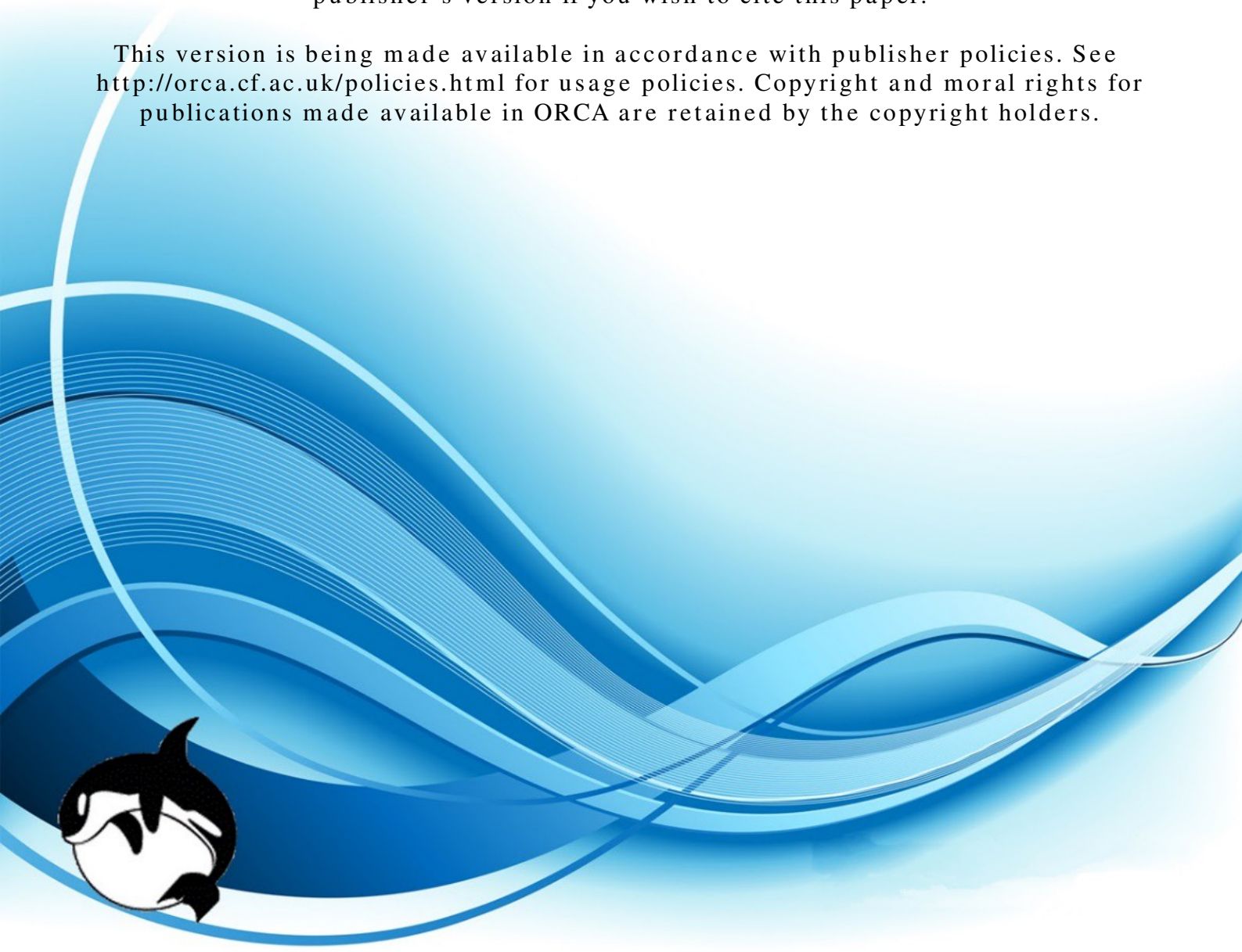

Temporal binding and internal clocks:
No evidence for general pacemaker slowing

Richard Fereday, Marc J. Buehner

Cardiff University

Author Note

Richard Fereday and Marc J Buehner, School of Psychology, Cardiff University, 70 Park Place, Cardiff, CF10 3AT, Wales, UK. The experiments reported in this article were conducted by RF as part of a doctoral thesis under supervision of MJB. Experiments 1 and 2 were presented at the 2015 meeting of the Cognitive Science Society in Pasadena, California, and appeared in Fereday, R., & Buehner, M. J. (2015). Temporal Binding and Internal Clocks: Is Clock Slowing General or Specific? Proceedings of the Thirty-seventh Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

We thank Richard Morey for his assistance with the Bayesian analyses.

Abstract

The perception of time is distorted by many factors (e.g., arousal, temperature, age etc.), but is it possible that causality would affect our perception of time? We investigate timing changes in the temporal binding effect, which refers to a subjective shortening of the interval between actions and their outcomes. Four experiments investigated whether binding may be due to variations in the rate of an internal clock. Specifically, we asked whether binding reflects changes to a general timing system, or a dedicated clock unique to causal sequences. We developed a novel experimental paradigm (embedded interval estimation procedure) in which participants made temporal judgments of either causal or non causal intervals, or the duration of an event embedded within that interval. Stimuli and modality were combined factorially, with interval markers and embedded events being either visual or auditory. While we replicated the temporal binding effect, we found no evidence for commensurate changes to time perception of the embedded event, which suggests that temporal binding is effected by changes to a specific and dedicated, rather than a general clock system.

Keywords: Temporal binding; Causality; Internal clock models; Time perception; Action-outcome learning

Public Impact Statement

Time perception is an intuitive and ubiquitous concept. Adages like ‘the watched pot never boils’ and ‘time flies when you’re having fun’ are ubiquitous in everyday parlance. Both sayings refer to the perceived rate of time: in the former, time appears to move slowly, while the opposite is true in the latter. Research finds that perceived time is affected by emotions, arousal, temperature, and modality of the content of experience (e.g., auditory or visual). An additional concept however, that also affects temporal judgments is causality; causes and effects are perceived as subjectively closer in time than two unrelated events, while duration timing tasks find that cause-effect intervals are judged as shorter than intervals separating unconnected events. Here we show that this causal binding effect is specific to the interval separating the cause from its effect: Causality slows down time perception only for the cause-effect interval; temporal perception for unrelated concurrent events remains unaffected. This finding is problematic for the folk notion of a single ‘internal clock’, but can be explained by theories postulating multiple clocks operating in parallel.

Temporal binding and internal clocks:
No evidence for general pacemaker slowing

Temporal binding refers to the perceptual attraction of actions to their effects (Haggard, Clark, & Kalogeras, 2002). When an action triggers a causal outcome, the action is perceived to occur later, and its outcome earlier, than two unrelated events. One suggestion is that temporal binding is a bi-directional constraint of Bayesian causal inference (Eagleman & Holcombe, 2002): The closer together two events occur in time, the more likely they will be judged as causally related (e.g. Buehner, 2005). Consequently, it follows that if two events are known to be causally related, they are also more likely to be contiguous in time.

Early experiments (Haggard et al., 2002) on temporal binding used the Libet clock method (Libet, Gleason, Wright, & Pearl, 1983), in which participants watch a fast-moving clock hand (one rotation every 2560ms) while experiencing different events. The participant has to report the hand position at the time when she perceives a target event. Judgment errors derived over repeated trials are then used as a proxy measure of event awareness. Using this method, Haggard et al. found a systematic shift in judgment errors for causal actions (key presses), which triggered an outcome (tone) after 250ms. More specifically, participants showed delayed awareness of their causal action, and early awareness of its consequence, relative to single-event judgment errors. In other words, actions and outcomes mutually attracted each other in subjective awareness. This temporal binding effect did not occur when participants reported the times of two unrelated events (Haggard et al., 2002).

