Received: 8 May 2019 | Revised: 7 August 2019 | Accepted: 13 August 2019

DOI: 10.1111/ejn.14551

RESEARCH REPORT

EJN European Journal of Neuroscience FENS WILEY

Rapid recalibration of temporal order judgements: Response bias accounts for contradictory results

Brendan Keane[1] | **Nicholas S. Bland**[2,3] | **Natasha Matthews**[4] | **Timothy J. Carroll**[1] | **Guy Wallis**[1]

[1]Centre for Sensorimotor Performance, School of Human Movement and Nutrition Sciences, University of Queensland, Brisbane, Qld, Australia

[2]Queensland Brain Institute, University of Queensland, Brisbane, Qld, Australia

[3]School of Health and Rehabilitation Sciences, University of Queensland, Brisbane, Qld, Australia

[4]School of Psychology, University of Queensland, Brisbane, Qld, Australia

Correspondence
Brendan Keane, University of Queensland, Building 26, Brisbane, Qld 4067, Australia.
Email: b.keane1@uq.edu.au

Funding information
Australian Research Council; Australian Government Research Training Program

The peer review history for this article is available at https://publo ns.com/publon/10.1111/EJN.14551

Abstract

Recent history influences subsequent perception, decision-making and motor behaviours. In this article, we address a discrepancy in the effects of recent sensory history on the perceived timing of auditory and visual stimuli. In the synchrony judgement (SJ) task, similar timing relationships in consecutive trials seem more synchronous (i.e. less like the repeated temporal order). This effect is known as rapid recalibration and is consistent with a *negative* perceptual aftereffect. Interestingly, the opposite is found in the temporal order judgement (TOJ) task (*positive* rapid recalibration). We aimed to determine whether a simple bias to repeat judgements on consecutive trials (choice-repetition bias) could account for the discrepant results in these tasks. Preliminary simulations and analyses indicated that a choice-repetition bias could produce apparently *positive* rapid recalibration in the TOJ and not the SJ task. Our first experiment revealed no evidence of rapid recalibration of TOJs, but *negative* rapid recalibration of associated confidence. This suggests that timing perception was rapidly recalibrated, but that the *negative* recalibration effect was obfuscated by a *positive* bias effect. In our second experiment, we experimentally mitigated the choice-repetition bias effect and found *negative* rapid recalibration of TOJs. We therefore conclude that timing perception is *negatively* rapidly recalibrated, and this is observed consistently across timing tasks. These results contribute to a growing body of evidence that indicates multisensory perception is constantly undergoing recalibration, such that perceptual synchrony is maintained. This work also demonstrates that participants' task responses reflect judgements that are contaminated by independent biases of perception and decision-making.

KEYWORDS
confidence, multisensory, perception, psychophysics

1 | INTRODUCTION

It is well-established that our perception of the relative timing of simple physical events is malleable. For instance, repeated exposure to temporal asynchrony between physical signals (e.g. an audio-track lagging behind the picture at a cinema) leads to a shift in perception such that the events seem to occur more synchronously (e.g. Fujisaki, Shimojo, Kashino, & Nishida, 2004; Vroomen, Keetels, De Gelder, & Bertelson, 2004). This effect is known as temporal recalibration and can be induced experimentally by presenting participants with prolonged periods of asynchronous audiovisual signals. Interestingly, recent work has found that this effect

Abbreviations: JND, just noticeable difference; PLC, point of least confidence; PSS, point of subjective synchrony; SJ, synchrony judgement; SOA, stimulus-onset asynchrony; TOJ, temporal order judgement.

Edited by Guillaume Rousselet.

does not require extensive periods of adaptation to manifest. Instead, a single presentation is sufficient to influence subsequent timing perception (*rapid* recalibration). However, while temporal recalibration is consistent across timing tasks, rapid recalibration effects appear to differ in direction based on the timing tasks participants are required to complete (Roseboom, 2019).

Van der Burg, Alais, and Cass (2013) found that participants' perception of the synchrony of audiovisual events is biased towards the temporal relationship presented on the previous trial. That is, participants were more likely to report that an auditory and visual stimulus had occurred simultaneously if they had been presented at a similar timing relationship on the previous trial. This is consistent with a negative (or repulsive) aftereffect, as participants were less likely to perceive the stimuli as having occurred in the temporal order of stimuli on the preceding trial, and more likely to instead perceive either synchrony or the opposite temporal order. Interestingly, this effect was of a similar magnitude to that found following prolonged adaptation (e.g. Vroomen et al., 2004). Van der Burg et al. posited that this effect might reflect a rapid sensory adaptation process. However, the interpretation of these results has been recently challenged (e.g. Roseboom, 2019; Simon, Noel, & Wallace, 2017).

