). This indicates that timing sensitivity remained constant across a range of durations, in line with my findings.

4.3.1. Switch latencies and timing sensitivity in temporal binding

In line with the majority of studies in time perception (Matthews et al., 2011; Penton-Voak et al., 1996; Wearden et al., 2007), the evidence from the current experiments do not support switch latency differences. Experiments 5 - 8 used a regression procedure commonly used to disentangle pacemaker from switch latency differences (Wearden et al., 1998). Inspection of the regression coefficients found no difference between causal and noncausal intercepts, which is a key indicator of switch latency differences. Thus, the results of my experiments are compatible with many others in time perception (see Matthews, 2011), which implicate changes in pacemaker rate as the root of temporal variability.

One might ask how and why an internal clock would slow down, and how this relates to Eagleman and Holcombe's (2002) sensory recalibration account of binding. However, there are a number of points to discuss before I address these questions. One relates to the variability in discrimination thresholds (JNDs) in the experiments. Although I replicated the binding effect with lower PSEs in causal trials in all experiments, I found a difference in discrimination thresholds; JNDs in Experiments 6 and 8 were more variable than in Experiments 5 and 7. Because participants always judged whether the comparison interval was shorter or longer than the reference interval (regardless of whether the comparison interval was first or second in the trial sequence), changing the sequence of this comparison interval affected discrimination judgments; JNDs were dispersed more widely when the

comparison interval was first in the trial sequence. This suggests that timing sensitivity per se was unaffected, but rather, implicates a decay in the representation of the comparison interval (Matthews, 2013). Importantly, this memory decay was not sufficient to adversely affect or even abolish the temporal binding effect. In principle then, I conclude that timing sensitivity (i.e., JNDs), increases linearly with Reference Duration.

Why did I not find a difference between slopes in Experiment 5? At best I attribute this to the random fluctuation inherent in psychological experiments. It cannot be simply due to the procedural difference between Experiment 5 and 6, because the same differences exist between Experiments 7 and 8, in which the results revealed a difference in slope coefficients between causal and noncausal conditions for both experiments. Furthermore, this difference was significant with only five (compared to ten) reference durations entered into the regression analyses. Importantly, I found a significant difference between causal and noncausal slope coefficients when the data from Experiments 5 and 6 were combined ($p < .01$). This suggests that the difference between slopes is a particularly robust effect. In sum, my results support the hypothesis of a slower pacemaker in temporal binding.

4.3.2. Pacemaker slowing or drifts in attention?

It is important to be clear about what the slope coefficient represents. A difference in slopes between causal and noncausal conditions represents a different number of pulses accumulated between conditions. Given an absence of an intercept difference, there are two possibilities: one is that there is indeed, a slower pacemaker, while the other is that pacemaker rate is constant, yet there are drifts in attention. The latter suggests that pulses are missed to due to non-focal attention, such that certain pulses are not accumulated. Both

accounts would explain the shallower slopes found in three out of four experiments (the exception is Experiment 5). I shall explore each option in turn.

Why would the pacemaker rate decrease in cause-effect intervals? Because outcomes followed causes in 100% of trials, then one answer is that clock speed is modulated by cause-effect contingency: if an action always produces a certain outcome with 100% contingency, then it becomes possible to reliably predict that pressing a key, for example, will always trigger a visual flash. Therefore, it might be that clock rate decreases as the causal strength between two events increases. One way to test this is to examine PSE and slope coefficients for different levels of cause-effect contingency. One caveat is necessary here though: contingency and predictability are insufficient in themselves to affect clock speed, because noncausal trial events are equally predictable. It is predictability engendered by *causality*, rather than mere association, that I posit as an explanation for clock slowing (Buehner, 2012).

An alternative possibility to clock slowing is that pacemaker rates remain the same in causal and noncausal conditions, but instead, an attentional drift in the former results in fewer pulses accumulated. Temporal binding might arise, for example, because less attention is devoted to the cause-effect interval but more to the outcome stimulus. For unrelated events (noncausal conditions), attention might be equally devoted to the interval and the stimuli marking the interval. The implication of this is that outcomes in causal trials would be perceived as longer. Temporal binding then, might be due to a bias in attention, such that causally related outcomes result in a bottom-up capture of attention. Note, however, that this is merely one suggestion of how attentional drifts might occur. The key point is that any lapse in attention during the interval would also result in fewer pulses accumulated, thus manifesting as a shallower slope. To be clear, though, this is not the same

as an attentional switch account, which posits different switch latencies between causal and noncausal conditions; none of the experiments support this notion, despite a numerical (and non-significant) intercept difference in Experiment 5. My results then, support a slower pacemaker in causal intervals, but drifts in attention *during* the interval should not be discounted.

To conclude this chapter, I have demonstrated that temporal binding is effected by a slower pacemaker rate: causal slopes are shallower than noncausal in three of the four experiments reported in this chapter. This supports the suggestions in Chapter 3, which posited a *specific* clock process. However, to further verify the clock slowing hypothesis, I conducted additional experiments.

5. Chapter 5: Verifying clock slowing in temporal binding with temporal discrimination tasks

The experiments in this chapter exploit the corollary that temporal resolution is necessarily impaired if the pacemaker operates at a slower rate. Specifically, if the contraction of time in temporal binding reflects a decrease in the rate of a specific clock system, then the impaired temporal resolution that follows generates an interesting prediction: The threshold required to temporally discriminate causal intervals should be higher than for noncausal intervals. This is because a slower pacemaker in causal intervals would produce fewer overall pulses, which in turn results in poorer temporal resolution (relative to noncausal intervals). This would not be the case if binding were effected by shifts in the events that delineate the interval (or by a difference in switch latencies). Only a subjective distortion of time *during* the interval could explain temporal discrimination differences between causal and noncausal intervals.

Participants completed a temporal discrimination task, using a similar method to the experiments in the previous chapter. However, rather than compare an extended temporal event to either a causal or noncausal interval, participants in the following experiments discriminated two identical (causal or noncausal) intervals. This is because I wished to explore whether a slower pacemaker also reduces temporal resolution. This cannot be fully ascertained from Experiments 5 – 8 because the same temporally extended stimulus was compared to either causal or noncausal intervals; in order to specifically measure discrimination thresholds in causal and noncausal sequences, it is essential that both intervals (in each condition) be the same. Also, the staircase procedure in the previous chapter's experiments targeted the PSE, whereas the following experiments were designed

specifically to probe to the JND with a higher temporal resolution. Additionally, duration perception is sensitive to stimuli structure, such that intervals with different markers are judged with less precision than intervals with identical markers (Grondin, 2010; Grondin et al., 2005). This might, at least in part, explain the wide dispersion of JNDs in Experiments 5 – 8 (i.e., reference and comparison intervals were marked by different stimuli). Therefore, to reliably derive a causal and noncausal JND, it was necessary to conduct further experiments, ones that (1) specifically target JNDs rather than PSEs, and (2) contain intervals delineated by the same events.

In causal trials participants discriminated two intervals, each delineated by a key press and visual flash. In noncausal trials, participants discriminated two intervals each delineated by two visual flashes. The duration of the first interval in both conditions was fixed for a block, while the second varied from trial to trial. Participants were asked to judge whether the second (variable) interval was shorter or longer in duration than the first (fixed) interval. I then computed the discrimination threshold (JND) for a range of fixed intervals. Larger JNDs in causal trials would indicate a slower pacemaker rate during causal intervals, relative to noncausal.

5.1. Experiment 9

5.1.1. Method

5.1.1.1. Participants

Fourteen students of Cardiff University (13 female, 1 male, M_{age} = 22.8 years, age range: 12-42) participated in exchange for course credit or £4 payment.

5.1.1.2. Apparatus and stimuli

The experiment was implemented in Psychopy (Pierce, 2007) on CRT monitors connected to Apple Mac computers, with resolution of 1280 x 1024 and refresh rate of 120

Hz. In both conditions (causal and noncausal) participants were presented with two intervals. In causal trials both intervals were between a key press and a visual flash (white square, 200 pixels^2), while two visual flashes marked the intervals in noncausal trials. All visual flash stimuli were 50 ms in duration. In causal and noncausal trials the first interval was a fixed duration, while the second always varied in duration from trial to trial. A fixation cross (60 pixels^2) appeared on screen before each interval and was displayed until the interval finished. All stimuli were presented centrally on screen.

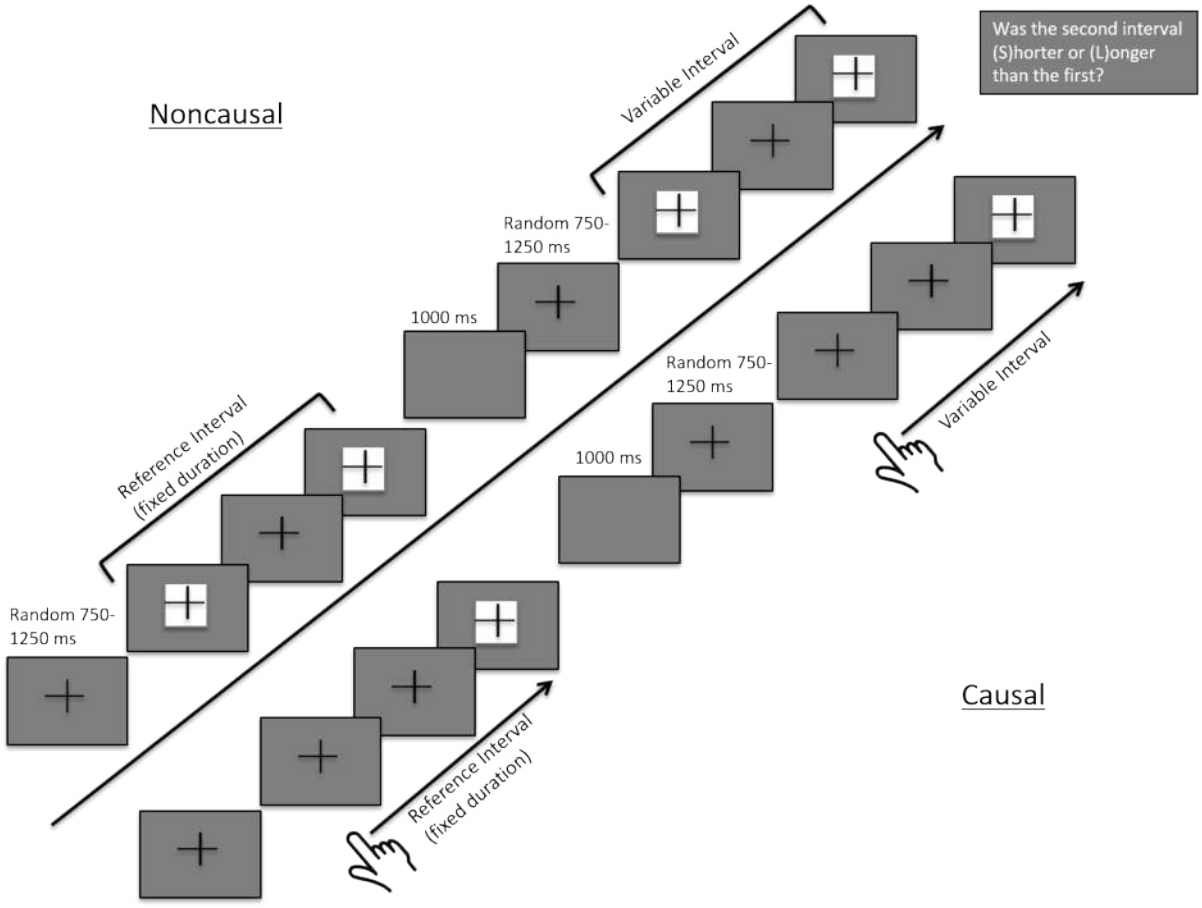

Figure 5.1. Trial structure of Experiment 9. Participants were presented with two intervals in each condition. The duration of the first interval was fixed for a block, while the second was variable. In causal trials, a key press triggered a visual flash after an interval; in noncausal trials, two visual flashes delineated an interval. Participants were asked whether the second interval was shorter or longer than the first.

5.1.1.3. Design and procedure

Two factors were employed in the study: Trial Type (causal, noncausal) and Reference Duration (250, 450, 650 and 850 ms). The dependent variable was the participants' discrimination judgments. I operationalized discrimination as the just noticeable difference (JND), which is half of the difference between 0.75 and 0.25 responding probability from a fitted psychometric function. Participants completed practice blocks of causal and noncausal trials, following the procedure from Experiments 5 – 8 (i.e., minimum 12 trials and 2 reversals per block). Thereafter, participants alternated between causal and noncausal experimental blocks, with one block of each reference duration, comprising roughly 50 trials per block (two interleaved staircases, each ending after a minimum of 25 trials and 4 reversals. See specifics below).

Figure 5.1 shows the trial structure of the experiment. Causal trials began with the display of a fixation cross. Participants then made a voluntary key press, which triggered a visual flash after a fixed interval (250, 450, 650 or 850 ms). During this interval the fixation cross remained on screen, but disappeared immediately following the flash. Participants then experienced an inter-stimulus interval (ISI) of 1000 ms before the fixation cross was again displayed. As in the first interval, participants made a voluntary key press that triggered a visual flash after a random interval (which began with durations of 50% and 150% of the reference duration for each respective staircase. See below for specifics). The fixation cross again disappeared immediately following the flash. Participants were then asked whether the second (variable) interval was shorter or longer than the first (fixed) interval by pressing the S or L keys, respectively. The screen then blanked for a random duration between 1500 - 2300 ms before the next trial began.

Noncausal trials followed a similar procedure. Trials began with the display of a fixation cross, for a random time (750 - 1250 ms). A visual flash then marked the beginning of the interval, followed by another visual flash after a fixed duration (250, 450, 650 or 850 ms). After an ISI of 1000 ms the fixation again appeared (random duration: 750 - 1250 ms,) before the second interval began, which was delineated by two visual flashes. The fixation cross remained on screen throughout each interval, disappearing for the ISI period. As in causal trials, participants were asked to press the S or L keys if they perceived the second interval to be shorter or longer than the first. Following the participants' response, the screen blanked for a random duration between 1500 and 2300 ms.

The experiment began with either a causal or noncausal block of trials (counterbalanced between participants), thereafter alternating between causal and noncausal trial blocks. The duration of the fixed interval was selected at random from one of the four Reference Duration intervals, and remained fixed for each block. The variable interval was controlled by a staircase procedure, using the same Kesten stochastic algorithm as in Experiments 1 - 4 (Kesten, 1958; Treutwein, 1995). However, the convergence thresholds of each staircase were 0.25 and 0.75, respectively, and each block ended after a minimum of 4 reversals and 25 trials per staircase. In other respects, the procedure remained the same as the previous experiments.

Participants first completed practice blocks, with a fixed reference duration of 600 ms. These were shorter than the main experimental blocks and were programmed to end after a minimum of 2 reversals and 6 trials per staircase. After completing practice blocks participants moved on to the main experimental blocks. Throughout the experiment, participants were given a short 2-minute break between blocks.

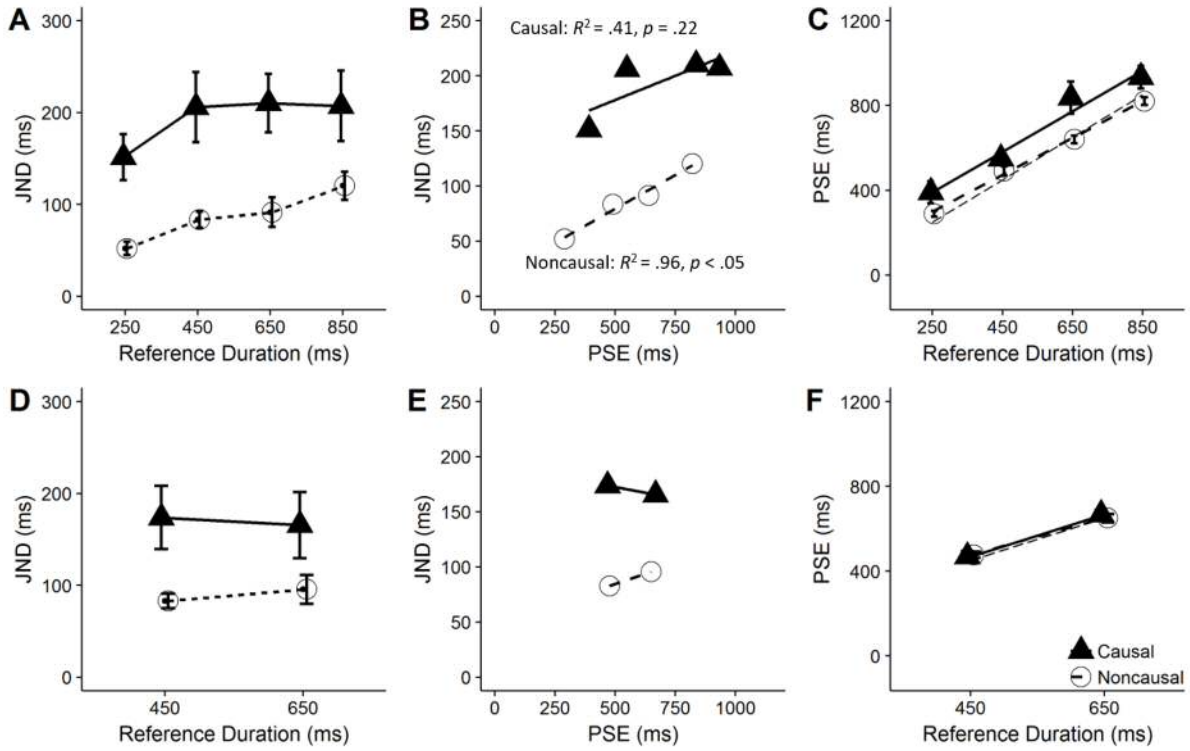

Figure 5.2. Results of Experiments 9 (upper row: A-C) and 10(lower row: D-F). (Left panels: A & D): JNDs are higher in causal conditions, than noncausal, showing that participants found causal intervals harder to discriminate in both experiments. Error bars represent standard error. (Centre panels B & E): JNDs as a function of PSEs. In both experiments, noncausal JNDs increase linearly with PSEs, conforming to the scalar component of timing. Dashed and solid lines show lines of best fit. (Right panels: C & F): In Experiment 9, causal PSEs are overestimated, with respect to the veridical (shown by a lighter/shorter dashed line). This is not the case for Experiment 10.