Studies using the stimulus anticipation method (SAM) have replicated and expanded upon the temporal binding effect (Buehner & Humphreys, 2009; Buehner, 2012). In the SAM, participants have to press a key in anticipation of a target event. A series of studies using the SAM has repeatedly demonstrated early anticipation of target events that were

triggered by a causal relation, compared to targets that were equally predictable, but were merely associated with the predictor, rather than caused by it. Specifically, Buehner & Humphreys (2009) found that it is not sufficient for an intentional action to be followed by the target – the action has to cause it. Furthermore, even a non-intentional mechanical cause that triggers an outcome results in binding (Buehner, 2012). Thus, the ability to predict the outcome via association is not sufficient to elicit temporal binding. Instead, a more meaningful link (i.e., causality) is both necessary and sufficient to bind events together in subjective time.

However, temporal binding occurs not only in the form of shifts in the perception of events delineating an interval, but also manifests itself via direct distortions of time perception. For example, Humphreys & Buehner (2009) found verbal estimates of intervals separating causal actions from their outcomes to be reliably lower than estimates of intervals separating two unrelated events. These changes also occur when participants reproduce the experienced interval (by holding down a key for the duration of the experienced interval; Humphreys & Buehner, 2010), or when they compare target intervals to standards in a method of constant stimuli (Nolden, Haering, & Kiesel, 2012). Thus, temporal binding manifests itself both via shifts in event perception as well as a direct shortening of experienced time.

Theoretical accounts of temporal binding

Eagleman and Holcombe (2002) suggest that temporal binding arises due to a recalibration of sensory streams: Differences in processing latencies associated with different modalities are overcome by realigning sensory streams, thus ensuring a unitary percept. Because motor acts usually produce causal outcomes immediately, a delay between action and outcome forces a recalibration of the system. A short delay between an action and its

outcome, it is argued, can be adapted to, thus realigning the perceptual-motor system to bring action and outcome closer together in time.

In contrast, temporal binding could also arise due to changes in time perception. Our sense of time is distorted by various factors (for example arousal, Droit-Volet & Meck, 2007; Penton-Voak, Edwards, Percival, & Wearden, 1996). These distortions are typically explained by cognitive models (commonly referred to as 'internal clock' models), such as scalar expectancy theory (SET: Gibbon, Church, & Meck, 1984). Internal clock models contain a pacemaker-accumulator process that represents perceived durations: A pacemaker emits pulses at rate *r*, which are counted in an accumulator; changes to *r* affect temporal judgments, such that decreases and increases in *r* result in respectively fewer and more pulses accumulated in a given interval *I*. Consequently, increases and decreases in *r* will lead to changes in the temporal perception of *I*. Additionally, timing latencies can also influence the temporal perception of *I,* via changes in the precision of an attentional switch. To begin timing *I*, the switch closes and pulses are accumulated; likewise, to stop timing *I*, the switch opens and no further pulses are accumulated. A late closure of the switch will result in a latency L_c between actual and perceived onset of *I*. Similarly, earlier switch opening will result in a latency L_o between actual and perceived offset of *I*. Both L_c and L_o will result in fewer pulses accumulated in interval *I*. Thus, changes in either L_c, L_o (or both), or changes in *r* will the influence temporal perception of *I*.

An example of pacemaker rate change is evidenced in Wearden, Edwards, Fakhri, and Percival (1998). They found that auditory stimuli are perceived as longer then visual stimuli, and, more importantly, that regressions of subjective over objective durations yielded a higher slope for auditory compared to visual stimuli. In other words, the difference between judgments in the two modalities grew as a function of duration, as would be expected by differences in *r*. Might it be possible that changes in the causal nature of event sequences

likewise lead to modulation of pacemaker speed, whereby causal intervals are perceived as shorter due to a slower *r*? To date, there exists only one study that directly investigated this possibility (Wenke & Haggard, 2009).

Empirical evidence for clock slowing in temporal binding

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 slowing of *r* 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.

However, considering Wenke and Haggard's (2009) design reveals that we have to be careful in interpreting their results. It is possible that the higher discrimination thresholds they found in causal trials simply reflect a transient, effector-specific modulation of the pacemaker induced by motor action. For example, Tomassini, Gori, Baud-Bovy, Sandini, and Morrone (2014) investigated time perception surrounding the initiation of voluntary hand movements. Prior to and immediately following movement, participants received two tactile taps, separated by a fixed interval, followed by a second pair of variable-spaced taps toward

the end of each trial. Tactile pads were attached to the moving hand in one condition, and the static hand in another. Participants had to judge the duration of the second pair of variable-spaced taps in relation to the first, which allowed derivation of the subjective point of equality (PSE). Tomassini et al. found shorter PSEs for conditions where the taps were delivered to the moving hand relative to when the taps were delivered to the hand not implicated in the movement. In other words, Tomassini et al. found effector-specific time contraction immediately before and during hand movements. Note that in this study, participants merely moved their hand back and forth, with no noticeable outcome. Thus, although there was no temporal binding, there was evidence of transient, effector-specific temporal compression before and during hand movement. It is therefore likely, that Wenke and Haggard's (2009) result of poorer temporal discrimination following voluntary finger movement is attributable to the same processes, which are independent of temporal binding.

A final issue concerns the theoretical application of the internal clock in Wenke and Haggard's (2009) study. Internal clock models are primarily concerned with interval timing (Wearden, 2003) and are not traditionally applied to simultaneity judgments. Judging whether two events were simultaneous or not, while doubtlessly involving temporal processes, is not the same as judging the length of an interval. At best, the former probes interval perception implicitly, using an event perception measure (i.e., simultaneity judgments) as a proxy, while the latter directly taps internal clock processes. Eagleman (2008) and Pariyadath and Eagleman, (2007) provide tentative evidence that duration and simultaneity judgments indeed are underpinned by separate mechanisms. For example, Pariyadath and Eagleman used the oddball technique to study whether timing is a single entity (i.e., general) or is served by independent processes. Their participants viewed sequences of visual stimuli, during which an unexpected stimulus is embedded (e.g., a clock is presented during a series of shoe images). Results showed that duration judgments were longer for oddballs than for the

surrounding images, when in fact each image was displayed for the same duration. However, this subjective temporal expansion did not generalize to other temporal abilities such as critical flicker fusion. Eagleman (2008) noted that "duration, simultaneity, temporal order, flicker rate, and other judgments are underpinned by different mechanisms that normally concur but are not required to" (p. 134). This means that decreases in subjective duration do not necessarily guarantee decreases in thresholds on a temporal simultaneity task, or vice versa. Thus, there are several reasons why it is not clear whether the differential temporal discrimination reported by Wenke & Haggard's indeed reflects changes to *r*.

Scope

In light of these concerns, we set out to investigate more carefully whether temporal binding might implicate changes to *r*. It is important to note that changes to internal clock processes in temporal binding may manifest in two distinct ways: First, temporal binding might reflect a general slowing of the timing system; secondly binding might only affect clock processes that are specific to the action-outcome interval. According to the first possibility, a slowing of a general all-purpose clock would result in perceptual changes to *any* temporal stimulus presented simultaneously with the interval. According to the latter, effects of clock slowing will be limited to the causal interval. Therefore, a key question we ask is whether clock slowing is a general or specific process.