Roseboom (2019) hypothesised that if rapid recalibration reflects adaptation of sensory processing, it ought to occur in other timing tasks, which (presumably) rely on the same sensory processing mechanisms. To test this hypothesis, he had participants complete a synchrony judgement task (SJ; as in Van der Burg et al., 2013), a temporal order judgement (TOJ) task and a temporal magnitude judgement task. In the SJ task, Roseboom replicated the effects of Van der Burg and colleagues: participants were more likely to report synchrony when presented with a timing relationship similar to that presented on the previous trial (a *negative* rapid recalibration effect). However, participants displayed the opposite effect in both the TOJ task and temporal magnitude judgement task. That is, participants were *less* likely to perceive synchrony when presented with similar timing relationships in consecutive trials, and instead were more likely to report the repeated stimulus order (a *positive* rapid recalibration effect). Roseboom concluded that a purely sensory account of rapid recalibration cannot explain these discrepant effects, assuming that some sensory process generates a timing estimate which is then evaluated by subsequent decision-making processes. Instead, he proposed that rapid recalibration reflects changes in decisional processes and that these changes manifest differently for synchrony and temporal order judgement tasks.

The discrepancy in the direction of effects reported in TOJ and SJ tasks is difficult to reconcile with a purely sensory account of rapid recalibration and is inconsistent with temporal recalibration following prolonged exposure to asynchrony, where effects are similar across a range of timing tasks. Assuming rapid recalibration induces changes in timing perception consistent with those induced by prolonged exposure to asynchrony, both tasks should reveal *negative* rapid recalibration effects. However, since rapid recalibration effects depend on the previous trial, any serial dependency in the participants' responses may impact the observed recalibration effect. Due to the manner in which rapid recalibration is determined for order judgement tasks (TOJ and temporal magnitude judgements), the effect is vulnerable to contamination by choice-repetition biases (Pape, Noury, & Siegel, 2017). Similar ideas are now emerging in related fields, such as serial dependence of orientation perception (e.g. Gekas, McDermott, & Mamassian, 2019; Zhang & Alais, 2019).

Here, we aimed to resolve the key issue of whether *positive* rapid recalibration of TOJs reflects a choice-repetition bias, unrelated to timing perception. We first show (see Choice-repetition bias) that *positive* rapid recalibration effects may manifest during TOJ (but not SJ) tasks as a result of a choice-repetition bias. We then show that such a bias is present in Roseboom's data and predicts the discrepant rapid recalibration effects that he reported in temporal order and magnitude judgement tasks. In Experiment 1, we used confidence as a secondary probe of timing perception. We failed to replicate rapid recalibration of TOJs, but found evidence of *negative* rapid recalibration in associated confidence. In Experiment 2, we dissociated the preceding SOA from participants' preceding TOJ, thereby mitigating the effect of a choice-repetition bias on TOJs. Using this paradigm, we found evidence of *negative* rapid recalibration of TOJs. The results of these studies indicate that timing perception rapidly recalibrates consistently across timing tasks.

2 | CHOICE-REPETITION BIAS

In general, participants are reasonably sensitive and accurate when making judgements of stimulus order (e.g. Donohue, Woldorff, & Mitroff, 2010). If the visual stimulus was leading on the current trial, the participant would likely report a visual-lead. In the subsequent trial, regardless of the timing of the stimuli, a simple bias to repeat their previous TOJ would incline the participant to report another visual-lead stimulus-onset asynchrony (SOA). This would uniformly increase their probability of reporting a visual-lead for any SOA following a visual-lead. The consequence, then, is that the area under their discriminant function would increase, shifting their point of subjective synchrony (PSS) towards audio-lead SOAs. Similarly, if the previous trial was an audio-lead, the participant likely reported an audio-lead, and a choice-repetition bias would slightly increase the probability that the participant would do so again in the subsequent trial. This would shift the PSS for trials following audio-lead SOAs towards a visual-lead.

When combined, these choice-repetition bias effects produce a pattern of TOJs that mimics the *positive* recalibration effect reported by Roseboom (2019). Of course, a bias to *switch* between TOJs would mimic a *negative* rapid recalibration effect. Regardless of choice-history bias—either repetition or alternation—the preceding SOA is confounded by the preceding TOJ when task performance is greater than chance. We refer to a choice-repetition bias here as it would account for the discrepancy between TOJ and SJ, and is more often found than a choice-alternation bias (e.g. Pape et al., 2017; Zhang & Alais, 2019). Importantly, while a choice-repetition bias would also produce changes in SJ distributions, it would only do so in terms of the response distribution amplitude, with no change in PSS (in this case taken as the mean of the distribution). We propose that differences in how a choice-repetition bias manifests in these two tasks may explain the discrepancy in reported rapid recalibration effects.