5.1.2. Results

JND analysis. Control trials (trial 1 of each staircase) were removed before fitting psychometric functions (cumulative Gaussian). Figure 5.2(A) shows mean JNDs for Trial Type, plotted against Reference Duration. JNDs are clearly higher for the causal condition across all reference durations, indicating that participants found discriminating causal intervals more difficult than noncausal. A 2 x 4 (Trial Type [causal, noncausal] x Reference Duration [250, 450, 650,850]) ANOVA corroborated these impressions. I observed a main effect of Trial Type, $F(1,13) = 25.06$, $p < .001$, partial $\eta^2 = .67$, and Reference Duration,

$F(3,39) = 3.79$, $p < .05$, partial $\eta^2 = .23$. However, the Trial Type X Reference Duration interaction was not significant, $F(3,39) = .42$, $p = .74$, partial $\eta^2 = .03$.

Scalar variability. Inspection of Figure 5.2 (A) also shows a slight linear increase in noncausal JNDs as a function of Reference Duration. In contrast, causal JNDs show a linear relationship only between 250 and 450 ms reference durations. Confirmation of the scalar component of SET requires the discrimination threshold to increase in line with to-be-timed durations. To further explore the scalar component, I regressed JNDs onto PSEs for causal and noncausal conditions separately. According to Wearden and Lejeune (2008), the scalar component holds when a high R^2 is obtained (see Chapter 4, and footnote 1 in the same chapter). The results support the impression conveyed by Figure 5.2 (B), with a significant model for noncausal JNDs, $p < .05$, adjusted $R^2 = .96$, but no significant model for causal JNDs, $p = .22$, adjusted $R^2 = .41$. Thus, my results conform to the scalar property of timing only for noncausal intervals.

PSE analysis. A common finding in time psychology is the temporal order effect (TOE), where the second interval is judged as longer than the first, when the intervals are presented in series (Schab & Crowder, 1988). Additionally, in the current experiment reference intervals were always presented before variable duration intervals. I have no a priori reason why this order effect might have differentially affected causal and noncausal intervals, and therefore, conducted an analysis to determine if this occurred. I derived a PSE for each level of Reference Duration. Figure 5.2 (C) displays these results, showing a near perfect linear relation between noncausal PSEs and actual stimulus duration. In contrast, causal PSEs are consistently longer than veridical for all reference durations, suggesting a TOE, i.e., because the second interval required a longer real-time duration in order to be judged as equal in duration to the first. In other words, the second interval was actually

perceived as *shorter* than the first. It is not certain, therefore, to what extent this order effect explains the JND results (I return to this point in the following Discussion section).

5.1.3. Discussion

In line with the hypothesis of a slower clock in temporal binding, I found that causal JNDs were significantly larger than noncausal, for all reference durations tested. Interestingly, I found that discrimination sensitivity increased linearly with reference duration only for the noncausal condition. Discrimination in the causal condition, by contrast, was not linearly related to reference duration; JNDs increased only between the shorter durations tested (250 and 450 ms), but did not change for durations between 450 and 850 ms. Whilst this violates empirical scalar timing (Wearden & Lejeune, 2008; see also footnote in Chapter 4), it is not an uncommon finding, as Rammsayer & Ulrich (2001) found that the pulse rate decreased as the standard duration increased. However, given that mean PSEs increased linearly with Reference Duration, my data then conforms to theoretical scalar timing.

A possible reason for JNDs violating scalar timing is due to differences in task difficulty between the current experiment, and those in Chapter 4. JNDs in the experiments of Chapter 4 all exhibited the scalar component, whilst the current experiment did not. A source of difference between these experiments is the staircase procedure. In the previous chapter's experiments, the target thresholds for the Kesten algorithm (Kesten, 1958) were set at 0.5; in the current Experiment 9, both staircases were set at 0.75 and 0.25, respectively. A target threshold of 0.5 would mean that upward and downward step sizes are equal. However, target thresholds other than 0.5 would result in step sizes that are different in downward than upward motion. Thus, unequal down-up step sizes in the current experiment might have increased task difficulty, relative to experiments 5 – 8.

An unexpected finding concerns the PSEs. While PSEs in noncausal conditions were generally veridical, causal PSEs were larger. This indicates that the second (comparison) causal interval was perceived as shorter than the first (reference) duration. Given that the data conform theoretically to scalar timing (i.e., PSEs increased linearly with reference duration), it is unknown if, and to what extent, this order effect contributes to the JND results. In order, therefore, to rule out this order effect and provide more compelling evidence of differences in the discrimination thresholds of causal and noncausal intervals, I conducted a further experiment. Experiment 10 is a variation of Experiment 9, with the exception that the order of reference and variable duration intervals are randomised from trial to trial.

5.2. Experiment 10

This experiment differed to the first only in terms of the sequence of reference and comparison intervals in each trial. I randomised the location of the reference and comparison intervals and asked participants to judge which interval was the longest. However, because researchers often conduct temporal discrimination studies with few standards (e.g., Rammsayer & Ulrich, 2001), I employed only two reference interval durations. In other respects, the procedure remained the same as in Experiment 9.

5.2.1. Method

5.2.1.1. Participants

Six students of Cardiff University (5 female, 1 male, M_{age} = 26.8 years, age range: 22-33) participated in exchange for £2 payment or course credit.

5.2.1.2. Apparatus and stimuli

This remained the same as the previous experiment.

5.2.1.3. Design and procedure

This remained the same as the previous experiment with a few exceptions. I randomised the sequential position of the reference and variable intervals, so that the reference interval could be presented either first or second in a particular trial. Also, I used two standard durations (450, 650 ms) instead of four. Participants were asked which interval was longest, and responded by pressing either the left or right arrow key to correspond to the first or second interval, respectively. In all other respects the procedure remained the same as above.

5.2.2. Results

JND analysis. As in Experiment 9, I removed the first trial of each staircase before fitting cumulative Gaussian functions. Mean JNDs are shown in Figure 5.2 (D). Consistent with the previous results, JNDs were higher in causal than noncausal conditions. ANOVA supports these results, with a main effect of Trial Type, $F(1,5) = 11.19$, $p < .05$, partial $\eta^2 = .69$. In contrast to the previous experiment I found no effect of Reference Duration, $F(1,5) = .02$, $p = .90$, partial $\eta^2 < .01$, and no Trial Type x Reference Duration interaction, $F(1,5) = .12$, $p = .74$, partial $\eta^2 = .02$.

Scalar variability. Regarding the scalar property, Figure 5.2 (E) also shows a concordant increase in JNDs in line with PSEs for the noncausal condition. By contrast, causal JNDs are inversely related to PSEs, with a lower JND in the 650 condition compared to the 450. In line with the previous experiment then, only noncausal trials conform to the scalar property of time. Because regression is not meaningful with only two reference durations, I restrict my observation to visual inspection only.

PSE analysis. A final and important observation concerns order effects. Given that the intervals in the trial sequence were random (i.e., fixed duration intervals could either be

first or second in a particular trial), I expected PSEs of both conditions to be veridical. Indeed, this is what I find, as Figure 5.2 (F) shows. Thus, the finding that JNDs in causal trials are higher than noncausal, cannot be attributed merely to order effects, and instead, is likely due to poorer temporal resolution, achieved by a different pacemaker rate.

5.2.3. Discussion

The results of the present experiment complement those of Experiment 9. I found that causal JNDs were significantly larger than noncausal. Importantly, this is independent of the order of reference and comparison intervals, with PSEs for both conditions following a linear and veridical trend. Also complementing Experiment 9 is the finding, based on visual inspection, that discrimination sensitivity increased linearly with reference duration in noncausal trials. In contrast, Figure 5.2 (E) shows a decrease in discrimination sensitivity from 450 ms to 650 ms in causal trials. As noted in the Discussion of Experiment 9, this might have occurred because of the difference in staircase parameters between the current chapter's experiments and those in Chapter 4; the different step sizes might have differentially affected Trial Type levels, such that discrimination in causal trials violates scalar timing. This finding may then simply reflect task difficulty, such that greater attention is required to discriminate intervals with smaller step sizes (Rammsayer & Ulrich, 2001). This is particularly bolstered by the finding that arousal (e.g., which here can be attributed to task difficulty) increases subjective time by increasing pacemaker rate (Gil & Droit-Volet, 2012; Mella, Conty, & Pouthas, 2011). However, given that mean accuracy is preserved, one can argue that the data conform theoretically to scalar timing. Importantly though, the higher JNDs in causal trials support a slower clock (and poorer temporal resolution) in temporal binding.

5.3. Experiments 9 – 10: Discussion

In two experiments I have found higher JNDs when discriminating causal, than noncausal intervals. Also, I find that JNDs in the noncausal condition increase linearly with reference duration, therefore conforming to the scalar component of timing (Wearden & Lejeune, 2008). Interestingly, JNDs in the causal condition do not show this scalar component. Instead, the JNDs only differ between 250 and 450 ms, thereafter remaining the same. What do these results say about internal clock process in temporal binding? First, they suggest that cause-effect intervals are served by a slower clock than noncausal intervals. Second, the rate of the clock is not systematically slower in causal than noncausal intervals, but varies with the duration to-be-timed; timing is more sensitive at longer (e.g., 850) than shorter durations (e.g., 450 ms). Third, the higher JNDs in causal conditions support the hypothesis that a slower pacemaker necessarily leads to poorer temporal resolution during causal (binding) sequences.

The obvious question is: why would pacemaker rate vary only in cause-effect intervals? One possibility is that pacemaker rate varies in order to compensate for timing difficulty at different durations. For example, if short durations are harder to discriminate than long durations, then pacemaker rate should vary in order to compensate for this: a faster pacemaker rate would improve discrimination at short intervals (Killeen, 1992; Killeen & Fetterman, 1988). Of course, this would have affected causal and noncausal conditions equally. So the question is why this affects only causal intervals. A tentative possibility is that discriminating causal intervals is sufficiently harder than noncausal intervals. This suggests that pacemaker rate fluctuates in order to compensate for poorer discrimination at certain durations. However, there is reason to believe that this is simply an artefact of the experiment, rather than a finding unique to causal interval discrimination, for the following

reasons: (1) temporal discrimination tasks, in general, usually find that the CV (standard deviation/mean) is greater at durations below 100 ms (Killeen & Fetterman, 1992), thereafter remaining constant; and (2) the experiments in the previous chapter (Experiments 5 - 8) found that JNDs increased linearly with PSEs. The reason that causal trial JNDs in the current experiments do not exhibit this scalar component is likely due to the parameters of staircase procedure; response thresholds of 25% and 75% result in smaller step sizes than the parameters used in Experiments 5 - 8 (where target thresholds of 50% were employed). Hence, the smaller step sizes might simply have increased the task difficulty (Wearden & Lejeune, 2008), which can affect the scalar property.

Can other clock processes, for example switch latencies, explain the results? Consider that the latencies between the opening L_o, and closing L_c, are greater in causal than noncausal intervals. This cannot account for my results because the latencies would presumably be equal in both (reference and comparison) intervals. In other words, comparing intervals where the subjective start and end points are the same, would not result in different JNDs between causal and noncausal conditions. The same argument applies to a sensory recalibration perspective (Eagleman & Holcombe, 2002); the shifts would result in equally perceived intervals. It is difficult to see how causality would affect JNDs, if the intervals were effected by subjective differences in the start and end point. To account for my results therefore, there has to be a change in time perception *during* the causal interval.

The only reasonable explanation then, is that pacemaker rate slows down in causal intervals. A slower pacemaker would mean that fewer pulses are emitted, which decreases temporal acuity. The key question is what the pulses actually represent. One suggestion is that the pulses themselves represent a neural pacing signal (Grondin, 2010; Wenke &

Haggard, 2009), such that a slower pacemaker rate reflects a decrease in neural activity. The exact locus of this activity is unknown, with researchers often claiming that time perception is distributed throughout the brain (Grondin, 2010). However, establishing the locus of time perception in the human brain is beyond the scope of this thesis. The main questions, regardless of the locus of time perception, are whether clock pulses represent neural activity, and how this activity results in changes to pacemaker rate in causal sequences.

In an attempt to answer the questions posed above, one might consider Gestalt grouping principles (von Ehrenfels, 1890/1988). There are two reasons for this: One, is that durations are perceived as shorter when grouped by similarity (Zhou, Yang, Zhang, Zhang, & Mao, 2015), and two, neural activity appears reduced when items are perceptually grouped (Peterson, Gozenman, Arciniega, & Berryhill, 2015). If causality is a perceptual grouping factor, then it might explain why causal intervals are perceived as shorter, and why the decreased pacemaker rate implies less neural activity. I will elaborate more on this argument in the following chapter. For now however, it is sufficient to say that (1) causal intervals may be perceived as shorter, and (2) the rate of an internal clock during such intervals is slowed, because causality may be a Gestalt grouping factor, i.e., causally related events appear closer together because causality itself might be a grouping principle.

In conclusion, the experiments in the current chapter rule out explanations based on subjective shifts in the stimuli marking the interval. Instead, the results are in line with those of the previous chapter, which support a slower pacemaker in temporal binding episodes.

6. Chapter 6: General Discussion

Temporal binding may be characterised as a contraction of perceived time during cause-effect intervals. Since the first publication by Haggard et al. (2002), there has been a good deal of discussion of the factors that give rise to, and modulate the effect. A notable topic of contention has been debated about the factors that give rise to temporal binding – whether intentional action or causality are necessary for binding to occur (Buehner, 2012; Buehner & Humphreys, 2009; Engbert et al., 2008). Discussion has also centred on whether binding can be considered an implicit measure of agency (Christensen, Yoshie, Di Costa & Haggard, 2016; Engbert et al., 2008; Moore et al., 2009). The outcome of these academic debates has shown unequivocally that causality, and not intentional action, is necessary for binding (Buehner, 2012). Furthermore, referring to binding as an implicit measure of agency is moot, because no correlations have been reported between measures of temporal binding and explicit agency judgments (Dewey & Knoblich, 2014; Ebert & Wegner, 2010)[4]. However, despite inconsistencies, research into temporal binding has been fruitful since Haggard et al. (2002).

One area that has received little attention, though, concerns the mechanisms that effect binding, i.e., *how* does binding occur? There are two dominant theoretical underpinnings to consider: on one hand, event perception approaches posit that temporal binding is a product of sensory realignment (Eagleman & Holcombe, 2002). On the other hand, temporal binding might be explained by changes to our perception of time (Wenke & Haggard, 2009). The aim of this thesis was simply to explore whether changes in time

[4] Haggard (e.g., Christensen et al., 2016) often refers to binding as a sense of agency because of the orthogonal relationship between binding and explicit agency measures: shorter perceived intervals are usually associated with stronger ratings of self-agency, even though these measure do not correlate. Whether or not binding should be referred to as a sense of agency is likely to become a debatable topic in the future publications.

perception actually occur. To this end, I applied the most popular model of time perception, the SET model (Gibbon, Church, & Meck, 1984), to temporal binding. Internal clock models explain changes in time perception either via fluctuations in the rate of a pacemaker, or via different switch latencies. There is also evidence that time perception is not 'general', but is stimulus specific (Pariyadath & Eagleman, 2007); that is, changes in temporal resolution for one stimulus might not necessarily affect another simultaneously presented stimulus. Therefore, an investigation into to the contraction of time in binding should consider not only the different clock processes but also whether such processes are general or specific.