We developed a new *embedded interval estimation procedure* to address this question. 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 (Haering & Kiesel, 2014, 2015; Humphreys & Buehner, 2009, 2010), and does not merely reflect processes related to motor preparation and execution, we focused our efforts on a simple causal versus non-causal interval distinction: In the experiments reported here, 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. Critically, we embedded an additional event into 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 we 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 pacemaker has no information about causality until the outcome occurs, at which time the switch closes, thereby stopping pulses accumulating. However, Moore and Haggard (2008), using the Libet clock method described earlier (Libet et al., 1983), demonstrated that temporal causal 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 et al. (2009) showed that the extent of

binding varies as a function of the contingency, in line with a causal theory of binding. In our experiments 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 our manipulation of causality.

The four experiments reported in this paper 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. 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 pertains only to the cause-effect interval 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 (see Eagleman, 2008; Klapproth, 2011; we return to this point in the General Discussion). 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 1 for a schematic of model predictions). Because there are modality-specific differences in the timing of auditory versus visual events (Wearden et al., 1998), we decided to factorially

combine stimulus modality with event type in a 2 x 2 design to ensure that whatever effects we may find are not confined to a particular modality.

Experiments 1 and 2

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. To make it clear that the embedded event was independent of the action, we 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, we 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 we employed stimulus delivery times recorded from participants in a pilot experiment using the same algorithm[2]. This procedure ensured noncausal trials mirrored causal trials as closely as possible.

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.

Method

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) in Experiments 2. Students were provided with course credits or £5 payment.

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 [6.21° x 4.83°] 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 [9.19°]) served as the outcome on causal trials and marked the end of noncausal intervals. A black square (400 pixels^2 [9.19°]) was presented at the beginning of each trial in all experiments. All visual stimuli were presented centrally on the screen.

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 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, we 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.

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 we adopted an analogous procedure. We replayed values from causal trials of a pilot version of the experiment, where we 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, we 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 1 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 embedded event 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%.

To optimize the experience on noncausal trials, we screened previous participants' stimulus patterns 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 worked through the experiment in small groups of 10 - 15 students and the experiments took about 45 minutes.

Results

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 1 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 we considered removing these participants from analyses, doing so does not change the pattern of results in any of the four experiments (all interval judgment ps $< .05$, all embedded event judgment ps $> .05$), so we decided to report analyses based on the entire valid sample.

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 (classified as *Embedded Event Location: Before Interval,* and *During Interval*); excluding them does not change the pattern of results (see Online Supplementary Materials). All analyses were computed with respect to judgment errors, 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. Preliminary analyses indicated that there were no main effects of, or interactions associated with Embedded Event Duration. We therefore conducted the analyses by collapsing across Embedded Event Duration.

Interval Judgments: Experiment 1

Figure 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 also shows that intervals with an event presented during the interval were underestimated more than intervals with no event and an event before. We 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 revealed that the main effect of Trial

Type was significant at all three levels of Embedded Event Location (all *p*s < .05). Post hoc tests also revealed that trials with an event during the interval were judged as significantly different than trials with an event before, or trials without an embedded event (both *ps* < .01). Additionally, noncausal trials with an event before significantly differed to all other Trial Type x Embedded Event Location combinations (all *ps* < .001).

Interval Judgments: Experiment 2

Figure 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 (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 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 ($p = .10$). Post hoc tests also revealed that trials with an event during the interval were judged as significantly different than trials with an event before, or trials without an embedded event (both *ps* < .01). Additionally, noncausal trials with an event before significantly differed to all other Trial Type x Embedded Event Location combinations (all *ps* < .01).

Embedded Event Judgments: Experiment 1

Inspection of Figure 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, we 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 design, we compared a restricted model against the null (see Figure 5). The restricted model captured our original prediction, in which the means of events embedded during causal intervals would be subjectively shorter than those embedded during noncausal intervals. We did not specify any other order restrictions (i.e., we 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.

Embedded Event Judgments: Experiment 2

Figure 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 our 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 $\eta^2 < .001$. We also conducted a Bayesian analysis using the same procedure we 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.

Discussion

In both experiments we 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, we also found an interaction, whereby noncausal intervals with embedded event *before* the interval were perceived as longer than intervals with events embedded in other locations. While this is interesting, it is only tangential to our main purpose, and we return to this in the General Discussion. Our main concern was with temporal judgments for events embedded into causal and noncausal intervals. We 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 Supplementary Materials).

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. Our 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.

Experiments 3 and 4

In Experiments 3 and 4 we 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).

Method

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 another 34 in Experiment 4 (28 female, 6 male, M_{age} = 21.9 years, age range: 17-46), in exchange for course credits or £5 payment.

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 [6.21° x 4.83°] 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 [9.19°] pixels^2) was presented at the beginning of each trial in both experiments. The apparatus was the same as in Experiment 1.

Design and procedure

The procedure remained the same as Experiments 1 and 2.

Results

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 2. 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 we 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 (in addition to events that began *during* the interval but overlapped into the outcome), were removed and the data subjected to reanalysis (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 Supplementary Online Material).

Intervals Judgments: Experiment 3

Figure 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 (we explore reasons for this in the General Discussion). In general, Figure 5 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 revealed that the effect of Trial Type was significant at all levels of Embedded Event Location (all $ps < .001$). Post hoc tests also revealed that trials with an event during the interval were judged as significantly different than trials with an event before, or trials without an embedded event (both $ps < .01$). Additionally, noncausal

trials with an event before significantly differed to all other Trial Type x Embedded Event Location combinations (all *ps* < .01).

Interval Judgments: Experiment 4

Figure 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 $\eta^2 = .05$. Post hoc tests also revealed that trials with an event during the interval were not judged as significantly different than trials with an event before ($p = .13$), but were significantly different to trials without a *before* embedded event (both $p < .001$). Additionally, noncausal trials with an event before significantly differed to all other Trial Type x Embedded Event Location combinations (all *ps* < .001).

Embedded Event Judgments: Experiment 3

Figure 7 (A) shows that embedded events are overestimated, corroborating the finding of Experiment 1. Inspection of Figure 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 in favour of the null versus the alternative restricted model (see Embedded Event Judgments: Experiment 1 and Figure 4 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.

Embedded Event Judgments: Experiment 4

Figure 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 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 12 in favour of the null versus the alternative restricted model.

Discussion

The results of Experiments 3 and 4 replicate those of Experiments 1 and 2: We found a significant binding effect, regardless of Embedded Event Location and modality. Furthermore, Experiment 4 found general overestimation of noncausal intervals in all conditions, a pattern not found in the previous Experiments. The important point though, is that we replicated the binding effect in both experiments. Additionally, in Experiment 3 we 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, we will comment on this in the General Discussion. Regarding embedded event judgments, we 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, we 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 Supplementary Materials).