To test this prediction, we re-analysed data made available by Roseboom (2019; Audiovisual synchrony, temporal order and, magnitude judgements; https://doi.org/10.1037/xhp0000591.supp). We first partitioned participants' timing judgements into two data sets, based on their judgement in the preceding trial. We fitted logistic functions to temporal order and magnitude judgements, and Gaussian functions to synchrony judgements, separately for each partition of the data (including each participants' overall data set, ignoring stimulus order or judgement on the preceding trial). These functions provided us with an estimate of each participants' point of subjective synchrony (PSS) as a function of their preceding judgement. We took the difference in PSS values as an estimate of the effect of the preceding judgement on participants' timing perception. We used frequentist statistical tests to make inferences about the probability of effects of at least the observed magnitude under the null hypothesis, with an arbitrary significance criterion of $\alpha = 0.05$. Where possible, we also include effect size estimates (Cohen's *d*). This analytic approach is identical to that used to quantify rapid recalibration, except here we partitioned the data set based on the preceding judgement, as opposed to the temporal order of stimuli on the preceding trial.

Using this method, we found that participants in Roseboom's (2019) study exhibited striking biases to repeat choices on consecutive trials in both the SJ and TOJ tasks (and in the magnitude judgement task; not shown here). In the SJ task, we compared participants' judgements following trials where they reported synchrony against those following trials where they reported asynchrony (Figure 1a). Similarly, in the TOJ task, we compared participants' responses on trials following an audio-lead (AV) report versus a visual-lead (VA) report (see Figure 1b). The choice-repetition bias manifested as a change in the amplitude of the SJ distribution (see Figure 1a), with no systematic change in PSS. However, this

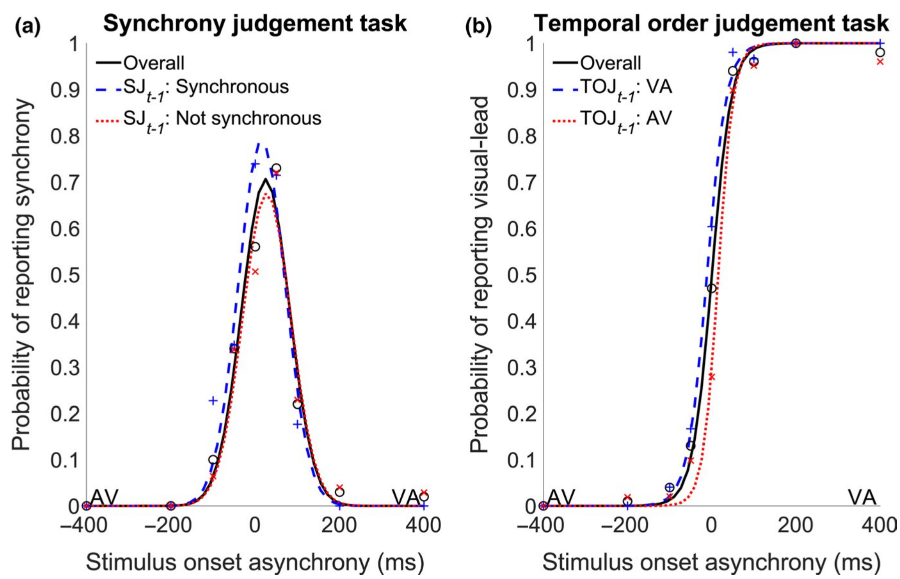

FIGURE 1 Choice-repetition biases in synchrony and temporal order judgements. These images are generated from a re-analysis of data reported in Roseboom (2019). In each panel, three data series are plotted: black, solid line, circle markers) overall response distributions, regardless of the judgement in the preceding trial; blue, dashed line, + markers) judgements from trials following a report of either synchrony or visual-lead; red, dotted line, x markers) judgements from trials following a report of either asynchrony or audio-lead. (a) A choice-repetition bias produces a shift only in the amplitude of the synchrony judgement response distribution, but not the mean. In this sense, the synchrony judgement task prevents any simple response bias (choice-repetition or otherwise) from producing false rapid recalibration effects. (b) A choice repetition bias in the temporal order judgement task produces shifts in the mean of the response distribution, unlike the synchrony judgement response distribution. Comparison of distributions reveals a positive rapid recalibration effect of temporal order judgements (despite being analysed in terms of preceding judgement, rather than preceding timing relationship) [Colour figure can be viewed at wileyonlinelibrary.com]

bias shifted the PSS in the TOJ task, consistent with *positive* rapid recalibration.