In applying an internal clock to temporal binding, I first explored whether temporal binding is effected by general or specific clock processes (Experiments 1 - 4, Chapter 3). In four experiments, participants judged the duration of intervals marked by different events: in the causal condition, the intervals were marked by a key press and either a visual or auditory outcome, while noncausal trial intervals were marked by a visual stimulus and either a visual or auditory outcome. Participants also judged the duration of an additional event embedded into certain trials: in one-third of trials a stimulus (auditory or visual) was scheduled to occur *before* the interval, in another one-third the stimulus was scheduled to occur *during* the interval, and in a further one-third the stimulus was *not* scheduled to occur. If the contraction of time in temporal binding is served by changes to a general clock process, specifically a general decrease in pacemaker rate, then it would effect events embedded during cause-effect intervals *in addition* to the overall cause-effect interval; intervals and events embedded during causal intervals would both be judged as shorter than those in noncausal conditions. However, if temporal binding is effected by changes to switch latencies or by perceived shifts in the time of the events marking causal intervals, then embedded event judgments ought not to differ between causal and noncausal

conditions. The results were uniform across all four experiments: (1) intervals were judged as shorter in causal than noncausal trials; (2) embedded events were judged the same in causal and noncausal trials. Importantly, interval judgments were not affected by the presence of embedded events. That is, the binding effect occurred regardless of the presence or location of embedded events. Also, embedded events were judged the same, whether presented before or during the interval. Together, this suggested that temporal binding was not effected by a general pacemaker slowing. Instead, I concluded that binding is affected either by differences in switch latencies or by clock processes specific to cause-effect intervals.

Second, I explored whether temporal binding is indeed effected by changes in pacemaker rate or by switch latencies (Experiment 5 - 8, Chapter 4). Following a procedure common in time perception research, that of conducting regression analyses to disentangle switch latencies and pacemaker rate changes, I investigated the slopes of various interval durations in causal and noncausal intervals. Participants judged whether a comparison duration interval was shorter or longer than a reference duration interval. In causal conditions, reference intervals were marked by a key press and flash, while noncausal trial reference intervals were marked by two flashes. In both conditions the duration of the intervals was fixed for a block. The comparison interval, meanwhile, varied from trial to trial and consisted of a single visual stimulus. I computed the PSE for a series of durations and conducted regression analyses to disentangle possible switch latency differences from pacemaker rate changes. In three out of four experiments (Experiments 6 - 8) I found a significant difference between causal and noncausal slopes, but no intercept difference. These results supported a slower pacemaker account of temporal binding, thus concluding

that temporal binding is effected by specific clock process dedicated to tracking cause-effect sequences.

Third, I verified the findings of Experiments 6 - 8 by examining the temporal discrimination thresholds within causal and noncausal conditions (Experiments 9 - 10, Chapter 5). If the pacemaker operates at a slower rate in cause-effect intervals then temporal resolution should be poorer, resulting in higher discrimination thresholds. Participants temporally discriminated two key press-flash intervals in causal trials, and two flash-flash intervals in noncausal trials (using a reference and comparison method similar to Experiments 5 - 8). The results showed that JNDs for a range of durations were higher in causal than noncausal trials, supporting a slower pacemaker in cause-effect intervals. Thus, the present body of work finds that temporal binding is effected by changes to the rate of an internal clock.

6.1.1. Is the compression of time around motor commands different to temporal binding?

Research shows that time perception is affected by the mere presence of motor action. Tomassini et al. (2014) have shown a perceived compression of time around a moment of action. Participants received tactile stimulation (consisting of two pairs of taps) to either a moving hand or a static hand. The interval between the reference pair was fixed, while the interval between the comparison pair of taps varied from trial to trial. By scheduling the reference pair of taps to occur before and after hand movements, Tomassini et al. were able to identify whether perceived time changed as a result, or in anticipation of, hand movements. Their results showed a subjective compression of time just prior to and immediately following hand movements. Importantly, this compression was effector-specific, as compression was found only for the hand that moved. This complements the results of Wenke and Haggard (2009), who also found evidence for effector-specific clock

slowing in temporal binding. However, a key difference between Tomassini et al., and Wenke and Haggard is that no causal relation existed in the former. Instead, participants simply moved their hand, with no causal outcome (i.e., the hand movements did not cause the taps). Motor actions in Wenke and Haggard, in contrast, always triggered an auditory outcome after an interval (in addition to triggering tactile stimuli). If the contraction of time in Tomassini et al. occurs between two tactile stimuli, then what accounts for the contraction between a tactile and non-tactile stimulus (e.g., a key press and tone)? Presumably, a tactile-specific contraction would not occur between tactile and non-tactile stimuli. Therefore, the results of Tomassini et al. cannot explain temporal binding, in which intervals are often delineated by a key press and auditory or visual outcome. I propose that Tomassini et al.'s finding, and those of Wenke and Haggard simply reflect a process independent of temporal binding, one that is specific to motor commands and not affected by causality.

6.1.2. Is clock slowing driven by causality or intentional action?

One issue with the present experiments is that causality and intentional actions were compounded. While noncausal intervals were always delineated by either auditory or visual stimuli (or a combination of both), causal intervals were always delineated by intentional action and auditory/visual stimuli. How can we be sure that the results are driven by causality and not merely by intentional action? Humphreys and Buehner (2009) were able to isolate intentional action and causality. In their study, participants had to synchronise a key press with a tone (t1) that marked the start of the interval. The authors employed stimulus anticipation to measure binding, which involved pressing a key to anticipate the outcome stimulus. In causal trials, the key press at time t1 was causally related to the outcome stimulus. However, in noncausal trials, neither the key press nor the tone (t1) caused the

outcome stimulus. The authors showed a temporal binding effect only when the action at time t1 caused the outcome stimulus. In trials where the key press (t1) did not cause the outcome stimulus, no binding effect was found. Thus, the study demonstrated that intentional action itself is not sufficient. Instead, it is the causal link between action and outcome that is necessary for temporal binding, over and above association. In the current experiments, stimuli marking causal intervals were related by causality, whereas stimuli marking noncausal intervals were related by mere association. Thus, one likely interpretation is that causality, and not intentional action is driving clock slowing.

To be certain that pacemaker rate is driven by causality, in the present experiments, however, further research is advisable. Given that Tomassini et al. (2014) found a contraction of time surrounding hand movements only, it would be beneficial to replicate Experiments 5 – 8 with a different causal manipulation. One solution is to replace the key press with the mechanical apparatus from Buehner (2012). Finding shallower slopes in mechanical-causal, compared to noncausal conditions, would support the notion of causality induced changes to pacemaker rate.

6.1.3. Clock slowing in temporal binding is specific

The evidence from the current experiments supports the notion of a slower specific clock in temporal binding episodes. This is in line with work by Eagleman and colleagues (Paryiadath & Eagleman, 2007, 2008) who found that time is not one entity (Eagleman, 2008). In a series of studies, the authors used the oddball paradigm to assess whether subjective temporal expansion affects all simultaneously presented stimuli, or whether it is specific to the oddball stimulus only. In the oddball illusion, a sequence of stimuli is presented, in which one stimulus is markedly different from the other stimuli (e.g., presenting an image of a car in a series of strawberry images). Eagleman found that the

subjective expansion of time for the oddball stimulus did not generalise to additional, simultaneously presented stimuli, such that a tone stimulus showed no expansion effect. This was also the case when the images where flickered; a perceived expansion of time should result in a decreased flicker frequency threshold (and better temporal acuity). Their results, however, showed no such increase in flicker frequency acuity (see also Stetson et al., 2007). Eagleman suggests that time perceptions function much like visual perception, "in which it is understood that vision emerges as the collaboration of many subpopulations that code for different aspect of scenes (motion, position, colour, and so on)" (Eagleman, 2008, p. 134). Eagleman further suggests that duration judgments, flicker rates, and simultaneity judgments (among others) are underpinned by mechanisms that usually work in concert but can be dissociated in laboratory settings. Wenke and Haggard (2009) is an example of a study that employed different measures: intervals were measured with verbal estimation, while discrimination thresholds were measured with simultaneity judgments. According to the logic of Eagleman (2008) then, the results of the discrimination judgments might be independent of the processes governing the contraction of the interval. Similarly, the embedded events in my Experiments 1 – 4 are judged by processes independent of the contraction in the causal intervals. Given the results of the present experiments, in addition to those of Eagleman, the implication is that the notion of specific clocks is not unique to temporal binding, but is a product of time perception in general.

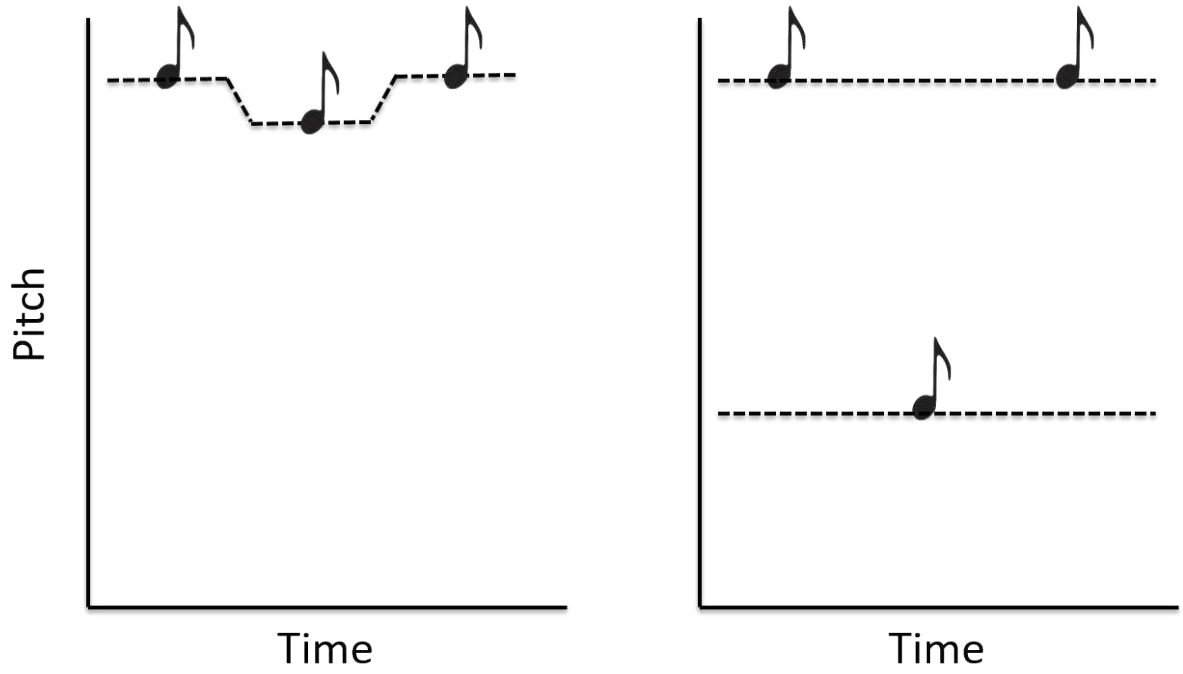

Figure 6.1. Sequences are composed of three-tone segments, which are repeated for a few seconds (e.g., the pattern 'up-down-up, up-down-up', and so on, is heard). The sequences are perceived differently depending on pitch difference. When the pitch is similar a single stream is heard, whereas increasing the pitch difference results in two simultaneous but distinct streams. The dashed line represents the object streams (based on Dannenbring and Bregman, 1976).

The logical question now is to ask what determines the specificity of time perception. It is not sufficient to simply conclude that duration judgments are governed by a specific group of neurons, distinct from those involved in other temporal judgments (e.g., simultaneity[5]), because duration judgments were required both for intervals and embedded events in Experiments 1 - 4. Also, it is not sufficient to conclude that causality engages a specific clock. One must ask *how* and *why* causality would engage a specific clock. One novel suggestion is that time perception is governed by the perceptual organization of the stimuli,

[5] It is not obvious whether internal clock models can be applied equally simultaneity and interval estimation judgments. The latter require explicit judgments of duration, whereas the former involve and implicit awareness of duration: judging whether two stimuli are consecutive or simultaneous (in time) undoubtedly necessitates implicit awareness of the duration separating the stimuli.

such that items grouped by similarity are more likely to be judged by a single independent clock. My chief motivation for this suggestion is that temporal judgments differ, depending on the relatedness of one stimulus to another set of stimuli. For example, van Noorden (1975) used the stimuli from Dannenbring and Bregman (1976; Bregman, 1990) in which a sequence is constructed from three tones, the second of which is lower in pitch than the outer tones (see Figure 6.1). The classic finding, from Dannenbring and Bregman (1976), is that the tones are perceived as a single auditory stream when the second, lower-pitched tone, is similar in pitch to the outer tones (Figure 6.1, left panel). When the second tone is significantly different to the outer tones then two simultaneous streams are heard (Figure 6.1, right panel). In other words, the perceptual grouping of the stimuli is differentially affected by the similarity of the tone pitches: when the pitches are similar the sequence 'up-down-up' is heard, whereas a substantial difference in pitch results in two simultaneous streams ('up-up' and 'down--down'). Van Noorden built upon this finding to determine whether this perceptual grouping affects time perception. To achieve this, van Noorden used a variant of Dannenbring and Bregman's stimuli, whereby the second stimulus was offset, so that its occurrence was delayed with respect to the midpoint between the outer tones (Figure 6.2). He found that the offset was only detected if the second stimulus differed significantly in pitch from the outer tones; when the pitch difference was small, the offset stimulus was actually perceived as equidistant between the outer tones. How does this relate to time perception? If a single timer was used to make judgments about the temporal offset of the second stimulus, then adjusting the pitch should not have affected the results (i.e., the offset should have been detected regardless of pitch). However, the fact that the temporal offset was only detected when the streams were segregated, suggests the involvement of multiple clocks. Indeed, it is difficult to explain van Noorden's results

without invoking multiple clock systems. One caveat however: I acknowledge that van Noorden (1975) asked participants to judge the time of the event occurrence, rather than its duration. While internal clock models are traditionally associated with duration judgments, I nevertheless feel that van Noorden's evidence is sufficiently supportive of multiple timers, rather than a single clock system. Analogously to temporal binding then, which is measured with event and duration judgments, I posit that van Noorden would have found similar results with duration judgments: for example, the second stimulus might be perceived as different (shorter or longer) when perceived in the same stream, and thus effected by a single clock, relative to the when perceived as separate groups. To be clear though, I posit perceptual grouping as a basis for judging the duration of stimuli according to single or multiple clocks; I am not suggesting that perceptual grouping is itself a theory of time perception. My argument, then, is that time perception is specific to the perceptual groups, such that similar items are grouped together and their temporal judgments effected by a single clock.

My claim of a Gestalt basis of time perception is further strengthened by evidence from Zhou et al. (2015). Zhou et al. found that temporal judgments of visually presented stimuli are also affected by similarity, such that similar items are judged as shorter than dissimilar. This merely demonstrates that a Gestalt basis (of time perception) is not limited to the auditory domain. Additionally, Rajendran, Harper, Willmore, Hartmann, & Schnupp (2013) used the stimuli from Dannenbring and Bregman (1976) to study the effects of temporal uncertainty on perceptual grouping. In their study the second stimulus was temporally jittered from trial to trial, so that its exact temporal onset was unpredictable. They found that this temporal jitter led to multiple simultaneous groups, such that the jittered stimulus and the outer stimuli were perceived as separate streams (relative to trials

in which no jitter was applied). Thus, temporal expectancy and similarity affect the perceptual grouping of stimulus sets; a temporally unexpected stimulus will be grouped separately to a temporally expected stimulus. Also, stimuli that are dissimilar (along a particular dimension) will be grouped separately.