General Discussion

We developed a new procedure - the *embedded interval estimation procedure* - to study the potential implication of internal clock models in temporal binding. Using this method, we 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, we 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, we found evidence of 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 our failure to find an effect of causality on embedded event judgments, might simply be due to the (relatively) short durations we employed for embedded events; differences in *r* might be more likely to manifest at longer intervals. It is certainly true that differences in *r* are more likely to manifest, the longer the duration to be timed. However, Humphreys & Buehner (2009) found that participants gave shorter estimates for causal, compared to noncausal intervals, at durations as short as 150 ms. Thus, causality-induced changes to interval estimation *do* manifest even at very short durations. Because the shortest embedded event duration in the present study was twice this duration (i.e., 300 ms), we would argue the absence of a causal vs non-causal difference for embedded event judgments indeed reflects an absence of a difference in *r*, rather than the inability of such a difference to manifest due to short duration.

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, we 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 found 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, we 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, we 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 Supplementary Online Materials). In general then, this pattern represents positive evidence against a difference between 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 we 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 (we thank an anonymous reviewer for this suggestion).

Importantly however, the 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 our results say about how temporal binding relates to time perception itself? At a minimum, our results show that temporal binding does not occur due to general slowing of the pacemaker. Does this rule out a clock-slowing account of temporal binding altogether? Not necessarily. It could be entirely possible that temporal binding selectively affects a dedicated time-keeping processes, allocated to keeping track of action-outcome (or cause-effect) intervals. 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.

One study that provides evidence for changes in *r* during temporal binding is Humphreys and Buehner (2009). 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. Fereday, Buehner & Rushton (in preparation) replicated this result using a temporal discrimination paradigm. Importantly, they also found impaired temporal resolution during causal sequences: In a separate series of experiment, participants discriminated between either two key press – flash intervals, or two flash – flash intervals (in causal and non-causal conditions, respectively). Across a range of durations (250 – 850 ms), the discrimination thresholds were larger for causal than non-causal intervals, meaning that temporal acuity was reduced in causal conditions, as would be predicted by reduced *r*.

How do our findings relate to Wenke and Haggard (2009)? As noted in the Introduction, it is likely that Wenke and Haggard's results do not reflect clock slowing but instead simply reflect motor-specific contraction of discrimination thresholds at the effector, which is independent of binding (Tomassini et al., 2014). Further, internal clock models are concerned with interval perception, and it is difficult to see how they may be applied to simultaneity judgments.

Is binding effected by shifts in event perception?

Until now we have discussed our 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, our 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.

A different perspective to temporal binding, one that does not explicitly draw on processes of time perception, is that binding reflects subjective shifts in event perception (Eagleman & Holcombe, 2002): The consequences of causal actions are perceived earlier, relative to other events occurring in the environment. Importantly, this perspective need not be mutually exclusive from an approach inspired by internal clock models: A slower *r* for action-outcome timing, for example, can lead to a subjective contraction of time (i.e., binding), which in turn would result in the outcome perceived sooner. Importantly, the outcome would only be perceived sooner (relative to other events in the environment), if the slowing of *r* is specific to the cause-effect interval. In this case, the action-outcome clock rate would differ to that of other timing streams, meaning that the outcome would be perceived earlier relative to other events timed at a standard clock rate. A general slowing of *r*, in contrast, would not lead to an earlier perception of the outcome relative to other events, because all events would now be judged in relation to the *same* slowed clock speed, and thus would shift by the same amount.

Alternatively, if temporal binding were effected via changes in switch latency L_c and L_o (as opposed to, or in addition to causality-specific changes in *r*), such changes in latency could correspond to a perception that the event marking the start of the interval was perceived subjectively *later*, while the event marking the end of the interval was perceived subjectively *earlier*. Consequently, the idea that temporal binding is effected via changes to clock processes that are specific to the causal interval can account not only for distortions of interval perception as reported in this paper, but also for the perceived shifts in event perception reported elsewhere in the literature (e.g. Buehner, 2012; Buehner & Humphreys, 2010; Haggard et al., 2002). A key task for future research is to determine whether such causality-induced changes to clock processes reflect differences in clock rate or timing latencies (or both).

A cue combination account?

Temporal binding has recently been interpreted in light of a cue-combination perspective (Moore & Fletcher, 2012; Wolpe, Haggard, Siebner, & Rowe, 2013; Kawabe, Roseboom, & Nishida, 2013). This traces back to seminal work conducted by Ernst and Banks (2012) in the domain of multi-sensory integration. Ernst and Banks found that judgments of stimulus height were based on an integration of visual and haptic cues. Importantly, they found that cue integration operates in a statistically optimal manner, such that cues associated with less variability (i.e. those that are more reliable) are weighted more heavily when forming an integrated judgment: systematic degradation of one signal (optic vs haptic) led to increased weighting of the other signal. In the context of temporal binding, the different cues that could be available are predictions of an internal forward model of the sensory consequences of one's action, proprioception of the action (or perception of the start of the interval), as well the feedback of the sensory consequence / end of interval itself (arguably, higher level beliefs about causality and agency might also be considered as cues here); these cues can then be combined to yield a posterior estimate of either the action-outcome interval, the sense of agency, or the time of action and the time of outcome. Interestingly, and in parallel with Ernst and Banks, Wolpe et al., using an event-perception paradigm, found reduced temporal binding of actions to their effects in situations where the outcome signal was degraded. It is not clear, however, how such accounts apply in situations involving prospective timing. Presumably, one could argue that the perceived time of the start and end of the to-be-timed interval is subject to biases, while time perception itself still operates according to SET. In the context of our experiments, the salient difference between causal and non-causal conditions was that the former employed a key press, whereas the latter a click-sound as the start of the to-be-timed interval; the end point was always the same. One could thus argue that the internal representation of the time of the key press would have

been more precise (and therefore more reliable) than the representation of the click sound. Applying the above logic then suggests that the perceived time of the end of the interval would be biased more strongly towards a key press (start of causal trials) than towards a click sound (start of non-causal trials). Conversely, the perceived time of the click (marking the start of a non-causal interval) would be subject to greater bias from the event marking the end of the interval than the perceived time of an action (marking the start of a causal interval). With respect to interval estimation – the variable of interest in the work presented here – this means that there are two tendencies operating in opposite directions. Yet we observed a clear and unambiguous temporal binding effect, such that causal intervals were always judged as shorter than non-causal intervals, regardless of stimulus modality, and, importantly, regardless of whether or not there were embedded events present, and regardless of whether the embedded events occurred before or during the interval. A cue combination account therefore cannot explain the overall binding effect purely on the basis of differential signal precision. However, if one adds prior beliefs (in this cause about causality) to the account, it can (Moore & Fletcher, 2012). Importantly, our work has shown that temporal binding is not driven by bottom-up biases to local (intervening, embedded) events, but that it is rooted in a persistent linkage between cause and effect. At the same time, the effect is highly specific, and limited to the cause-effect interval, and does not extend to other temporal events, even when they occur concurrently with the causal interval. A cue-combination interpretation of our results therefore would be that causality biases the judgments of elapsed time between cause and effect, and that this bias is highly specific to the causal interval, such that other, concurrently occurring events are not subject to the same bias.

In sum, our 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 these accounts is by systematically investigating clock processes in causal and noncausal intervals with regression models. Other work from our lab (Fereday, Rushton & Buehner, in preparation) found conclusive evidence for causality-specific changes to clock rate. For now, however, we conclude that the perceptual shifts associated with temporal causal binding occur due to changes in time-keeping processes specific to the causal interval. Whether these changes reflect a causality-specific slowing of clock-rate, changes to switch-latencies, or systematic biases induced via cue-combination is a matter to be explored elsewhere.

References

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

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. (2005). Contiguity and covariation in human causal inference. *Learning and Behavior, 33*(2), 230-238.