We then investigated the contribution of choice-repetition bias effects to the apparent rapid recalibration effects. For each participant, we calculated their rapid recalibration effect by computing the difference in their PSS for TOJs following AV (i.e. audio-lead SOA) trials versus those following VA (i.e. visual-lead SOA) trials. In both the temporal order and magnitude judgement tasks, paired t tests (and Cohen's d effect size statistics) indicate that the choice-repetition bias effect ($M = -99.00$ ms, $M = -70.65$ ms, respectively) is significantly larger ($\alpha = 0.05$) than the rapid recalibration effect ($M_{\text{TOJ}} = -26.00$ ms, $t_{18} = -3.44$, $p = .003$, $d = -0.75$; $M_{\text{Mag}} = -33.54$ ms, $t_{16} = -3.60$, $p = .002$, $d = -0.77$), with medium/large effect sizes. Further, the magnitude of these two effects was positively correlated in the TOJ ($r_{17} = .71$, $p = .001$) and magnitude judgement tasks ($r_{15} = .65$, $p = .005$), even capturing the *negative* rapid recalibration of participants with choice-*alternation* biases. Critically, and consistent with a choice-repetition bias, we found no evidence that the central-tendency of synchrony judgement distributions changed as a function of the preceding SJ ($t_{18} = -1.29$, $p = .215$, $d = -0.13$), but significant changes in the amplitude of the synchrony judgement distribution ($t_{17} = 6.58$, $p < .001$, $d = 0.74$; see Figure 1a).

We then attempted to quantify the magnitude of the choice-repetition bias required to elicit a false rapid recalibration effect. We simulated observers in a TOJ task with varying probabilities of simply repeating their preceding TOJ. On each trial, a simulated observer perceived an SOA drawn from a random Normal distribution with a mean of the actual SOA and variance of 80 ms, and compared this to a fixed criterion at 0 ms to make a temporal order judgement. On some proportion of trials, the simulated observer simply repeated the previous TOJ, regardless of SOA. We simulated 20 of these observers completing an experiment using the same SOAs as Roseboom (2019), with the same number of trials, split into the same number of testing blocks. This experiment was simulated 1,000 times per repetition rate. We found that the proportion of trials required to produce significant rapid recalibration effects was alarmingly low (see Figure 2).

We found that the probability of finding a significant 'rapid recalibration' effect exceeded 90% with a choice-repetition rate of only 4%. That is, if participants repeat their preceding TOJ on only four out of every 100 trials, perhaps due to lapses in concentration, there is a ~93% chance of finding evidence of a *positive* rapid recalibration effect. This is alarming, and plausible. Pape et al. (2017) found that participants were strongly biased to repeat their orientation judgements in consecutive trials, even with intervening button-presses and a random response-mapping on each trial. Verstynen and Sabes (2011) asked participants to reach to targets incrementally rotating around a circular track. Despite the target location being perfectly predictable, participants were biased to reach towards the location where the target had just been. Marinovic, Poh, de Rugy, and Carroll (2017) reported a similar effect, where participants were consistently biased towards an irrelevant location to which they had previously moved, even when given explicit instruction that the location was irrelevant. The bias effect remained even when participants were provided a long time (~1 s) to prepare each movement after target presentation. These results suggest that choices, and the motor behaviours that enact them are reliably biased towards preceding instances, despite explicit instructions, predictable patterns, and ample preparation time.

We took a very simple approach to quantifying the repetition rate in Roseboom's (2019) data. In the temporal order and magnitude judgement tasks, we computed the probability of each response being the same as the response made on the preceding trial (based only on stimulus order for the magnitude judgement task). We then subtracted the probability of the correct response repeating in consecutive trials, to estimate the repetition bias over and above the repetition built into the actual stimulus order. In temporal order and magnitude judgement tasks, we found that Roseboom's participants repeated their responses on ~56% of trials, while the correct response (dictated by the stimulus order) was repeated only ~50% of the time (as expected). According to our simulations, a 6% choice-repetition rate is extremely likely (>99%) to produce false evidence of *positive* rapid recalibration. In fact, of our 1,000 simulations of a repetition rate of 6%, every single sample returned a significant *positive* rapid recalibration effect.

In our first experiment, we aimed to replicate the reported *positive* rapid recalibration of TOJs, while also measuring participants' confidence in their TOJs. This design allows us to exploit the shape of the confidence distribution, which ought to be immune to a choice-repetition bias (as in SJ tasks). If timing perception rapidly recalibrates *negatively*, similar timing relationships in consecutive trials should seem more synchronous, or less ordered, and so be associated with lower confidence. In essence, confidence in timing judgements should rapidly recalibrate in line with rapid recalibration of timing perception itself.

3 | EXPERIMENT 1—METHODS

Fifty adults (21 male, 29 female; ages ranged 18–33 years, $M = 21.83$) were paid to participate in a single one-hour testing session involving an audiovisual TOJ task. Participants gave written informed consent before testing started and were remunerated $20/hr for their time. This study was approved by the ethics committee of the School of Human Movement and Nutrition Sciences at The University of Queensland (project number 2018002472).