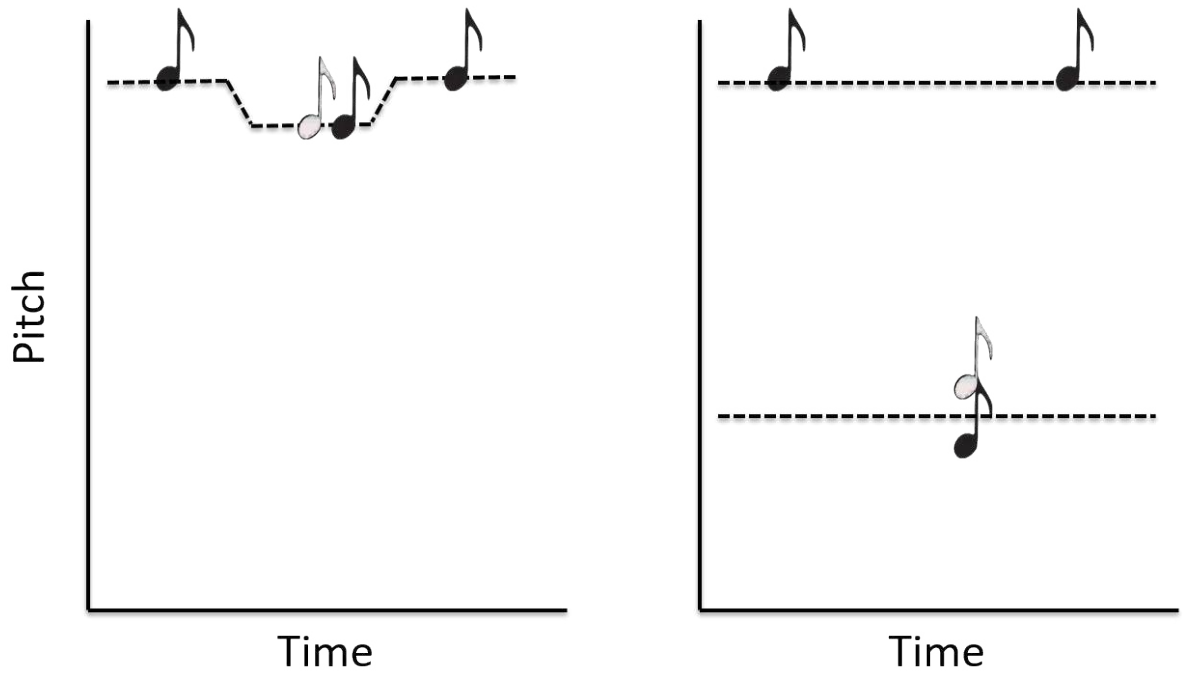

Figure 6.2. Schematic of the stimuli in van Noorden (1975). The temporal offset is perceived when the streams are segregated (i.e., ungrouped: right panel). The white-coloured tone denotes the perceived time of the offset stimulus.

As an interim summary I propose the following: auditory stimuli of similar pitch are grouped into the same perceptual stream (Bregman, 1990). Temporal uncertainty is an additional factor that affects perceptual grouping, such that stimuli occurring at unpredictable times are grouped separately to predictably occurring stimuli (Rajendran et al., 2013). Perceptual grouping affects temporal awareness, such that judging whether a stimulus is placed equidistantly between outer stimuli is dictated by pitch similarity; a

temporal offset is not detected when the stimulus is integrated into the same stream as the outer stimuli (van Noorden, 1975).

How does this relate to temporal binding? I propose that causality itself is a grouping factor. This is because (1) temporal unpredictability and similarity affect the perceptual grouping of events (Bregman, 1990; Rajendran et al., 2013); (2) grouped items are timed by independent clocks (Dannenbring & Bregman, 1976); and (3) similar items are judged as shorter than dissimilar items (Zhou et al., 2015). Strictly speaking, this is not a novel idea. There is a rich literature documenting the categorization of objects based on causal information (Sloman & Lagnado, 2005; Sloman, Love, & Ahn, 1998). The main premise of this literature is that objects are assigned to a category if they share the same causal features (Hayes & Rehder, 2012). One study, for example, presented a fictitious animal (a "Taliboo") to 7 – 9 year olds (Ahn, Gelman, Amsterlaw, Hohenstein, & Kalish, 2000). The children were told that the Taliboo had promicin in their nerves, which caused thick bones and large eyes. Ahn et al. then presented the children with two test Taliboos, each missing a feature: one missed the cause (i.e., promicin), and the other one of the outcome features (thick bones or large eyes). When asked which animal made the most likely Taliboo, the children preferred the test animal with a missing outcome feature over the animal with the missing cause. In other words, children judged the animal with preserved causal information (in which promicin caused one of the outcome features), as a more likely Taliboo candidate than the animal without any causal information. Thus, this demonstrates that causality influences category membership, such that objects with a similar cause are grouped together.

In line with the above, I propose that causality is not only a grouping factor that affects category membership, but might also affect temporal judgments. Specifically, the

interval between items that are causally related is not only likely to be judged as shorter than unrelated (along the lines of Zhou et al., 2015), but will be judged by the same clock process. One way to test my hypothesis is to conduct a variant of the embedded event estimation procedure, but manipulate the causal relation between the embedded events and the interval. This might be achieved by varying the key press-embedded event contingency: rather than targeting a 50-50 balance for events presented *before* and *during* the interval, an experiment might use an additional 25-75 balance, such that more events occur during the interval. Because outcomes always follow causes (Hume, 1739/1888), there would be a greater likelihood of embedded events being grouped with the intervals, either by causality or associationism (in causal and noncausal conditions, respectively). If causality (or simply association) is a grouping factor that affects the operation of single or multiple clocks, then a contingency x trial type interaction would be expected: embedded events should be judged as shorter in causal versus noncausal trials, with events in the 25-75 condition being judged as shorter than those in the 50-50 contingency condition. Additionally, embedded events should also be judged as shorter in 25-75 than 50-50 conditions, in causal and noncausal trials. This is because the embedded event would be associated with the cause (i.e., key press or first interval-marker), assuming that association truly is a grouping factor. The key point I propose is that temporal binding is selectively effected by a specific clock process because causality itself is a grouping factor.

6.1.4. Outcome predictability and attentional drifts

One might ask how the pacemaker might slow down, when no information about causality is given until the outcome has occurred, at which point the pacemaker has stopped emitting pulses. A possibility I discussed in chapter 4 is that pacemaker rate might be modulated by causality. Because contingency is a key determinant of perceived causality

(Allen & Jenkins, 1979; Cheng, 1997), then differing levels of contingency would affect outcome predictability. Studies using the Libet clock method (Libet et al., 1983) have shown that binding is driven by predictive, in addition to retrospective, components (Moore & Haggard, 2008; Moore et al., 2009). Moore and colleagues contrasted a predictable condition (75% key press-outcome contingency) with an unpredictable condition (50% key press-outcome contingency). Their results showed that the reported time of action shifted towards expected (but absent) outcome tones in the 75% compared to the 50% condition. In other words, participants experienced a binding effect when when no outcome actually occurred. This demonstrates that binding is driven by predictability, in addition to retrospective awareness. Would predictability directly modulate clock speed? Manipulating cause-effect contingency is a way to measure this. For instance, if contingency is linearly related to pacemaker rate, such that increased contingency results in higher pacemaker rates, then two predictions follow: (1) a regression of perceived onto actual intervals would reveal shallower slopes as pacemaker rate decreases; (2) discrimination thresholds of two causal intervals (e.g., the paradigm in Experiments 9 - 10) would increase as pacemaker decreases, as a result of poorer temporal resolution. However, this paradigm would necessitate a different approach regarding staircases, because this procedure would require intervals that are marked by two events. An alternative, instead of including trials where no outcome occurs, is to manipulate the contingency with which an event causes an outcome; on certain trials, for example, an outcome might follow a key press without having been caused by it. Additionally, it may be that predictability, and not causality, would affect pacemaker rate. In this case, slope differences might be observed between various contingency levels in noncausal trials. In contrast, if causality drives the pacemaker rate,

then slope differences would only manifest in causal trials. In sum then, I propose that pacemaker rate might be modulated by cause-effect contingency.

Instead of clock slowing *per se*, one option is that pulses are missed due to an imbalance of attention in causal versus noncausal trials. For example, fewer pulses might accumulate in causal intervals because more attention is focused on the outcome stimulus. This would result in shorter perceived cause-effect intervals but a longer perceived outcome stimulus, relative to noncausal intervals. As noted in chapter 4, a way to test this hypothesis is to measure the subjective duration of an outcome stimulus, in addition to the action-outcome interval. More attention to the outcomes in causal trials would results in longer perceived outcome durations than in noncausal trials. Thus, temporal binding might arise due to attention being focused *away* from the interval and *toward* the outcome. To date, no research has investigated this. Empirical work then, is necessary to establish whether our results occur due to an imbalance of attention in causal versus noncausal trials.

6.1.5. Reconciling event with time perception theories

It is important to note that event and time perception accounts of temporal binding are not mutually exclusive. Sensory recalibration (Eagleman & Holcombe, 2002; Parsons et al., 2013) proposes a realignment of sensory streams, such that the delay between an action and outcome can be adapted to, bringing the events closer together in time. Time perception theories would account for binding by a change in temporal resolution during the cause-effect interval, such that the rate of an internal clock is slowed down. The present experiments support this latter notion, particularly Experiments 9 - 10, which found that the slower clock in cause-effect intervals affect temporal resolution; the higher discrimination thresholds in causal intervals supported the notion of a slower clock, relative to the thresholds in noncausal intervals. It is difficult to see how Eagleman's sensory recalibration

approach can explain these findings; a shift in the perceived time of events marking the intervals would not affect discrimination judgments. Instead, the results of the JNDs (Experiments 9 - 10) can only be explained by a change in our temporal acuity during cause-effect intervals. Does this mean that sensory recalibration is redundant? Not necessarily, as the two approaches can be reconciled. For example, a slower pacemaker would result in a contraction of time, and thus, the outcome being perceived as earlier. However, one problem with this hypothesis is that it assumes time perception to be responsible for the perceived change in events, i.e., that changes in time perception somehow *cause* the perceived shift in events. Alternatively, it may be that perceived event shifts and time perception changes are represented by the same underlying process; event and time perception approaches manifest from a common cause. More empirical work is needed to investigate the relation between subjective event and time perception changes. For now, the point is wish to stress is that dismissing either internal clocks or sensory recalibration approaches is naïve: research should determine the commonality between event and time perception methods.

6.1.6. Conclusion

Temporal binding is affected by the same principles that govern causal inference. Contiguity (Shanks, Pearson & Dickinson, 1989) and contingency (Allen, 1993) both modulate the strength of the binding effect. The purpose of this thesis, however, was to determine *how* binding is effected. That is, what are the processes that give rise to temporal binding. Specifically, I investigated whether changes in time perception underlie binding, by focussing on the pacemaker rate of internal clocks. My findings have demonstrated that temporal binding is effected by a slower pacemaker. This is in line with previous studies (Humphreys & Buehner, 2009) that demonstrated a slower pacemaker in causal intervals.

The results are also in line with a larger body of evidence implicating pacemaker rate in temporal illusions in general (e.g., Penon-Voak et al., 1996; Wearden et al., 1998). Future research is necessary to investigate (1) whether the results generalise to other causal scenarios, such those involving mechanical causal action (Buehner, 2012), and (2) the extent to which changes in the magnitude of the temporal binding effect is rooted in causality, such as whether pacemaker rate is affected by cause-effect contingency. To assume a broader perspective, research into time perception and perceptual grouping would provide a great and necessary insight into the recent notion of multiple clocks (Klapproth, 2011). A potential start would be to determine whether perceptual grouping underpins the amount of clocks used.

References

Ahn, W., Gelman, S. A., Amsterlaw, A., Hohenstein, J., & Kalish, C. W. (2000). Causal status in determining feature centrality. *Cognition, 69*, 135-178.

Allan, L.G. (1993). Human contingency judgments: rule based or associative? *Psychological Bulletin, 114,* 435-448.

Allan, L. G. (1998). The influence of the scalar timing model on human timing research. *Behavioural Processes, 44,* 101-117.

Allan, L., & Gibbon, J. (1991). Human bisection at the experimental mean. *Learning and Motivation, 22*, 39-58.

Avni-Babad, D., & Ritov, I. (2003). Routine and the perception of time. *Journal of Experimental Psychology: General, 132*(4), 543-550.

Bar-Haim Y., Kerem A., Lamy D., Zakay D. (2010). When time slows down: the influence of threat on time perception in anxiety. *Cognition and Emotion, 24*, 255–263.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

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

Block, R. A., & Gruber, R. P. (2014). Time perception, attention, and memory: a selective review. *Acta Psychologica, 149*, 129-133.

Bregman, A. S. (1990). Auditory Scene Analysis. Cambridge, MA: MIT Press.

Brown, S. W. (2006). Timing and executive function: Bidirectional interference between concurrent temporal production and randomization tasks. *Memory & Cognition, 34*, 1464-1471.

Brown, S. W. (2008). Time and attention: Review of the literature. In S. Grondin (Ed.), *Psychology of time* (pp. 111-138). Bingley, U.K.: Emerald Group.

Brown, S. W., & Merchant, S. M. (2007). Processing resources in timing and sequencing tasks. *Perception & Psychophysics, 69*, 439-449.

Bryce, D., Seifried-Dubon, T., & Bratzke, D. (2015). How are overlapping time intervals perceived? Evidence for a weighted sum of segments model. *Acta Psychologica, 156*, 83-95.

Buehner, M. J. (2012). Understanding the past, predicting the future: causation, not intentional action, is the root of temporal binding. *Psychological Science, 23*(12), 1490-1497.

Buehner, M. J. (2015). Awareness of voluntary and involuntary causal actions and their outcomes. *Psychology of Consciousness: Theory, Research, and Practice, 2*(3), 237-252.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*(10), 1221-1228.

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgment of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology, 56A,* 865-890.

Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking & Reasoning, 12*(4), 353-378.

Buhusi, C. V., & Meck, W. H. (2009). Relativity theory and time perception: Single or multiple clocks? *PLoS ONE, 4*(7), e6268.

Chen. L., & Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Attention, Perception, & Psychophysics, 75*(5), 790-811.

Cheng, P. W. & Buehner, M. J. (2012). Causal Learning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 210-233). Oxford, England: Oxford University Press.

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.

Cornsweet, T. N. (1962). The staircase-method in psychophysics. *American Journal of Psychology, 75*, 485-491.

Cousineau, D. (2005). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*, 42-45.

Cravo, A. M., & Baldo, M. V. C. (2008). A psychophysical and computational analysis of the spatio-temporal mechanisms underlying the flash-lag effect. Perception, 37(12), 1850-1866.

Cravo, A. M., Claessens, P. M. E., & Baldo, M. V. C. (2011). The relation between action, predictability and temporal contiguity in temporal binding. *Acta Psychologica, 136*(1), 157-166.

Cravo, A. M., Haddad, H., Claessens, P. M. E., & Baldo, M. V. C. (2013). Bias and learning in temporal binding: intervals between actions and outcomes are compressed by prior bias. *Consciousness and Cognition, 22*(4), 1174-1180.

Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America, 34,* 582-593.

Dannenbring, G.L. & Bregman, A.S. (1976). Stream segregation and the illusion of overlap. *Journal of Experimental Psychology: Human Perception and Performance, 2, 544-555*.

De Gelder, B., & Bartelson, P. (2003). Multisensory integration, perception and ecological validity, Trends in Cognitive Sciences, 7(10), 460-470.

Dennet, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Desantis, A., & Haggard, P. (2016). Action-outcome learning and prediction shape the window of simultaneity of audiovisual outcomes. *Cognition, 153*, 33-42.

Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition, 20*(4), 1211-1220.

Dewey, J. A., & Knoblich, G. (2014). Do implicit and explicit measures of the sense of agency measure the same thing? *PLoS ONE, 9*(10), e110118.

Dixon, N. F., & Spitz, L. (1980). The detection of audiovisual desynchrony. *Perception, 9*(6), 719-721.

Droit-Volet, S., Fayolle, S. L., & Gil, S. (2011). Emotion and time perception: effects of film induced mood. Frontiers in integrative Neuroscience, 5(33), 1-9.

Droit-Volet, S., Brunot, S., & Niedenthal, P.M. (2004). Perception of the duration of emotional events. *Cognition and Emotion, 18*, 849-858.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11,* 504-513.

Eagleman, D. M. (2008). Human time perception and its illusions. Current Opinion in Neurobiology, 18, 131-136.

Eagleman, D. M., & Holcombe, A. O. (2002). Causality and the perception of time. *Trends in Cognitive Sciences, 6*(8), 323-325.

Ebert, J. P., & Wegner, D. M. (2010). Time warp: authorship shapes the perceived timing of actions and events. *Consciousness and Cognition, 19*(1), 481-489.

Engbert, K., & Wohlschlaeger, A. (2007). Intentions and expectations in temporal binding. *Consciousness and Cognition, 16*(2), 255-264.

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

Ferrara, A., Lejeune, H., & Wearden, W. (1997). Changing sensitivity to duration in human scalar timing: An experiment, a review and some possible explanations. *Quarterly Journal of Experimental Psychology, 50B*, 217-237.

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

Gebhard, J. W., & Mowbray, G. H. (1959). On discriminating the rate of visual flicker and auditory flutter. *American Journal of Psychology, 72*, 521-529.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279-325.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), Timing and time perception (pp. 52-77). Annals of the New York Academy of Sciences, 423.

Gil, S., & Droit-Volet, S. (2012). Emotional time distortions: The fundamental role of arousal. *Cognition and Emotion, 26*(5), 847-862.