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 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., & Humphreys, G. R. (2010). Causal contraction: spatial binding in the perception of collision events. *Psychological Science, 21*(1), 44-48.

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

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.

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.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513.

Eagleman, D. (2008). Human time perception and its illusions. *Current Opinon 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.

Fereday, R., Bueher, M., & Rushton, S. (2016). Causality impairs temporal resolution: accounting for temporal binding with the internal clock model. *Manuscript in preparation*.

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

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.

Hume, D. (1888). *Hume's treatise of human nature* (L.A. Selby-Bigge, Ed.). Oxford, England: Clarendon Press (Original work published 1739).

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.

Kawabe, T., Roseboom, W., & Nishida, S. (2013). The sense of agency is action-effect causality perception based on cross-modal grouping. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1763), 20130991–20130991.

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

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

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59–68.

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.

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].

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

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

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*, 307-320.

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

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.

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: Hogrefe & Huber.

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. *Quarterly Journal of Experimental Psychology*, *51B*, 97-120.

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

Wolpe, N., Haggard, P., Siebner, H. R., & Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research*, *229*(3), 467–474.

Footnotes

1 Naturally, the second shock might also be bound to the action. No previous research has investigated whether action-effect binding extends to multiple events caused by the same action. However, previous research (e.g. Haggard et al., 2002) showed that the extent of binding decreases with increasing action-effect intervals. Consequently, one could expect that the second shock would be bound to the action to a lesser extent than the first shock, and this would lead to an overall lengthening of the subjective inter-shock interval – which would give rise to an apparently higher simultaneity detection threshold.

2 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 durations (interval duration, time of key press, time of embedded event etc) replayed in noncausal blocks. This pilot experiment used the same prediction algorithm as the current experiments.

Tables

Table 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).

	Experiment 1				Experiment 2			
	Causal		Noncausal		Causal		Noncausal	
Judgment	BI	DI	BI	DI	BI	DI	BI	DI
Intervals	50.16 (11.09)	44.35 (12.63)	41.94 (6.28)	53.63 (6.35)	47.66 (9.71)	45.55 (11.55)	44.38 (7.04)	51.72 (6.64)
Embedded Event	50.40 (10.02)	45.16 (11.76)	46.05 (5.62)	49.27 (6.23)	51.41 (10.34)	41.95 (12.28)	46.48 (5.53)	48.67 (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), $F(1, 31) = 7.79$, $p < .01$. We attribute this to random fluctuation.

Table 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).

	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).

Figures

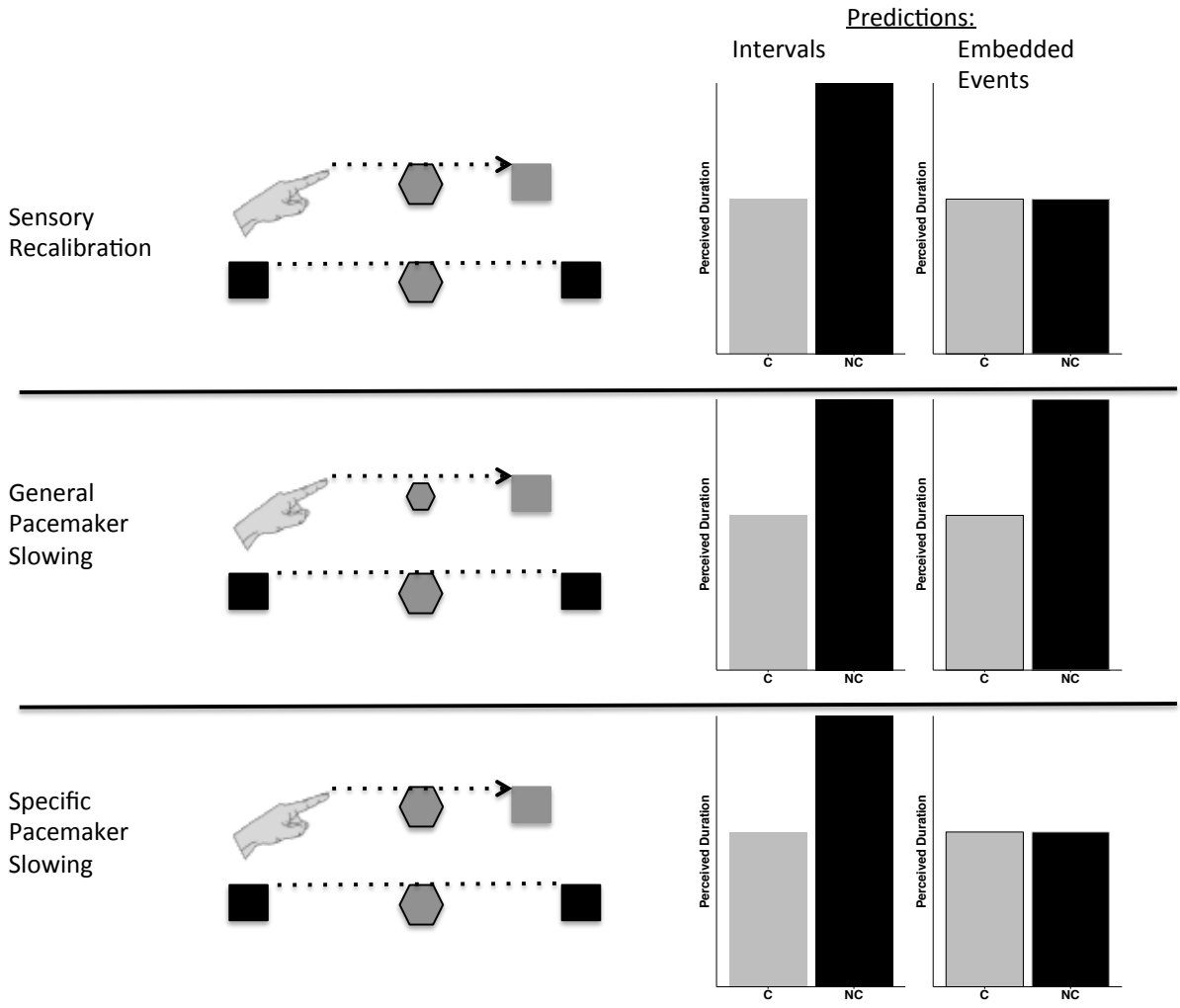

Figure 1. Predictions for trials with an embedded event *during* the interval. Interval judgments: we predict a temporal binding effect regardless of theoretical approach. Embedded event judgments: Only a general pacemaker slowing will result in shorter perceived embedded events during causal intervals. C: Causal; NC: Non-Causal.