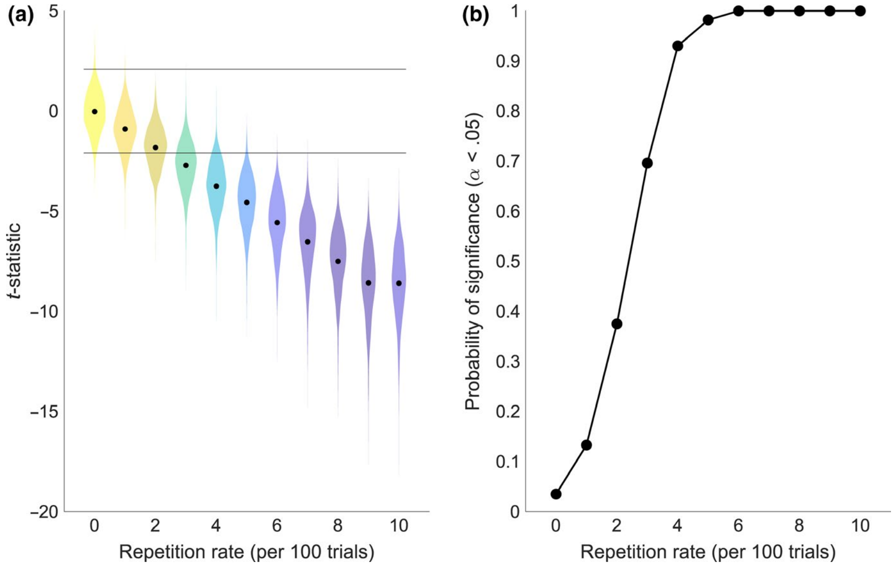

FIGURE 2 Results of simulated experiments. We simulated the effect of choice-repetition rates ranging from 0 to 10 repetitions per 100 trials, in experiments with a sample size of 20. Note that each participant in these simulated experiments had the same repetition-rate and perceptual sensitivity, but a random trial order. (a) We conducted 1,000 virtual experiments for each repetition-rate, and tested for rapid recalibration. Black horizontal bars indicate the t-statistic corresponding to significance ($\alpha < 0.05$) in a paired-samples *t*-test with $df = 19$. (b) The proportion of *p*-values from these *t*-tests that exceed the significance criterion. That is, the probability of finding a significant rapid recalibration effect as a function of repetition-rate. Note that the simulated observer was coded to make completely independent responses across trials (i.e., no form of 'rapid recalibration' was built into the simulated observer), except for the variable proportion in which they simply repeated their preceding TOJ regardless of current-trial SOA [Colour figure can be viewed at wileyonlinelibrary.com]

Participants reported having normal or corrected-to-normal hearing and vision and were naïve to the purposes of the experiment. Two participants were removed from analyses because their sensitivity to temporal order was extremely poor (just-noticeable difference greater than the sampling range, indicating they could not reliably discern the temporal order of even the most asynchronous stimuli, offset by 300 ms).

Note that the sample of participants analysed here is aggregated from two iterations of this experiment. The first sample consisted of 20 participants, completing 600 trials of a TOJ task with a confidence judgement. We decided to re-run this study with another (independent) sample as the original sample showed no evidence of rapid recalibration of TOJs in either direction, and we were concerned we had found a Type 2 error. The second iteration of the experiment comprised 50% more participants ($N = 30$), and 20% more trials (720). These participants also completed a short familiarisation block of trials before starting the experiment, with no feedback, to give them the opportunity to ask specific questions about the task. We believe these samples are appropriately matched in terms of experimental procedures and demographic properties. Analysis of each sample separately yields results consistent with the aggregate sample, and so for simplicity we report the pooled sample.

Participants sat in a dimly lit room with their chin placed in a chin rest 57 cm from the computer monitor. A small red LED (0.5 cm diameter; 0.5 degrees of visual angle, DVA) was attached to the centre of the monitor, and speakers were hidden behind the monitor. On each trial, the red LED was illuminated for 20 ms, and the speakers played a square-wave tone of 440 Hz for 20 ms. The auditory stimulus was onset at one of six SOAs relative to the visual stimulus ($\pm$300, $\pm$200 and $\pm$ 100 ms; i.e. never synchronous). Each SOA was sampled 100 times in pseudorandom order for the first 20 participants, and 120 times each for the remaining 30 participants. Trial structure is schematised in Figure 3. Audio and visual stimuli were generated by an independent microcontroller. Stimulus presentation timing was validated prior to data collection using an oscilloscope and also measured during testing using another independent microcontroller. Online measurement allowed for rejection of trials in which stimulus presentation timing deviated from intended timing. Note that no trials were rejected on this basis. Stimulus presentation timing did not deviate from intended timing by more than 0.24 ms on any trial for any participant ($M = -6$ μs, $SD = 86$ μs).