Greville, W. J., & Buehner, M. J. (2015). Temporal predictability enhances judgments of causality in elemental causal induction from both observation and intervention. *The Quarterly Journal of Experimental Psychology*, *69*(4), 678-697.

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin, 127*(1), 22-44.

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception & Psychophysics, 72*(3), 561-582.

Grondin, S. (2014). Why studying intermodal duration discrimination matters. *Frontiers in Psychology: Perception Science, 5*, 628.

Grondin, S., Roussel, M. E., Gamache, P. L., Roy, M., & Oullet, B. (2005). The structure of sensory events and the accuracy of time judgments. *Perception, 34*, 45-58.

Gu, B-M., & Meck, W. H. (2011). New perspectives on Vierordt's law: memory mixing in ordinal temporal comparison tasks. *Lecture Notes in Computer Science, 6789*, 67-78.

Haering, C., & Kiesel, A. (2014). Intentional binding is independent of the validity of the action effect's identity. *Acta Psychologica, 152*, 109-119.

Haering, C., & Kiesel, A. (2015). Was it me when it happened too early? Experience of delayed effects shapes sense of agency. *Cognition, 136,* 38-42.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382-385.

Hayes, B. K., & Rehder, B. (2012). The development of causal categorization. *Cognitive Science, 36*, 1102-1128.

Hume, D. (1739/1888). A treatise of human nature. In L. A. Selby-Bigge (Ed.), *Hume's treatise of human nature*. Oxford, UK: Clarendon Press.

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1542-1549.

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental Brain Research, 203*(2), 465-470.

Johnson, C. A., Chauhan, B. C., & Shapiro, L. R. (1992). Properties of staircase procedures for estimating thresholds in automated perimetry. *Investigative Ophthalmology and Visual Science, 33*(10), 2966-2974.

Keetels, M., & Vroomen, J. (2008). Temporal recalibration to tactile-visual asynchronous stimuli. *Neuroscience letters, 430*(2), 130-134.

Keetels, M., & Vroomen, J. (2012). Perception of synchrony between the senses. In M. T. Murray, & M. M. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 147-178). Boca Raton, Florida: CRC Press.

Kesten, H. (1958). Accelerated stochastic approximation. *Annals of Mathematical Statistics, 29*, 41-59.

Killeen, P. R. (1992). Counting the minutes. In F. Macar, V. Pouthas, & W. Friedman (Eds.), *Time, action cognition: Towards bridging the gap* (pp. 203-314). Dordrecht, Netherlands: Kluwer.

Killeen, P. R., & Fetterman. J. G. (1988). A behavioural theory of timing. *Psychological Review, 95*, 274-295.

Klapproth, F. (2011). Temporal decision making in simultaneous timing. *Frontiers in Integrative Neuroscience, 5*(71), 1-10.

Knoblauch, K., & Maloney, L. T. (2012). *Modeling Psychophysical Data in R*. New York: Springer.

Leek, M. R. (2001). Adaptive procedures in psychological research. *Perception & Psychophysics, 63*(8), 1279-1292.

Lejeune, H., & Wearden, J. H. (2006). Scalar properties in animal timing: conformity and violations. *The Quarterly Journal of Experimental Psychology, 59*(11), 1875-1908.

Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69*(6), 1918-1929.

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.

Matthews, W. J. (2011). Can we use verbal estimation to dissect the internal clock? Differentiating the effects of pacemaker rate, switch latencies, and judgment processes. *Behavioural Processes, 86*, 68-74.

Matthews, W. J. (2013). How does sequence structure affect the judgment of time? Exploring a weighted sum of segments model. *Cognitive Psychology, 66*, 259-282.

Matthews, W. J., & Meck, W. H. (2014). Time perception: the bad news and the good. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(4), 429-446.

Matthews, W.J., Stewart, N., & Wearden, J.H. (2011). Stimulus intensity and the perception of duration. *Journal of Experimental Psychology: Human Perception and Performance, 37(1),* 303-313.

McCormack, T., Brown, G. D., Maylor, E. A., Darby, R. J., & Green, D. (1999). Developmental changes in time estimation: comparing childhood and old age. *Developmental Psychology, 35*(4), 1143-1155.

Mella, N., Conty, L., & Pouthas, V. (2011). The role of physiological arousal in time perception: Psychophysiological evidence from an emotion regulation paradigm. *Brain and Cognition, 75*(2), 182-187.

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition, 17*(1), 136-144.

Moore, J., Lagnado, D., Deal, D. C., & Haggard, P. (2009). Feelings of control: contingency determines experience of action. *Cognition, 110*, 279-283.

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency. *Consciousness and Cognition, 21*(1), 546-561.

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition, 18*(4), 1056-1064.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research, 17*(1), 154-163.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*, 61-64.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor (Version 0.9.11-3)[Computer software].

Mutter, S. A., DeCaro, M. S., & Plumlee, L. F. (2009). The role of contingency and contiguity in young and older adults' causal learning. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 64*(3), 315-323.

Nolden, S., Haering, C., & Kiesel, A. (2012). Assessing intentional binding with the method of constant stimuli. *Consciousness and Cognition, 21*(3), 1176-1185.

van Noorden, L. P. A. S. (1975). Temporal coherence in the perception of tone sequences. Unpblished doctoral dissertation. Technische HogeschooI Eindhoven, Eindhoven, The Netherlands.

Noulhiane, M., Mella, N., Samson, S., Ragot, R., & Pouthas, V. (2007). How emotional auditory stimuli moderate time perception. *Emotion*, 7(4), 697-704.

Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. PLoS ONE, 11, e1264.

Parsons B., Novich, S. D., & Eagleman, D. M. (2013). Motor-sensory recalibration modulates perceived simultaneity of cross-modal events. *Frontiers in Psychology*. 4:46.

Penton-Voak, I. S., Edwards, H., Percival, A., & Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(3), 307-320.

Peterson, D. J., Gozenman, F., Arciniega, H., & Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Attention, Perception & Psychophysics. 77*(7), 2270-2283.

Pierce, J. (2007). Psychopy – Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Poppel, E. (1988). Mindworks: *Time and Conscious Experience*. New York: Harcourt Brace Jovanovich.

Rajendran, V. G., Harper, N. S., Willmore, B. D., Hartmann, W. M., & Schnupp, J. W. H. Temporal predictability as a grouping cue in the perception of auditory streams. *Journal of the Acoustical Society of America, 134*(1), EL98-EL104.

Rammsayer, T. H. (1999). Neuropharmacological evidence for different timing mechanisms in humans. *The Quarterly Journal of Experimental Psychology, 52B*, 273-286.

Rammsayer, T. H., & Ulrich, R. (2001). Counting models of temporal discrimination. *Psychonomic Bulletin & Review, 8*(2), 270-277.

Repp, B. (2005). Sensoriomotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992.

Repp, B. (2011). Tapping in synchrony with a perturbed metronome: the phase correction response to small and large phase shifts as a function of tempo. *Journal of Motor Behavior, 43*(3), 213-227.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989).Temporal contiguity and the judgment of causality by human subjects. *Quarterly Journal of Experimental Psychology Section B- Comparative and Physiological Psychology, 41*(2), 139-159.

Sloman, S., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29*, 5-39.

Sloman, S., Love, B. C., & Ahn, W. (1998). Feature centrality and conceptual coherence. *Cognitive Science, 22*, 189-228.

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology, 13*, R519-R521.

Spielmann, M., Schroger, E., Kotz, S. A., Pechmann, T., & Bendixen, A. (2013). Using a staircase procedure for the objective measurement of auditory stream integration and segregation thresholds. *Frontiers in Psychology, 4*, 534.

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*(5), 651-659.

Stetson, C., Fiesta, M. P., & Eagleman, D. M. (2007). Does time really slow down during a frightening event? PLoS ONE, 2, e1295.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: implicit estimation of sound-arrival time. *Nature, 421*, p. 911.

Tamm, M., Jakobsen, A., Havik, M., Burk, A., Timpmann, S., Allik, J., Oopik, V., & Kreegipuu, K. (2014). The compression of perceived time in a hot and cold environment depends on physiological and psychological factors. *The Quarterly Journal of Experimental Psychology,* 67(1), 197-208.

Tomassini, A., Gori, M., Baud-Bovy, G., Sandini, G., & Morrone, M. C. (2014). Motor commands induce time compression for tactile stimuli. *The Journal of Neuroscience, 34*(27), 9164-9172.

Treisman, M. (1963). Temporal discrimination and the indifference interval: implications for a model of the internal clock. *Psychological Monographs, 77*, Whole number 576.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research, 35*(17), 2503-2522.

Tsakiris, M., & Haggard, P. (2003). Awareness of somatic events associated with a voluntary action. *Experimental Brain Research, 149*, 439–446.

Tse, P. U., Rivest, J., Intriligator, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics, 6*6(7), 1171-1189.

Von Ehrenfels, C. (1890/1988). On "Gestalt Qualities". In B. Smith (Ed. & Trans.), Foundations of Gestalt Theory (pp. 82-117). Wien: Philosophia Verlag.

Wearden, J. H. (1991). Human performance on an analogue of an interval bisection task. *The Quarterly Journal of Experimental Psychology, 43B*, 59-81.

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, 18, 134-144.

Wearden, J. H. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and Mind II: Information Processing Perspectives* (pp. 21-39). Gottingen, Germany: Hogrefe & Huber.

Wearden, J. H. (2004). Decision processes in models of timing. *Acta Neurobiologiae Experimentalis, 64*, 303-317.

Wearden, J. H. (2008). Slowing down an internal clock: implications for accounts of performance on four timing tasks. *The Quarterly Journal of Experimental Psychology, 61,* 264-275.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology*, *51B*, 97-120.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: conformity and violations. *The Quarterly Journal of Experimental Psychology. 61*(4), 569-587.

Wearden, J. H., & McShane, B. (1988). Interval production as an analogue of the peak procedure: evidence for similarity of human and animal timing. *The Quarterly Journal of Experimental Psychology, 40B*, 363-375.

Wearden, J. H., Norton, R., Martin, S., Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 716-729.

Wearden, J. H., O'Rourke, S. C., Matchwick, C., Min, Z., & Maeers, S. (2010). Task switching and subjective duration. *The Quarterly Journal of Experimental Psychology, 63*(3), 531-543.

Wenke, D., & Haggard, P. (2009). How voluntary actions modulate time perception. *Experimental Brain Research, 196*(3), 311-318.

Welch, R. B., Duttonhurt, L. D., & Warren, D. H. (1986). Contributions of audition to temporal rate perception. *Perception & Psychophysics, 39*(4), 294-300.

Wohlschlaeger, A., Haggard, P., Gesierich, B., & Prinz, W. (2003). The perceived onset time of self- and other-generated actions. *Psychological Science, 14*(6), 586-591.

Wolpert, D. M., & Ghamramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience, 3*, 1212-1217.

Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control, *Neural Networks, 9*(8), 1265-1279.

Zakay, D., & Block, R. A. (2004). Prospective and retrospective duration judgments: An executive-control process. *Acta Neurobiologiae Experimentalis*, 64, 319-328.

Zhao, K., Chen ,Y-H., Yan, W-J., & Fu, X. (2013). To bind or not to bind? Different temporal binding effects from voluntary pressing and releasing actions. *PLoS ONE, 8*(5), e64819.

Zhou, B., Yang, S., Zhang, T., Zhang, X., & Mao, L. (2015). Situational context is important: perceptual grouping modulates temporal perception. *Cognitive Processes, 16*, pp. S443-S447.

7. Appendix A: Supplementary analysis 1: Experiments 1 – 4

7.1.1. Results

7.1.1.1. Data analysis

In the main results section I classified intervals as *before* or *during* depending upon whether they began *before* or *during* the interval, respectively. In some cases this means that embedded events that began *before* the interval could have overlapped into the interval. Similarly, embedded events that began during the interval could have overlapped into the outcome stimulus. Thus, an embedded event that began before the interval might finish during the interval, yet be still be classified as a *before* trial. The concern with this approach to classification is that embedded events might be perceived in two segments instead of a single stimulus (Bryce, Seifried-Dubon & Bratzke, 2015; Matthews, 2013). In other words, the portion of an embedded event presented before the interval would be segment 1, and the portion occurring during the interval segment 2. If causal event-triggering actions affect pacemaker rate, then segment 2 would be temporally perceived differently to segment 1. Therefore, there is a chance that our classification process has somewhat confounded the embedded event judgment results. In order to avoid this confound and provide a more robust picture, I analyzed the data based on more rigorous criteria. Specifically, I eliminated trials where the embedded event overlapped into the interval (for *before* trials) and into the outcome stimulus (for *during* trials). The distribution of *before* and *during* embedded events for causal and noncausal trials are shown in Tables A1 (Experiments 1 and 2) and A2 (Experiments 3 and 4). In other respects, I applied the same exclusionary criteria as in the main results section.

The pattern presented by this analysis does not differ to the main results section. As the results show, the pattern is near identical to the main body of results that did not control for overlapping embedded events.

7.1.1.2. Interval Judgments: Experiment 1

Inspection of Figure 8.1 (panel A) finds causal intervals are judged as shorter than noncausal intervals for all levels of Embedded Event Location. Also, noncausal intervals are judged longest when an embedded event was presented before the interval. Statistical analysis confirm these impressions, with a significant effect of Trial Type, $F(1, 30) = 21.73$, $p < .001$, partial $\eta^2 = .42$, Embedded Event Location, $F(2, 60) = 23.12$, $p < .001$, partial $\eta^2 = .44$, and Trial Type x Embedded Event Location interaction, $F(2, 60) = 4.28$, $p < .05$, partial $\eta^2 = .13$. Simple effects analysis revealed that the main effect of Trial Type was significant at all three levels of Embedded Event Location (all *p*s < .05).

7.1.1.3. Interval Judgments: Experiment 2

Figure 8.1 Panel (B) shows the same pattern as intervals in Experiment 1. Causal intervals are again perceived as longer than noncausal, and noncausal intervals with a *before* embedded event are perceived as longer than other interval conditions. ANOVA confirms this with a significant effect of Trial Type, $F(1, 31) = 15.71$, $p < .001$, partial $\eta^2 = .34$, Embedded Event Location, $F(2, 62) = 6.40$, $p < .01$, partial $\eta^2 = .17$, and Trial Type x Embedded Event Location interaction, $F(2, 62) = 6.92$, $p < .01$, partial $\eta^2 = .18$. Simple effects analysis revealed that the main effect of Trial Type was significant at Embedded Event Locations *before* ($p < .001$), and *during* the interval ($p < .05$), but not at the *no event* level ($p = .10$).

7.1.1.4. Embedded Event Judgments: Experiment 1

Numerically, Figure 8.2 (panel A) shows that causal trial embedded events are judged as longer then noncausal, both when presented before and during the interval. However, these observations are not statistically significant, with no significant effect of Trial Type, $F(1, 30) = 1.57$, $p = .22$, partial $\eta^2 = .05$, Embedded Event Location, $F(1, 30) = 3.08$, $p = .09$, partial $\eta^2 = .09$, or Trial Type x Embedded Event Location interaction, $F(1, 30) = .22$, $p = .64$, partial $\eta^2 = .01$. I also conducted a Bayesian analysis using the same procedure I applied to embedded event judgments in Experiment 1 (see main results section of the manuscript). This yielded a Bayes factor of 8 in favour of the null versus the alternative restricted model.

7.1.1.5. Embedded Event Judgments: Experiment 2

In contrast to Experiment 1, embedded events in causal trials are judged as slightly shorter than noncausal (see Figure 8.1, panel B). However, this difference is only numerical, with ANOVA finding no significant effect of Trial Type, $F(1, 31) = .29$, $p = .59$, partial $\eta^2 = .01$, Embedded Event Location, $F(1, 31) = .96$, $p = .34$, partial $\eta^2 = .03$, or Trial Type x Embedded Event Location interaction, $F(1, 31) = .09$, $p = .76$, partial $\eta^2 < .01$. A Bayesian analysis found a factor 10 in favour of the null versus the alternative restricted model.

7.1.1.6. Interval Judgments: Experiment 3

The results display the same pattern as in experiments 1 - 3. These are also significant with ANOVA finding a significant effect of Trial Type, $F(1, 27) = 47.53$, $p < .001$, partial $\eta^2 = .64$, Embedded Event Location, $F(2, 54) = 12.47$, $p < .001$, partial $\eta^2 = .32$, and Trial Type x Embedded Event Location interaction, $F(2, 54) = 14.34$, $p < .001$, partial $\eta^2 = .35$. Simple effects analysis revealed that the effect of Trial Type was significant at all levels of Embedded Event Location (all *ps* < .001).