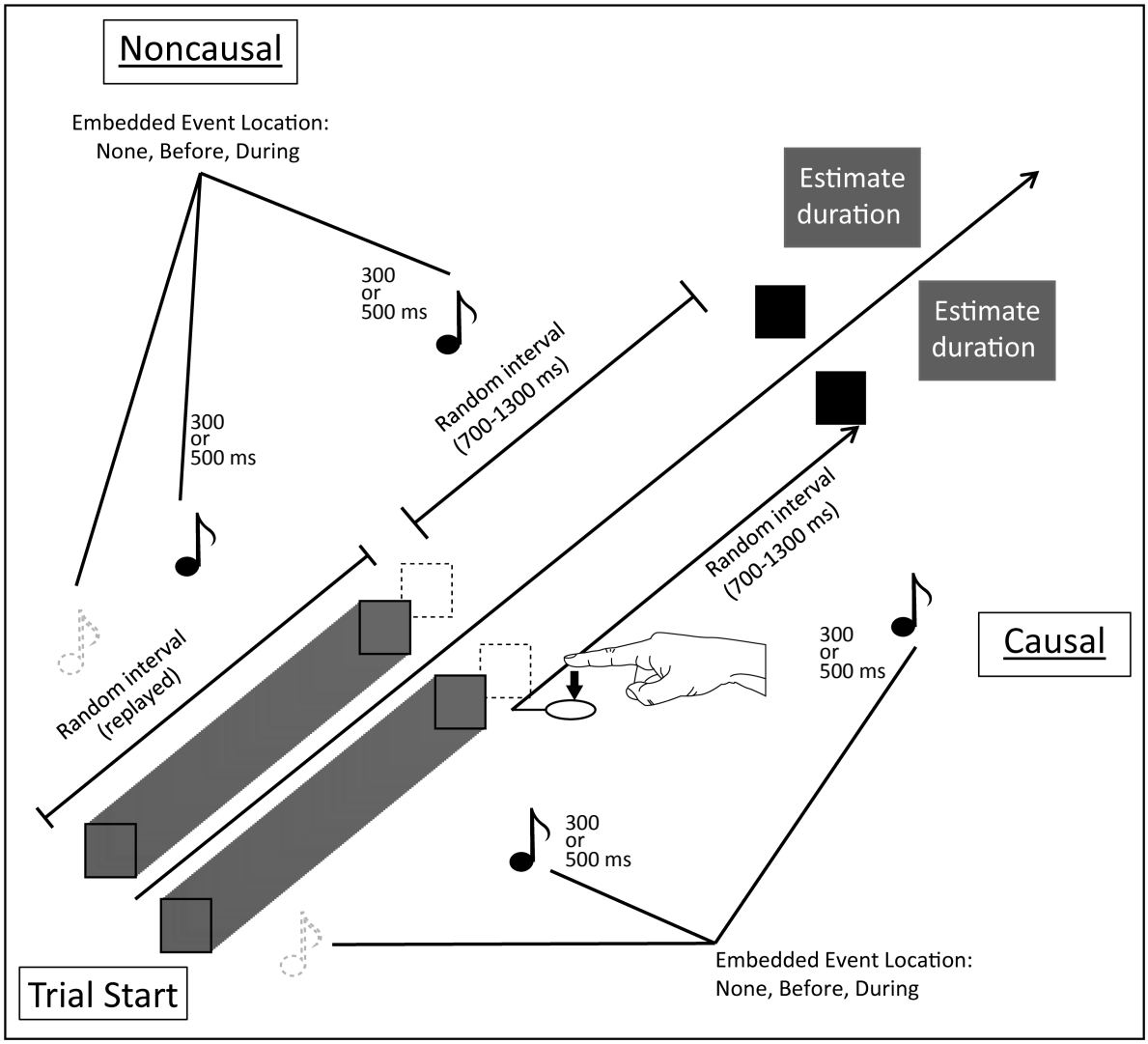

Figure 2. Trial structure of Experiment 1. In causal trials a black square was displayed until participants pressed a key, which led to the immediate disappearance of the square and triggered an outcome stimulus after a random interval (700 - 1300 ms). In noncausal trials a black square was displayed 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.

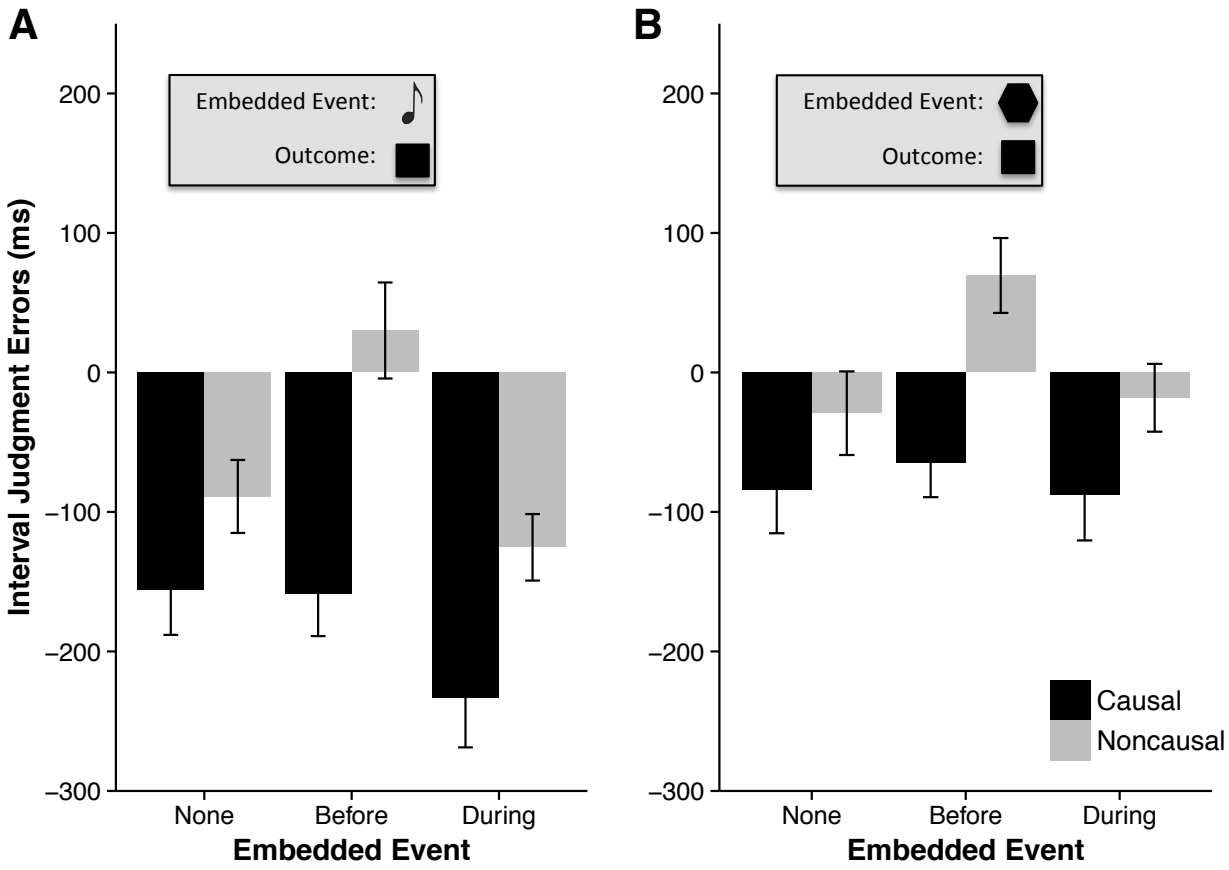

Figure 3. (A) Experiment 1. (B) Experiment 2. Mean interval judgment errors from causal and noncausal trials broken down by embedded event location. Error bars represent within-subjects confidence intervals (Cousineau, 2005; Morey, 2008).