Following stimulus presentation, participants were asked to report which of the audio and visual stimuli had occurred first. Participants were prompted by the appearance of a text message on screen behind the (now dim) red LED, informing them to press either the left or right mouse button to indicate their choice. Participants were then prompted to report their confidence (either high or low; see Figure 3) in

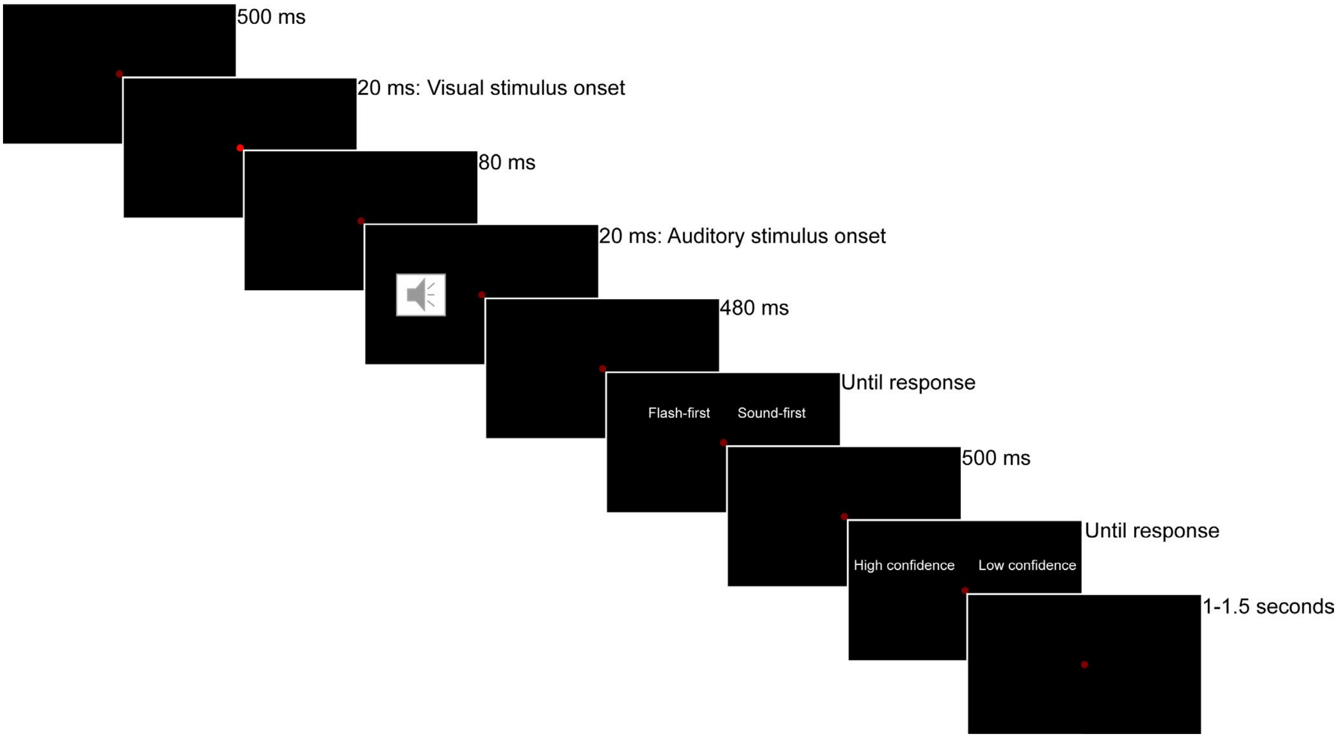

FIGURE 3 Schematic of the trial structure, with an SOA of 100 ms. Each trial was preceded by a period of silence and a blank screen lasting between 1 and 1.5 s. The trial period then started, lasting a total of 1.1 s. The visual stimulus appeared 500 ms into the trial period on every trial, while the auditory stimulus onset at any of the six SOAs relative to the visual stimulus (±300, ±200, and ± 100 ms). Note that the 'loudspeaker' symbol appearing in the figure is diagrammatic only and no visual cue was provided with audio onset. Note that the red LED was always visible, but only became illuminated (here in the second panel) to indicate the visual signal (i.e., it was otherwise dim) [Colour figure can be viewed at wileyonlinelibrary.com]

their TOJ. Response mappings for temporal order and confidence judgements were counterbalanced across participants. Participants were given a one minute break every 100 trials (approximately every seven minutes). The first trial, and the trial immediately following each break, was excluded from analyses since no trial immediately preceded them.