7.1.1.7. Interval Judgments: Experiment 4

Figure 8.3 (panel B) shows the same pattern as Experiment 4 interval judgments in the main body of the results. Causal intervals are judged as shorter than noncausal, yet noncausal intervals are not judged longer with an embedded event before the interval. ANOVA supports these observations, with a significant effect of Trial Type, $F(1, 30) = 74.70$, $p < .001$, partial $\eta^2 = .71$, Embedded Event Location, $F(2, 60) = 15.96$, $p < .001$, partial $\eta^2 = .34$, but no Trial Type x Embedded Event Location interaction, $F(2. 60) = 1.24$, $p = .30$, partial $\eta^2 = .04$.

7.1.1.8. Embedded Event Judgments: Experiment 3

Embedded events are judged as shorter in causal trials than noncausal (Figure 8.4, panel A). As in Experiments 1 and 2, these are only numerical differences, with no significant effect of Trial Type, $F(1, 27) = 2.76$, $p = .11$, partial $\eta^2 = .09$, Embedded Event Location, $F(1, 27) = 2.61$, $p = .12$, partial $\eta^2 = .09$, or Trial Type x Embedded Event Location interaction, $F(1, 27) = .03$, $p = .86$, partial $\eta^2 < .01$. I also conducted a Bayesian analysis using the same procedure applied to embedded event judgments in Experiment 1. This yielded a Bayes factor of 1 in favour of the *alterative* versus the null model.

7.1.1.9. Embedded Event Judgments: Experiment 4

Figure 8.4 (panel B) shows the same pattern as Experiment 3 (Figure 8.4, panel A). Causal trial embedded events are shorter than noncausal, but this difference is not significant. ANOVA finds no significant effect of Trial Type, $F(1, 30) = 1.52$, $p = .23$, partial $\eta^2 = .05$, Embedded Event Location, $F(1, 30) = .37$, $p = .55$, partial $\eta^2 = .01$, or Trial Type x Embedded Event Location interaction, $F(1, 30) = .44$, $p = .51$, partial $\eta^2 < .02$. A Bayesian analysis found a factor of 6 in favour of the null versus the alternative restricted model.

7.1.1.10. Pooled Analysis

I combined the data from all experiments and included Embedded Event Modality and Outcome Modality as between-subjects factors. The results revealed significant effects of Embedded Event Modality ($p < .001$), an Outcome Modality x Trial Type interaction ($p < .05$), and an Embedded Event Modality x Embedded Event Location interaction ($p < .05$). Importantly, I found no effect of Trial Type ($p = .21$) or Trial Type x Embedded Event Location interaction ($p = .94$). Lastly, I also compared the fit of the pooled data under the restricted model against the null, to yield a Bayes factor of 82 in favour of the null (see Embedded Event Judgments: Experiment 1 in main results section of manuscript).

Table 7.1

Mean percentage of trials on which an embedded event was experienced before and during the overall interval, for interval and embedded event judgment trials in Experiments 1 and 2 (standard deviation in parenthesis). The table shows data with overlapping embedded events removed.

	Experiment 1				Experiment 2			
	Causal		Noncausal		Causal		Noncausal	
Judgment	BI	DI	BI	DI	BI	DI	BI	DI
Intervals	32.18	32.42	27.18	36.94	33.83	32.66	28.52	35.39
	(16.52)	(11.96)	(7.90)	(9.28)	(14.06)	(11.55)	(9.50)	(8.71)
Embedded	33.63	32.90	29.35	36.69	38.28	31.33	27.27	36.95
Event	(13.66)	(10.63)	(6.92)	(7.84)	(13.45)	(10.96)	(7.50)	(7.64)

Note: BI, DI = Embedded Event occurred before/during the interval. Percentages do not sum to 100 because a mean range of 4.81 - 7.77% of trials contained embedded events that began during or after the outcome, and are thus not included. 2 x 2 (Trial Type [causal, noncausal] x Embedded Event Location [before, during]) ANOVAs in which the dependent variable was the proportion of trials with an embedded event, found no significant effects (*ps*>.05), with one exception: Embedded event judgment trials in Experiment 2 had more events in causal than noncausal trials (means of 34.81 and 32.11%, collapsed across BI and DI trials), $F(1, 31) = 4.39$, $p < .05$. I attribute this to random fluctuation.

Table 7.2

Mean percentage of trials on which an embedded event was experienced before and during the overall interval, for interval and embedded event judgment trials in Experiments 3 and 4 (standard deviation in parenthesis). The table shows data with overlapping embedded events removed.

	Experiment 3				Experiment 4			
	Causal		Noncausal		Causal		Noncausal	
Judgment	BI	DI	BI	DI	BI	DI	BI	DI
Intervals	29.91	35.09	26.70	36.16	36.61	34.11	31.61	34.27
	(13.01)	(11.35)	(7.67)	(6.58)	(21.58)	(14.73)	(9.84)	(8.12)
Embedded	28.93	35.80	26.70	40.36	35.73	33.87	27.18	36.45
Event	(13.13)	(10.07)	(8.82)	(6.86)	(23.28)	(13.02)	(8.08)	(5.69)

Note: BI, DI = Embedded Event occurred before/during the interval. Percentages do not sum to 100 because a mean range of 4.17 - 6.25% of trials contained embedded events that began during or after the outcome, and are thus not included. 2 x 2 (Trial Type [causal, noncausal] x Embedded Event Location [before, during]) ANOVAs in which the dependent variable was the proportion of trials with an embedded event, found no significant effects (*ps*>.05), except the following: First, Experiment 3 had more events embedded *during* than *before* the interval both for interval judgment trials (means of 35.63 and 28.30, collapsed across causal and noncausal trials, respectively), $F(1, 27) = 7.65$, $p < .05$, and embedded event judgment trials (means of 38.08 and 27.81, collapsed across causal and noncausal trials, respectively), $F(1, 27) = 16.02$, $p < .001$. Second, Experiment 4 had more events in causal than noncausal trials both for interval judgment trials (means of 35.36 and 32.94%, collapsed across BI and DI trials), $F(1, 30) = 6.11$, $p < .05$, and embedded event judgment trials (means of 34.80 and 31.82%, collapsed across BI and DI trials), $F(1, 30) = 7.00$, $p < .05$. I attribute this to random fluctuation.

7.2. Figures

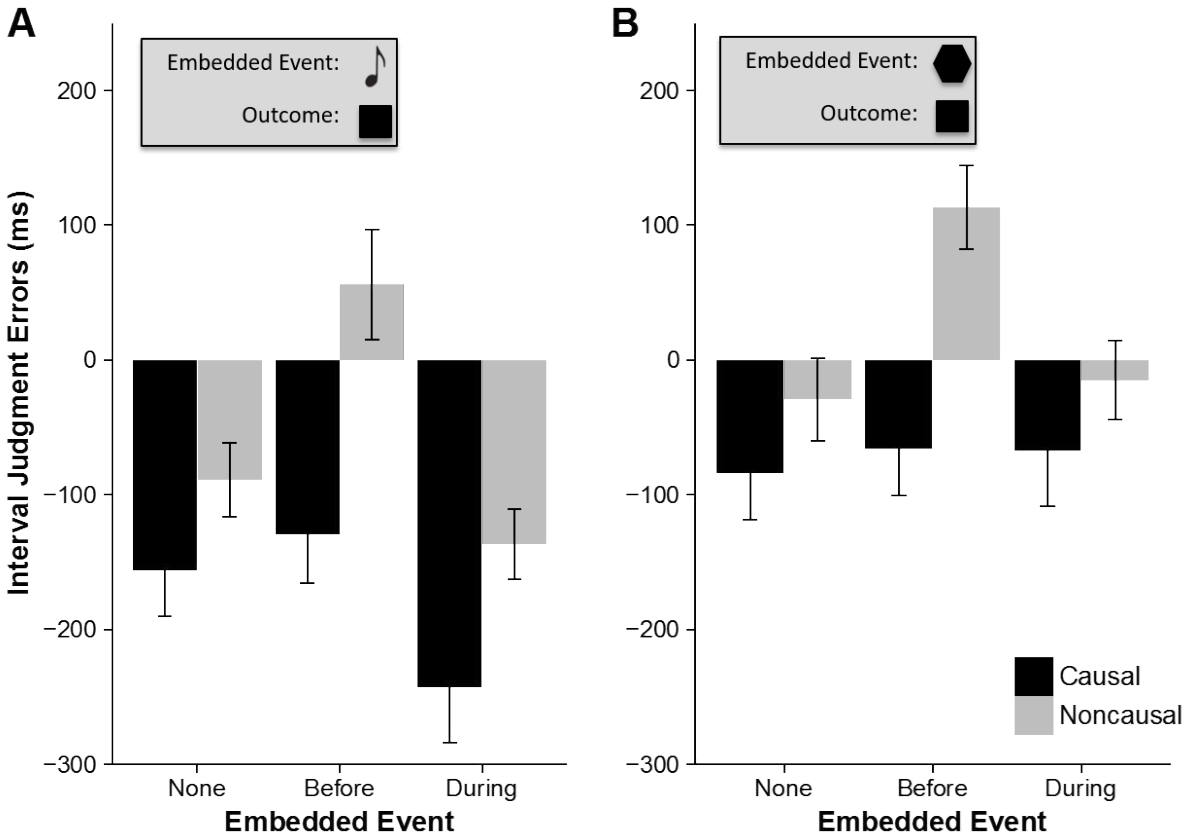

Figure 7.1. (A) Experiment 1. (B) Experiment 2. Data with embedded events removed. Mean interval judgment errors from causal and noncausal trials broken down by embedded event location. Error bars represent within-subjects confidence intervals (Cousineau, 2005; Morey, 2008).

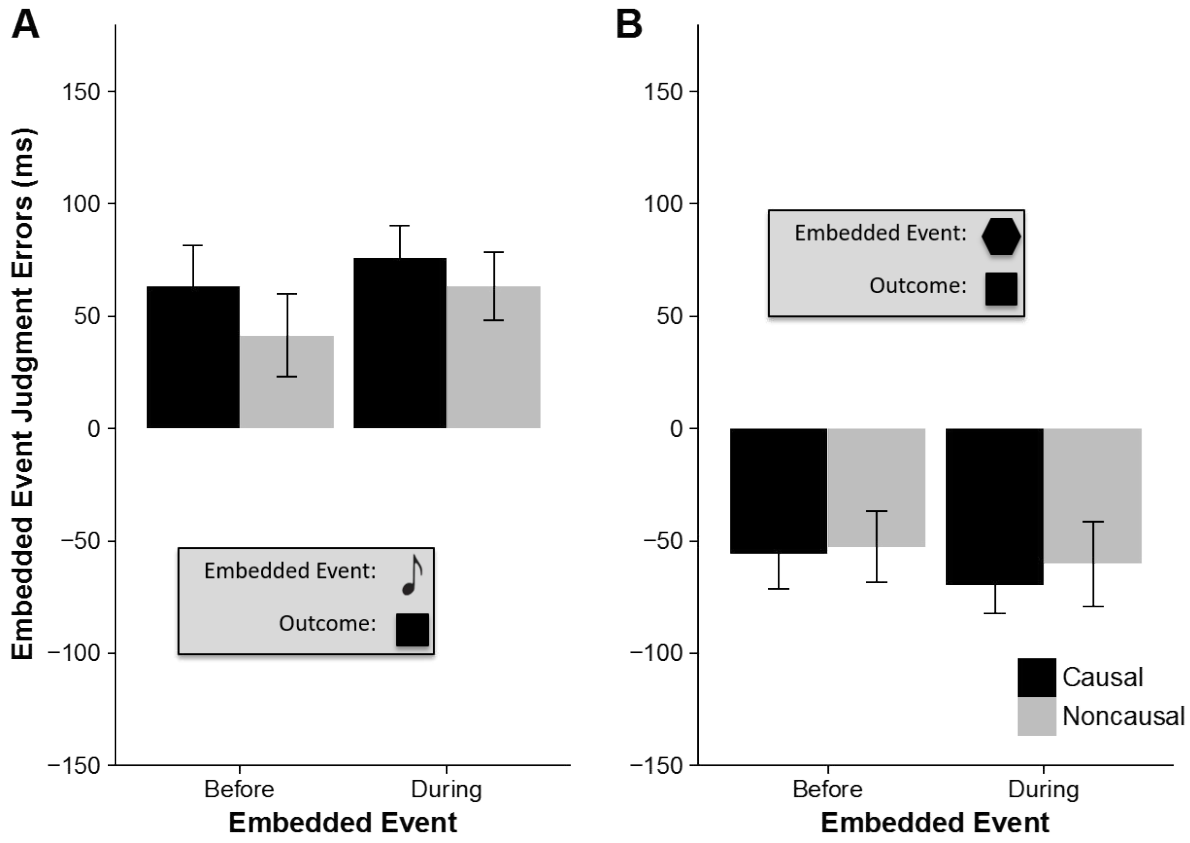

Figure 7.2. (A) Experiment 1. (B) Experiment 2. Data with embedded events removed. Mean embedded event judgment errors for events occurring before and during the interval, broken down by trial type. Error bars represent within-subjects confidence intervals.

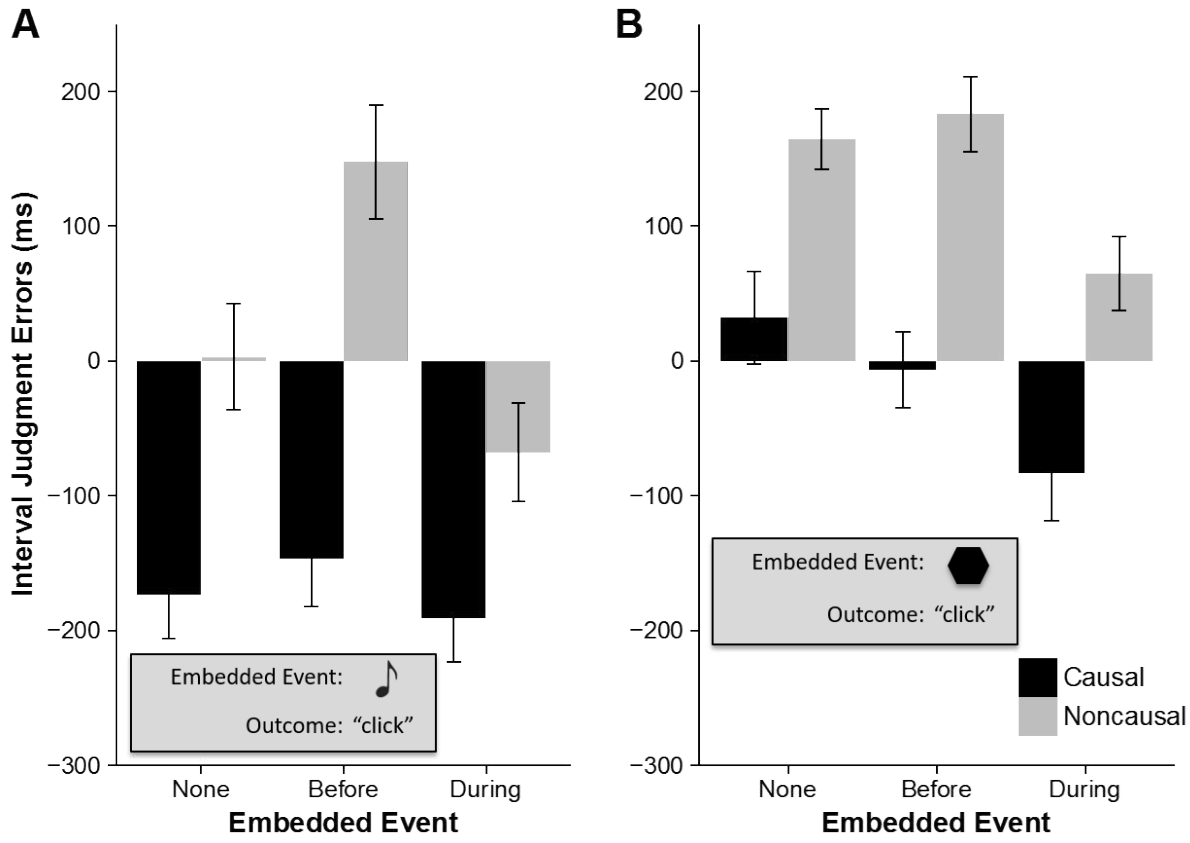

Figure 7.3. (A) Experiment 3. (B) Experiment 4. Data with embedded events removed. Mean interval judgment errors from causal and noncausal trials broken down by embedded event location. Error bars represent within-subjects confidence intervals.