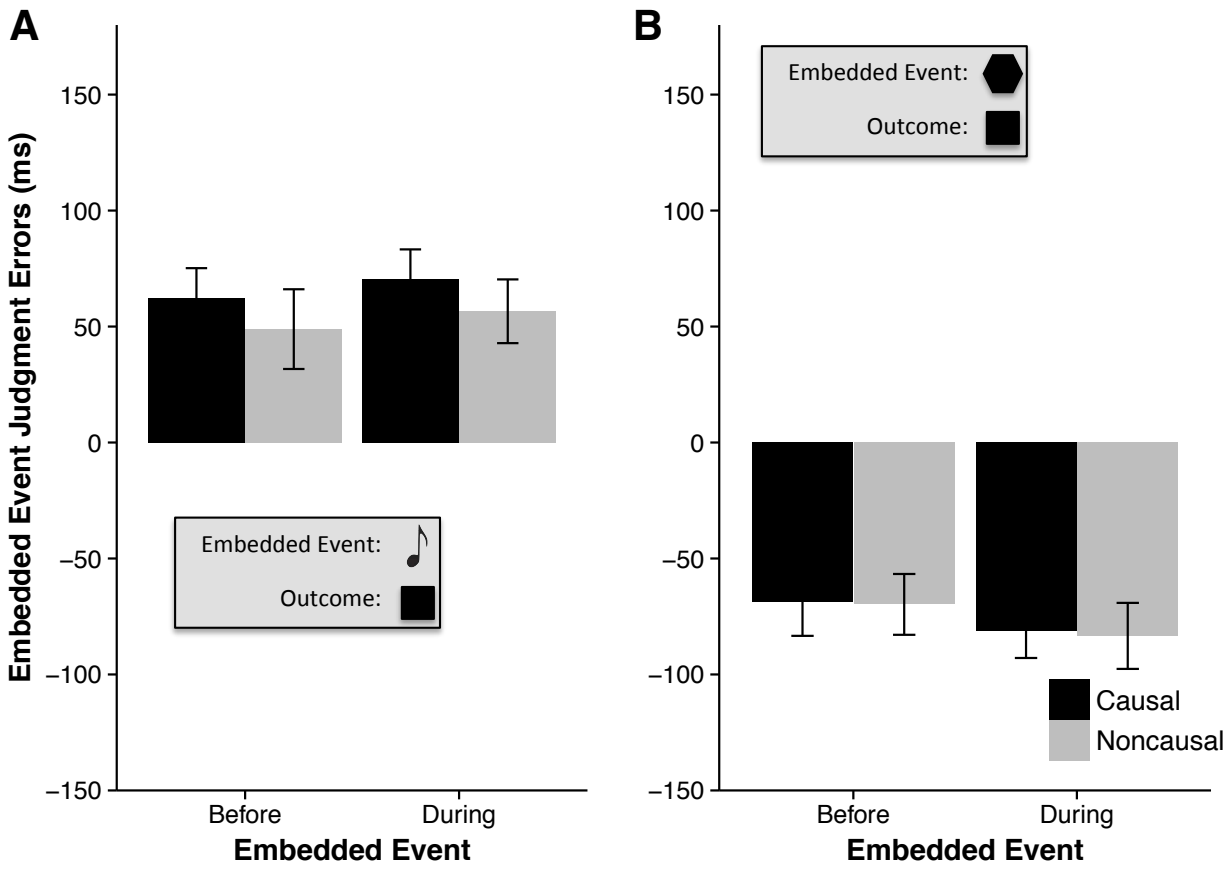

Figure 4. (A) Experiment 1. (B) Experiment 2. Mean embedded event judgment errors for events occurring before and during the interval, broken down by Trial Type. Error bars represent within-subjects confidence intervals.

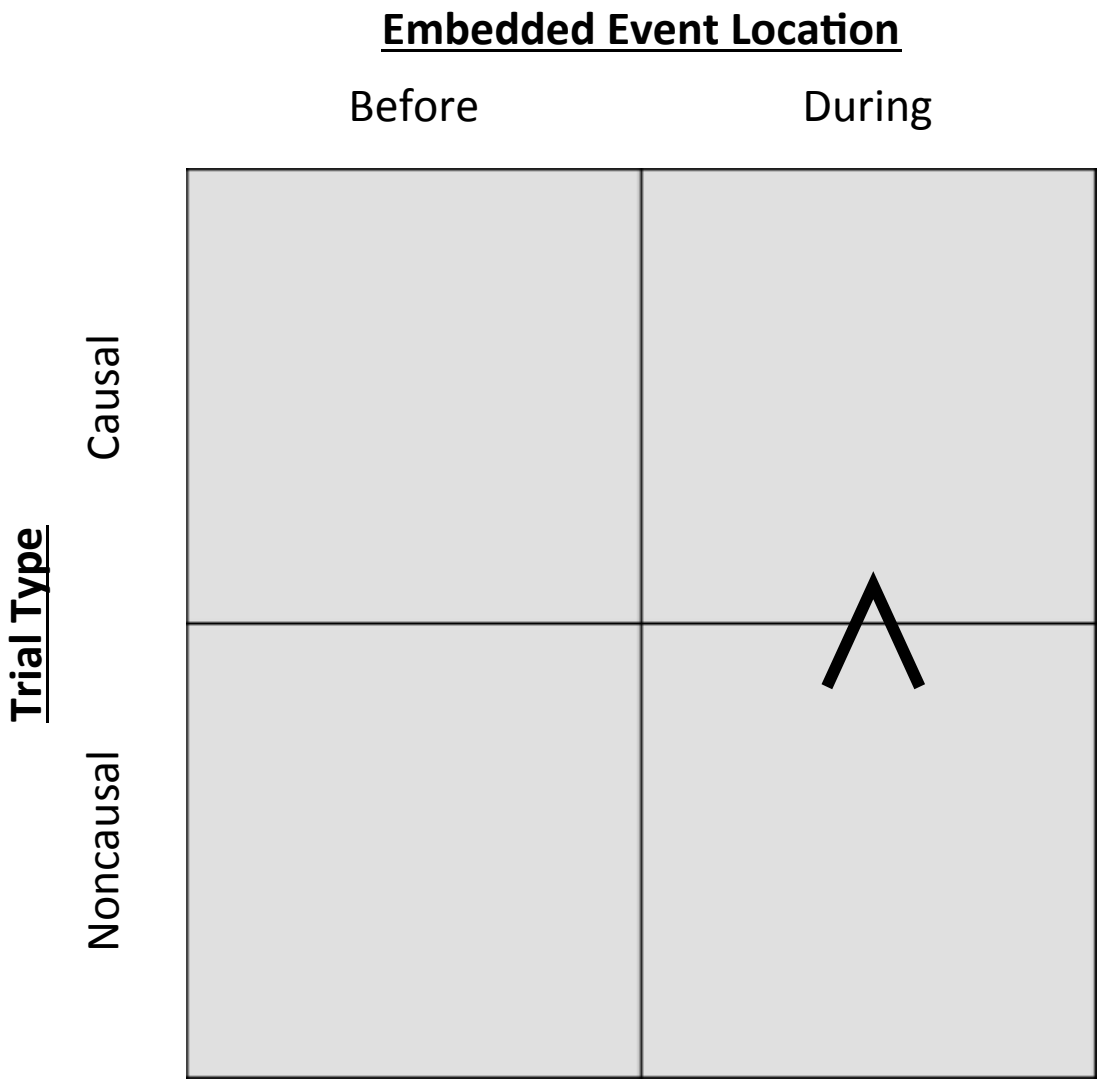

Figure 5. The restricted model entered into our Bayesian analysis. We specified a single constraint: events embedded during causal intervals would be judged shorter than events embedded into noncausal intervals. We made no assumptions about any other differences.