3.1 | Data analysis

As in our re-analysis of Roseboom's (2019) data, participants' TOJs were partitioned based on the temporal order of stimuli on the preceding trial, and logistic functions were fitted to these data sets. These fitted functions provided estimates of the PSS and sensitivity to temporal order (JND). We computed rapid recalibration by taking the difference in PSS values for the two partitions; the PSS when the previous SOA was a visual-lead minus the PSS when the previous SOA was an audio-lead.

We used a very similar process to quantify the effect of temporal order on the preceding trial on participants' confidence judgements. We fitted Gaussian functions to each participant's probability of reporting low confidence at each of the six SOAs, for the two data sets. This provided an estimate of the timing relationship for which the participant was least confident (point of least confidence, PLC), and the variance in their low-confidence responses (σ). As for order judgements, we estimated the effect of the preceding stimulus order on confidence judgements by computing the difference in PLCs.

Participants were excluded from analyses if their results for any variable of interest exceeded the sample mean by more than two standard deviations, or if their JND (or variance in low-confidence responses) exceeded 300 ms (the extent of the sampling range). Note that this was conducted separately for analyses of rapid recalibration, biases due to preceding response, and analysis of choice-repetition bias. We used this approach as each suite of analyses was based on model parameters from different partitions of the data, occasionally producing extreme parameter estimates (due to poor model-fits). Importantly, any participant with an extreme overall TOJ JND was excluded from all analyses. As in our preliminary work, we used frequentist statistical tests to estimate the probability that effects of at least the observed magnitude would be drawn from the null distribution, with an arbitrary significance criterion of $\alpha = 0.05$. Analysis code (in MATLAB) and raw data files have been made available online (Rapid recalibration of temporal order judgements; https://doi.org/10.17605/OSF.IO/CNQ4V).

4 | EXPERIMENT 1—RESULTS

4.1 | Rapid recalibration analysis

We first calculated participants' overall bias and sensitivity (PSS and JND, respectively) to audiovisual temporal order. On average, participants were accurate (M_{PSS} = 18.61 ms, *SD* = 102.10 ms) and precise (M_{JND} = 115.59 ms, *SD* = 73.73 ms). Paired *t* tests revealed no evidence that the temporal order of stimuli on the preceding trial influenced participants' perception of synchrony ($t_{46} = -1.52$, $p = .134$, $d = -0.11$). Similarly, we found no evidence that stimulus order on the preceding trial influenced participants' sensitivity to temporal order (as predicted; $t_{46} = -1.48$, $p = .147$, $d = -0.11$). The size of these effects (measured using Cohen's *d*) suggests that if there is an effect here, it is quite small. Contrary to predictions, and unlike previous reports (e.g. Van der Burg et al., 2013), we found no evidence that participants' sensitivity to temporal order was correlated with the magnitude of their rapid recalibration effect ($r_{45} = .01$, $p = .970$; Figure 4).

4.2 | Confidence analysis

We then fit Gaussian functions to participants' probability of reporting low confidence across each of the six SOAs, independently for each of the two data sets (SOA_{t-1}: VA and SOA_{t-1}: AV). This provided an estimate of the point at which each participant was least confident (PLC), and the variance in their confidence judgements. On average,