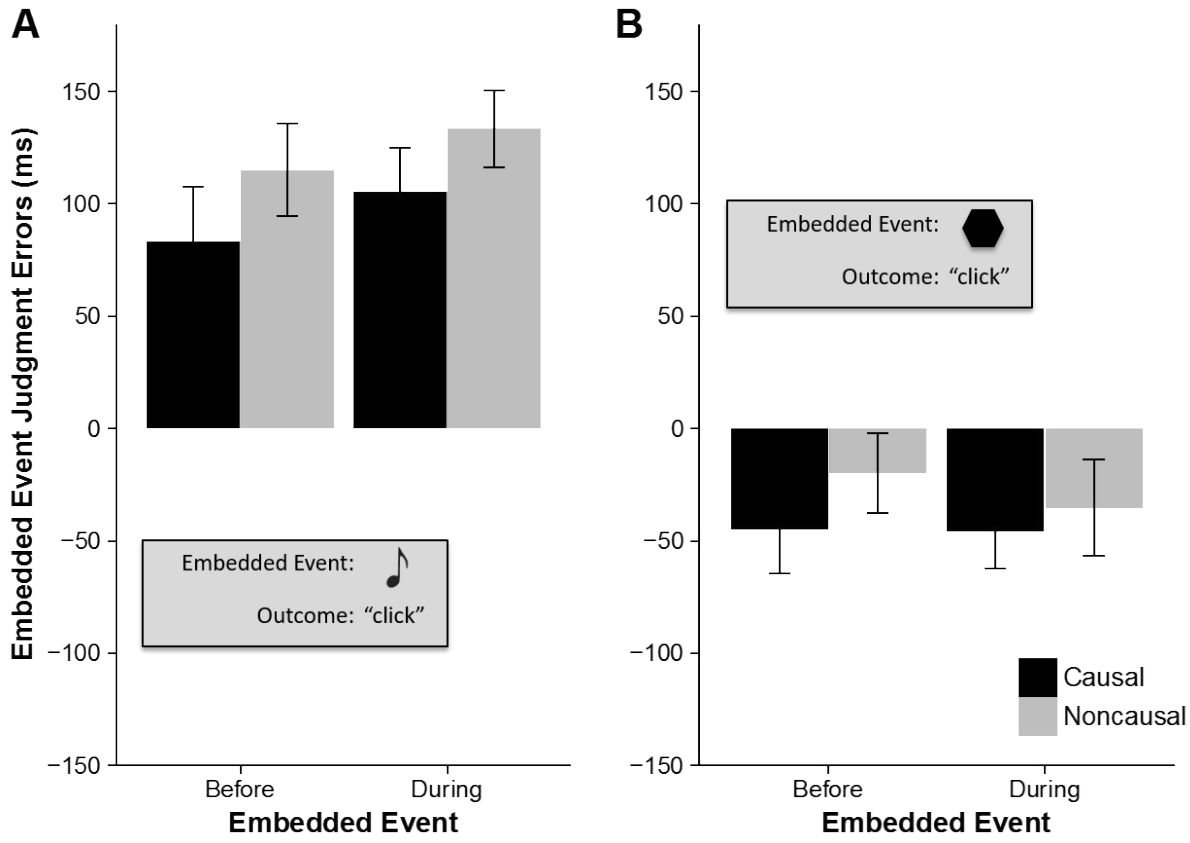

Figure 7.4. (A) Experiment 3. (B) Experiment 4. Data with embedded events removed. Mean embedded event judgment errors for events occurring before and during the interval, broken down by trial type. Error bars represent within-subjects confidence intervals.

8. Appendix B: Scatterplots showing estimated versus actual durations for Experiments 1 – 4

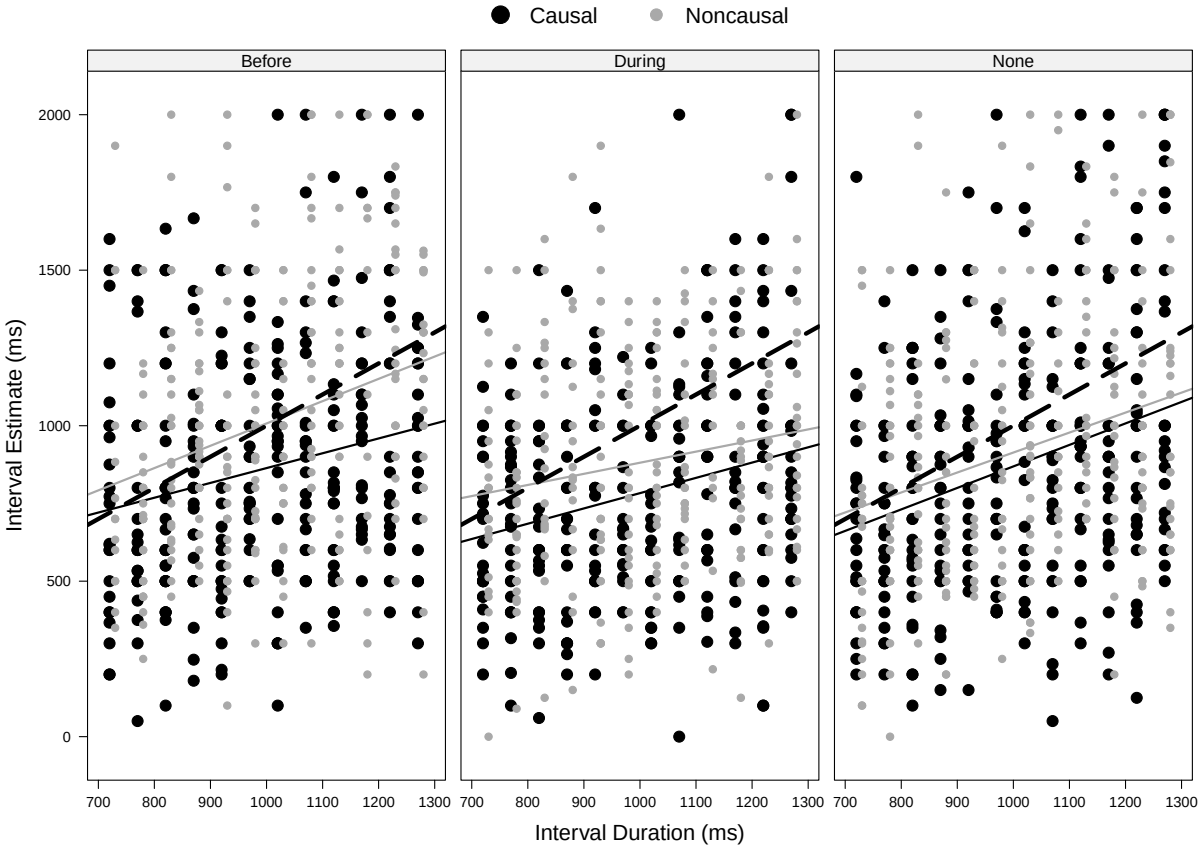

Figure 8.1. Interval judgments for Experiment 1. Scatterplot showing interval estimates plotted against interval durations. Because random durations were used, I binned the estimates into 12 bins within the interval range (700 - 1300 ms). The data points represent the mean of the binned estimates per participant. Regression lines are also shown (emboldened black dashed line).

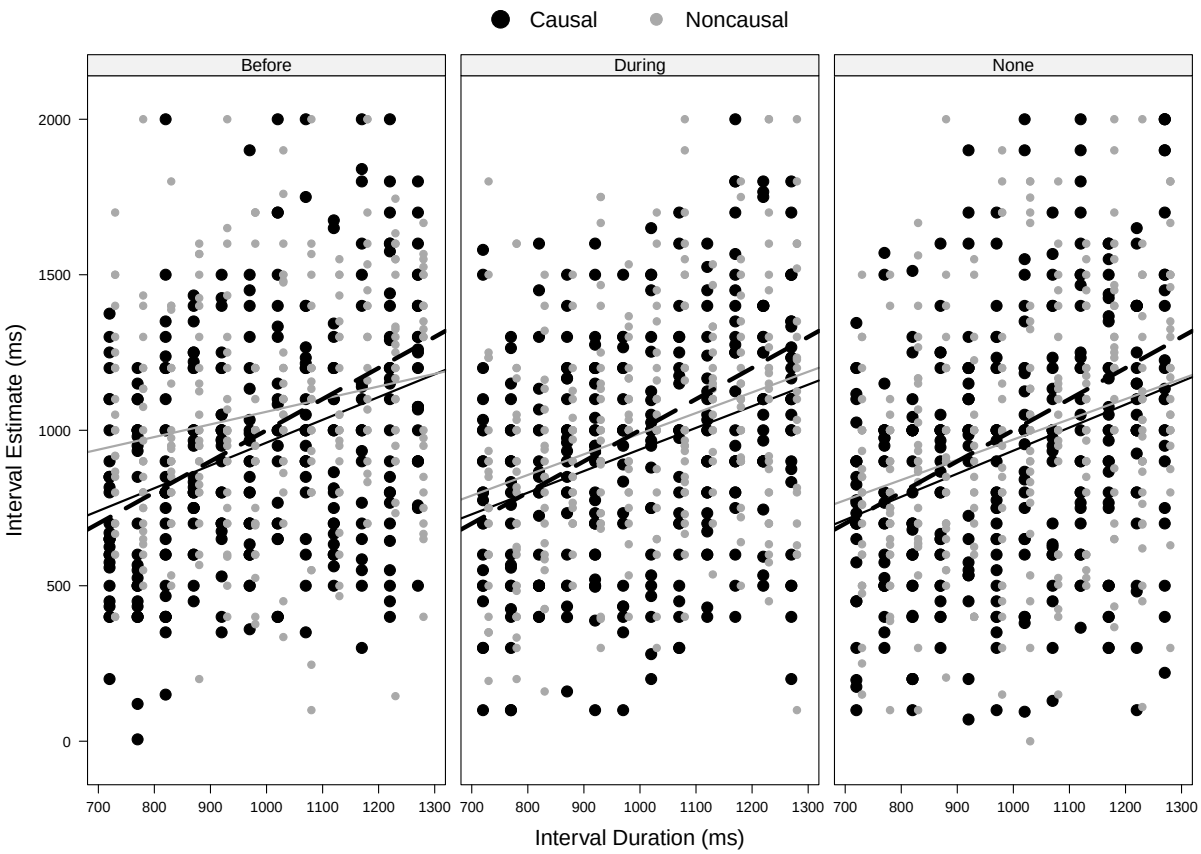

Figure 8.2. Interval judgments for Experiment 2. Scatterplot showing interval estimates plotted against interval durations. The data were binned following the same procedure used for Figure 8.1.

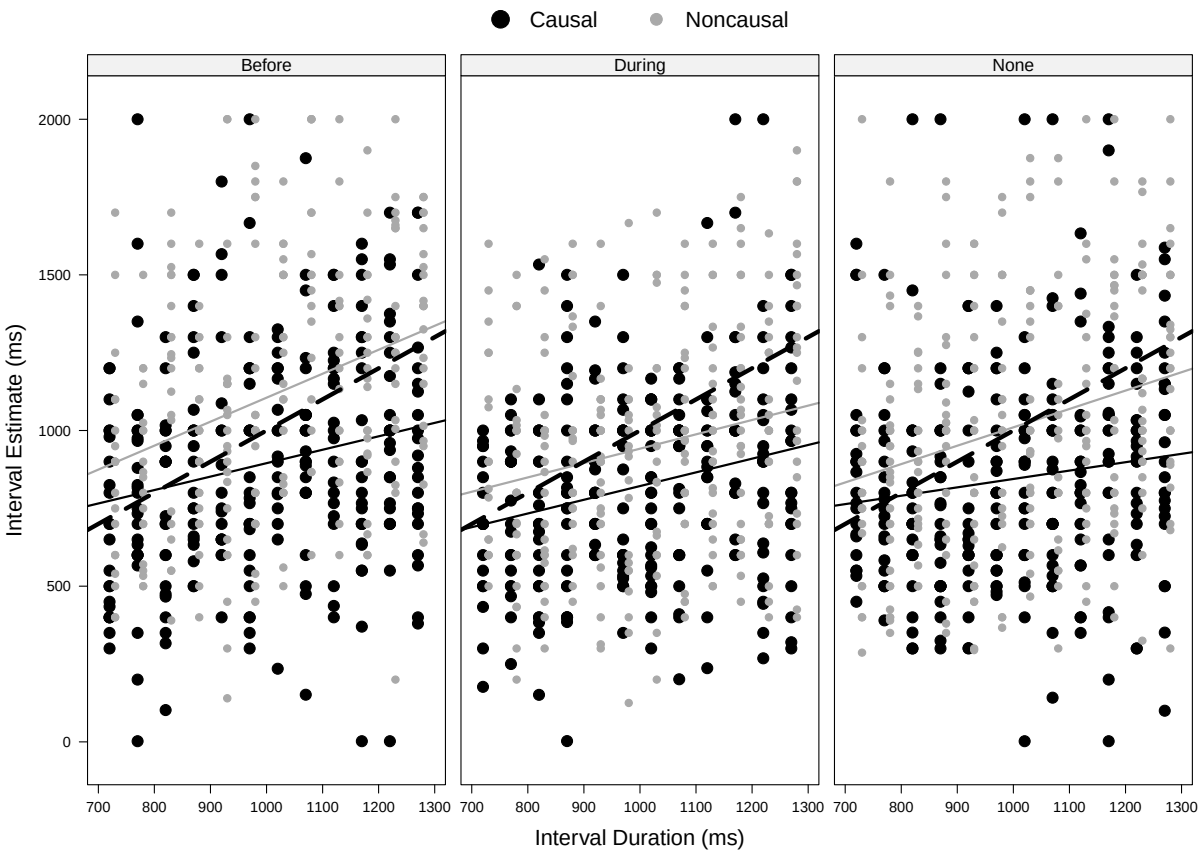

Figure 8.3. Interval judgments for Experiment 3. Scatterplot showing interval estimates plotted against interval durations. The data were binned following the same procedure used for Figure 8.1.

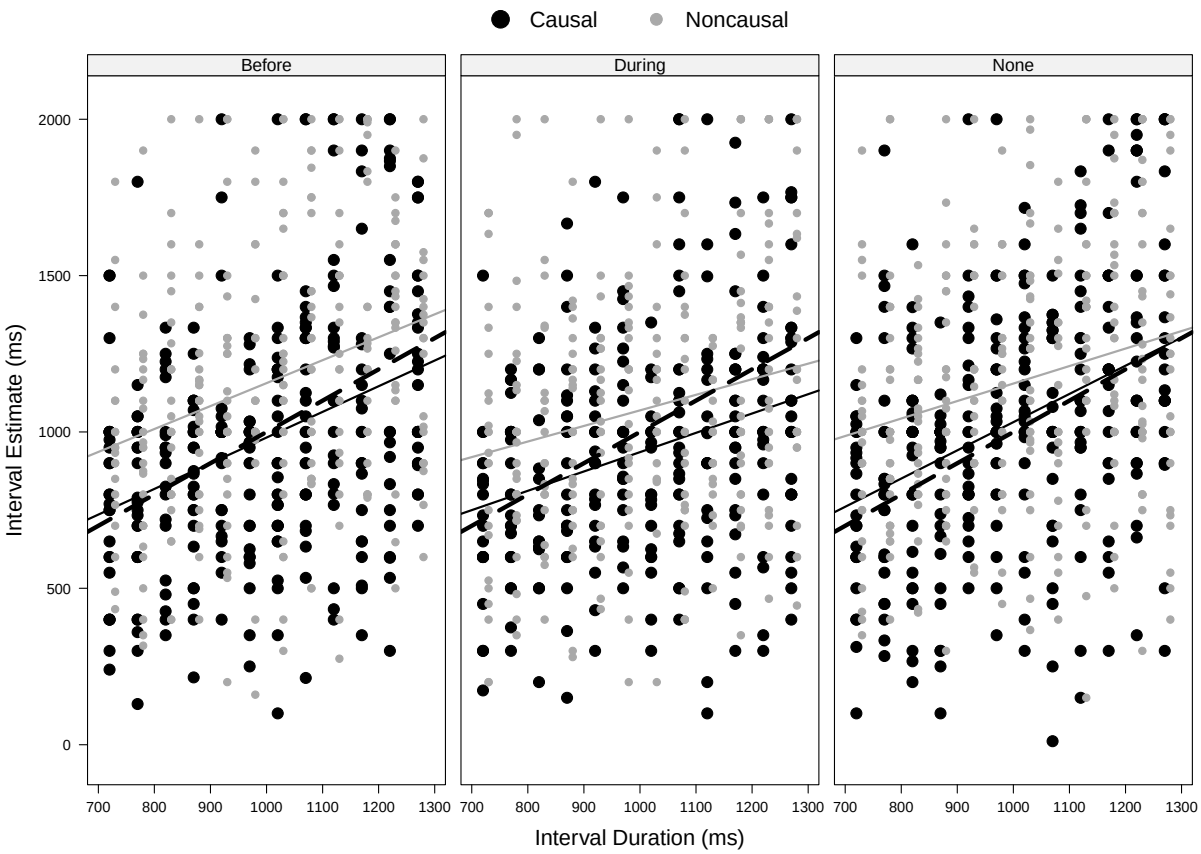

Figure 8.4. Interval judgments for Experiment 4. Scatterplot showing interval estimates plotted against interval durations. The data were binned following the same procedure used for Figure 8.1.