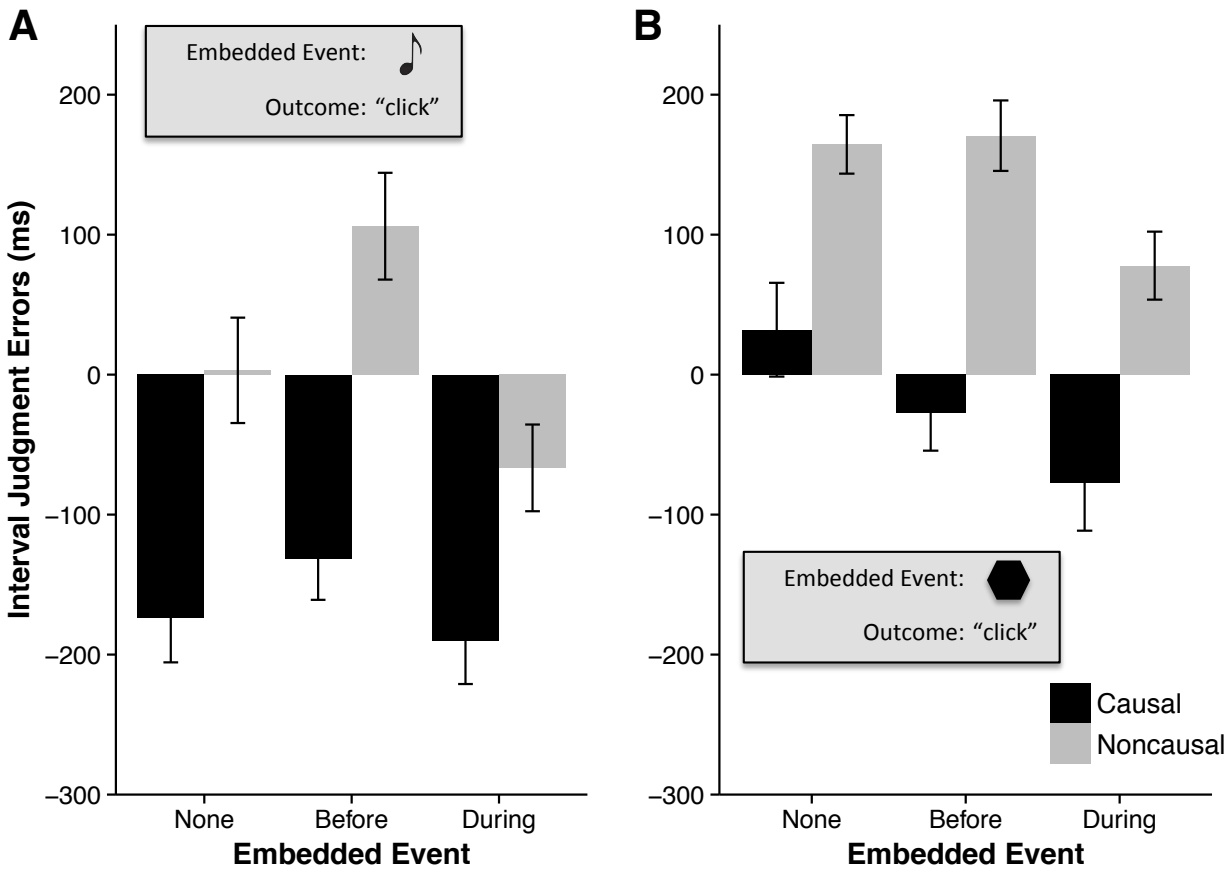

Figure 6. (A) Experiment 3. (B) Experiment 4. Mean interval judgment errors from causal and noncausal trials broken down by embedded event location. Error bars represent within-subjects confidence intervals.

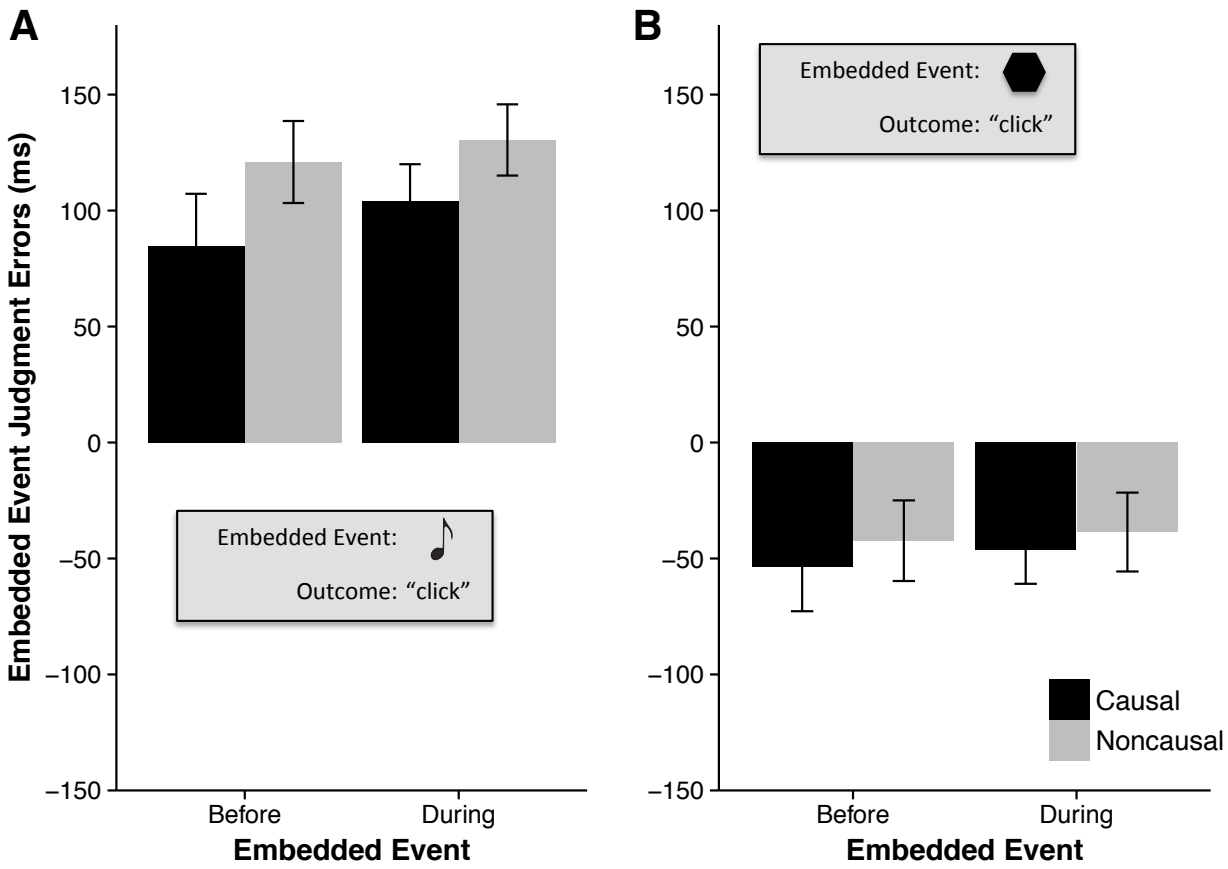

Figure 7. (A) Experiment 3. (B) Experiment 4. Mean embedded event judgment errors for events occurring before and during the interval, broken down by Trial Type. Error bars represent within-subjects confidence intervals.