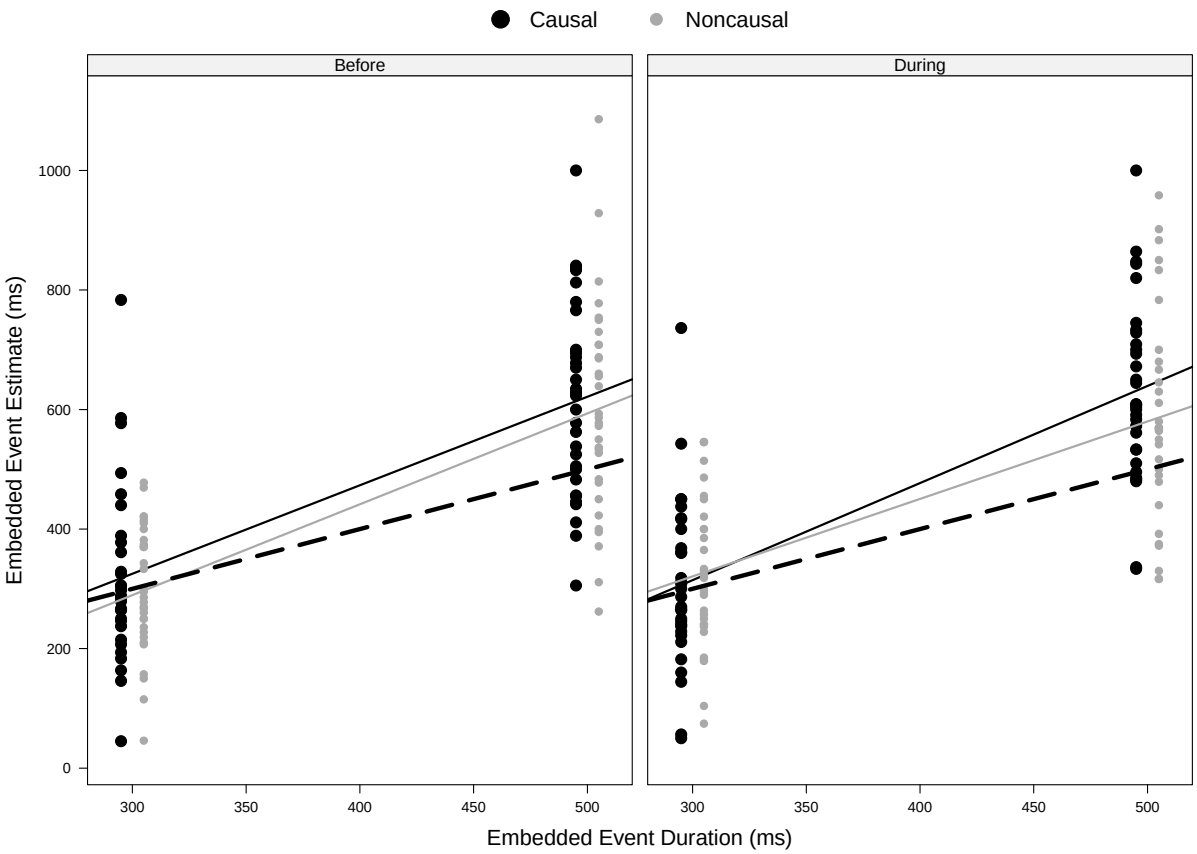

Figure 8.5. Embedded event judgments for Experiment 1. Estimates are plotted against actual durations. Regression lines are shown.

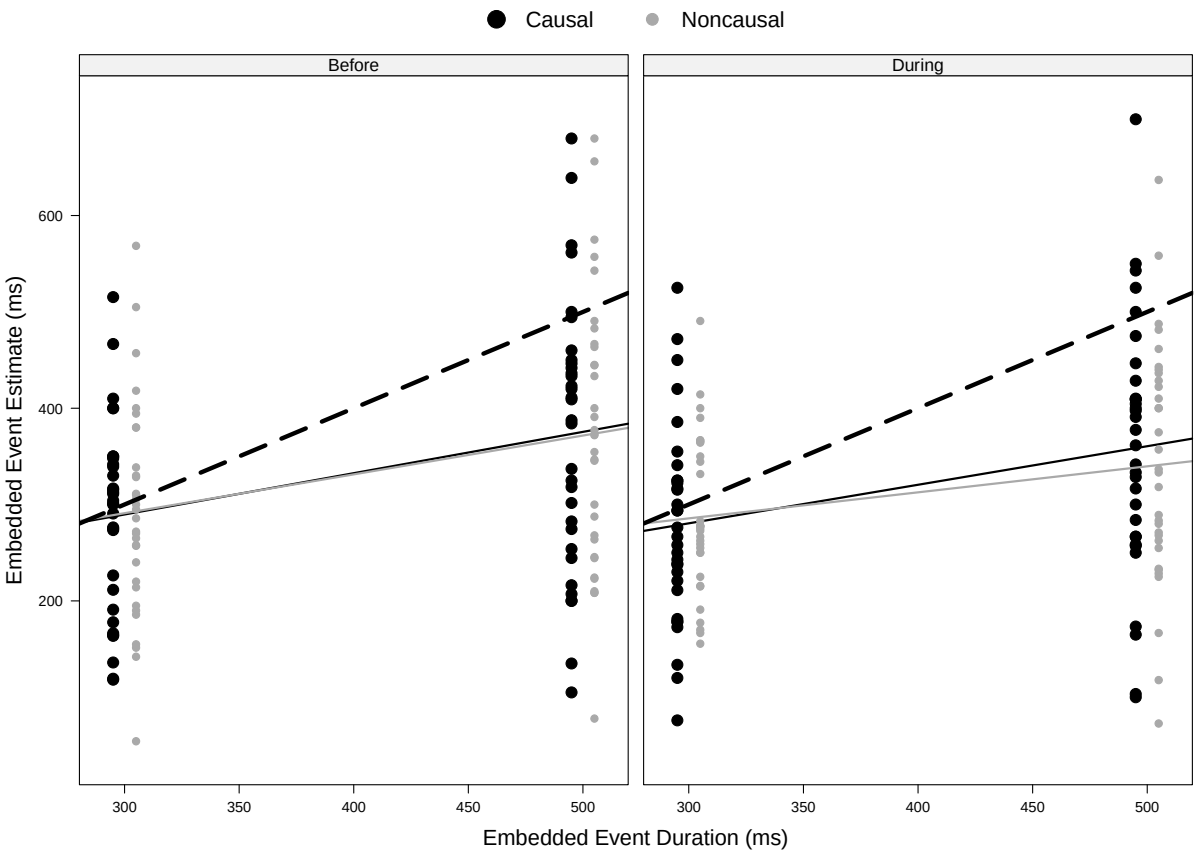

Figure 8.6. Embedded event judgments for Experiment 2. Estimates are plotted against actual durations.

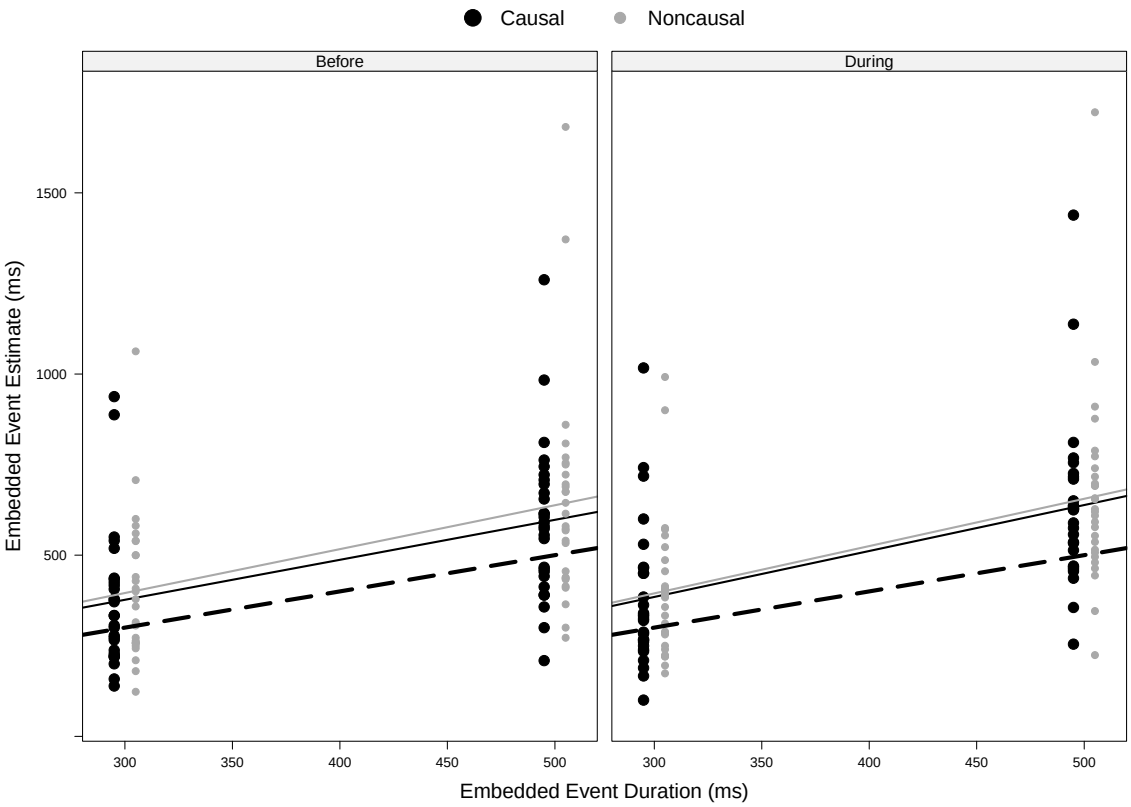

Figure 8.7. Embedded event judgments for Experiment 3. Estimates are plotted against actual durations.

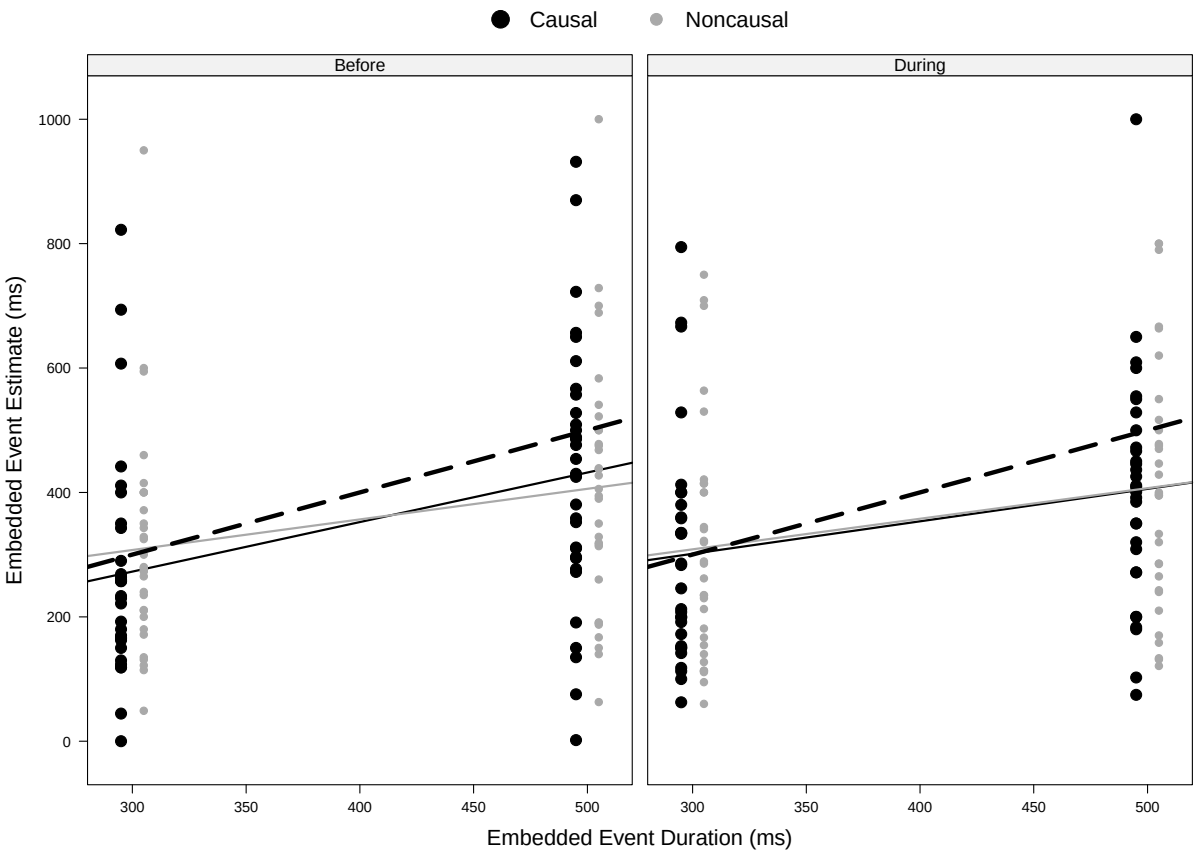

Figure 8.8. Embedded event judgments for Experiment 4. Estimates are plotted against actual durations.

9. Appendix C: Supplementary analysis 2: Experiments 1 – 4

9.1.1. Results

The following analyses are based on deviations in the intended 50-50 balance of *before* and *during* embedded events. Participants who experienced fewer than 35% or greater than 65% of before and during trials, were removed from the analyses.

9.1.1.1. Interval Judgments

Experiment 1. ANOVA found a main effect of Trial Type, $F(1, 23) = 23.47$, $p < .001$, partial $\eta^2 = .51$, Embedded Event Location, $F(2, 46) = 22.83$, $p < .001$, partial $\eta^2 = .50$, and a Trial Type x Embedded Event Location interaction, $F(2, 46) = 4.86$, $p < .05$, partial $\eta^2 = .17$.

Experiment 2. ANOVA found a main effect of Trial Type, $F(1, 19) = 5.06$, $p < .05$, partial $\eta^2 = .21$, but no Embedded Event Location, $F(2, 38) = 1.74$, $p = .19$, partial $\eta^2 = .08$, or a Trial Type x Embedded Event Location interaction, $F(2, 38) = 3.09$, $p = .06$, partial $\eta^2 = .14$.

Experiment 3. ANOVA found a main effect of Trial Type, $F(1, 20) = 39.27$, $p < .001$, partial $\eta^2 = .66$, Embedded Event Location, $F(2, 40) = 10.88$, $p < .001$, partial $\eta^2 = .35$, and a Trial Type x Embedded Event Location interaction, $F(2, 40) = 5.94$, $p < .01$, partial $\eta^2 = .23$.

Experiment 4. ANOVA found a main effect of Trial Type, $F(1, 20) = 49.74$, $p < .001$, partial $\eta^2 = .71$, and Embedded Event Location, $F(2, 40) = 8.64$, $p < .001$, partial $\eta^2 = .30$, but no Trial Type x Embedded Event Location interaction, $F(2, 40) = 2.56$, $p = .09$, partial $\eta^2 = .11$.

9.1.1.2. Embedded Event Judgments

Experiment 1. ANOVA found no effect of Trial Type, $F(1, 23) = .17$, $p = .69$, partial $\eta^2 = .01$, Embedded Event Location, $F(1, 23) = 1.48$, $p = .24$, partial $\eta^2 = .06$, nor a Trial Type x Embedded Event Location interaction, $F(1, 23) = .35$, $p = .56$, partial $\eta^2 = .02$.

Experiment 2. ANOVA found no effect of Trial Type, $F(1, 19) = .19$, $p = .67$, partial $\eta^2 = .01$, Embedded Event Location, $F(1, 19) = 2.67$, $p = .12$, partial $\eta^2 = .12$, nor a Trial Type x Embedded Event Location interaction, $F(1, 19) = .27$, $p = .61$, partial $\eta^2 = .01$.

Experiment 3. ANOVA found no effect of Trial Type, $F(1, 20) = 2.13$, $p = .16$, partial $\eta^2 = .10$, Embedded Event Location, $F(1, 20) = 2.91$, $p = .10$, partial $\eta^2 = .13$, nor a Trial Type x Embedded Event Location interaction, $F(1, 20) = .47$, $p = .50$, partial $\eta^2 = .02$.

Experiment 4. ANOVA found no effect of Trial Type, $F(1, 20) = 1.40$, $p = .25$, partial $\eta^2 = .07$, and Embedded Event Location, $F(1, 20) = .83$, $p = .37$, partial $\eta^2 = .04$, but no Trial Type x Embedded Event Location interaction, $F(1, 20) = 2.41$, $p = .14$, partial $\eta^2 = .11$.