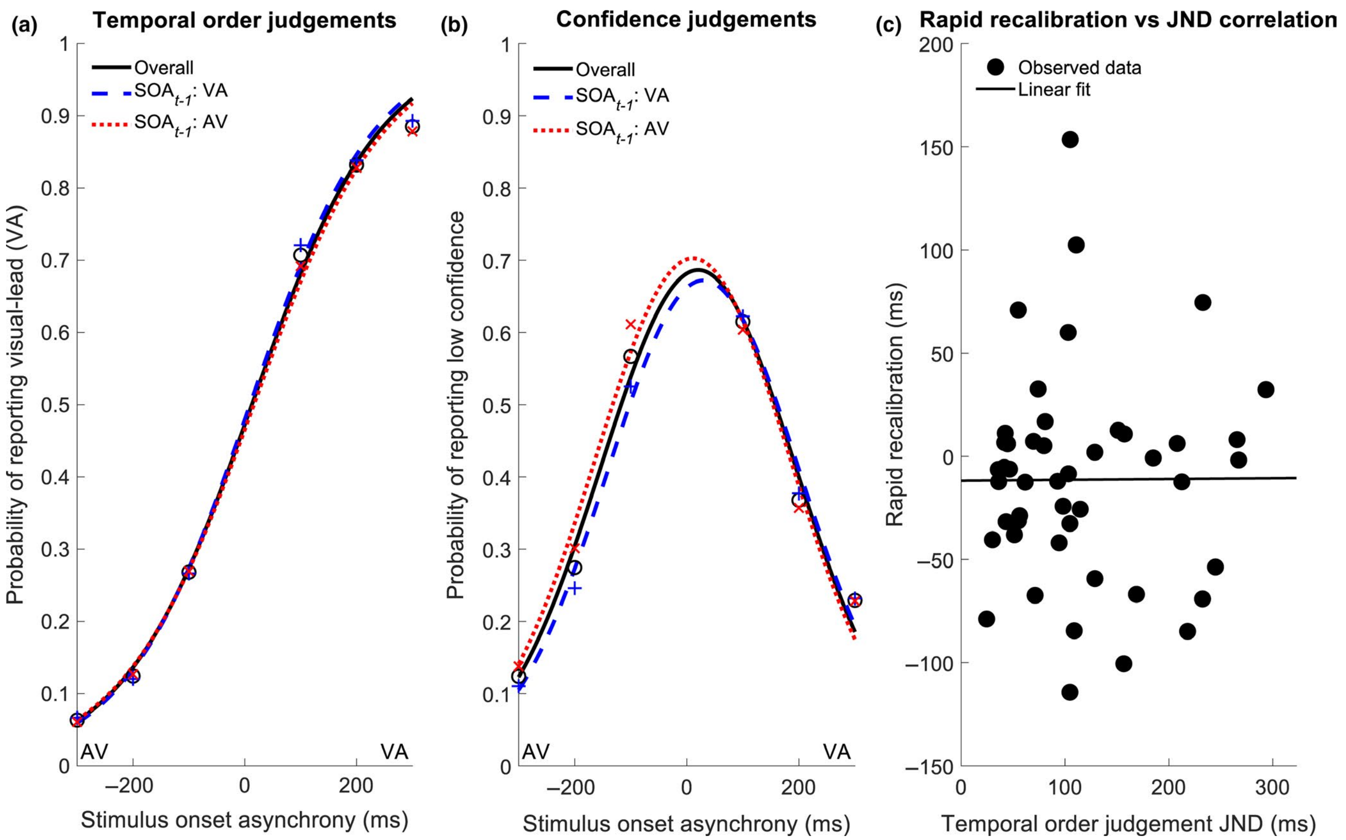

FIGURE 4 Temporal order and confidence response distributions, collated across participants, and correlation of rapid recalibration with timing sensitivity. Three data series are plotted: black, solid line, circle markers) overall response distributions, regardless of the timing relationship in the preceding trial; blue, dashed line, + markers) judgements from trials following a visual-lead; red, dotted line, x markers) judgements from trials following an audio-lead. (a) Participants' temporal order judgements indicate sensitivity to physical timing, as expected. However, there is no evidence of an effect of the order of stimuli on the preceding trial, contrary to predictions. (b) Overall, participants were most likely to make TOJs with low confidence when the physical timing offset was near zero, as expected. Participants were more likely to report low confidence for timing relationships similar to that presented in the preceding trial. This is evidenced by a significant difference between the mean of the red and blue distributions, corresponding to trials following audio-leads and audio-lags, respectively. This is consistent with a negative rapid recalibration of timing perception, with similar SOAs in consecutive trials seeming more synchronous, or less ordered, and therefore judged with lower confidence. (c) We found no evidence that participants' sensitivity to temporal order (measured using the JND) was linearly related to the magnitude of their rapid recalibration effect, contrary to the findings of Van der Burg et al. (2013) and Simon et al. (2017) [Colour figure can be viewed at wileyonlinelibrary.com]

participants reported least confidence for 36.92 ms visual-leads (SD = 49.65 ms), indicating they were sensitive to the difficulty of the task (mean σ = 189.63 ms, SD = 86.55 ms).

Paired t tests indicated that the temporal order of stimuli on the preceding trial reliably influenced participants' PLC (t_{47} = 2.98, p = .005, d = 0.32), consistent with *negative* rapid recalibration. Cohen's d indicates that this is a reasonably small effect; however, this is observed in a measure only indirectly probing timing perception. As predicted, neither a paired-sample t test nor Cohen's d indicated the temporal order of stimuli on the preceding trial influenced the variance in participants' confidence judgements (t_{47} = 1.07, p = .292, d = 0.06). We found no evidence for a relationship between participants' recalibration of low confidence and the variance in their reports of low confidence (r_{46} = −.16, p = .280).

4.3 | Choice-repetition bias analysis

We then quantified the effect of recent responses on participants' TOJs and associated confidence. Figure 5 shows participants' TOJs when split by their preceding TOJ (Figure 5a) and their confidence in the preceding TOJ (Figure 5c). Note that their confidence judgement for the preceding TOJ (Figure 5c) corresponds to the button-press immediately preceding the plotted TOJs. Paired t tests on model parameters indicated that participants' PSS varied as a function of the preceding TOJ (Figure 5a; t_{46} = −3.50, p = .001, d = −0.54),

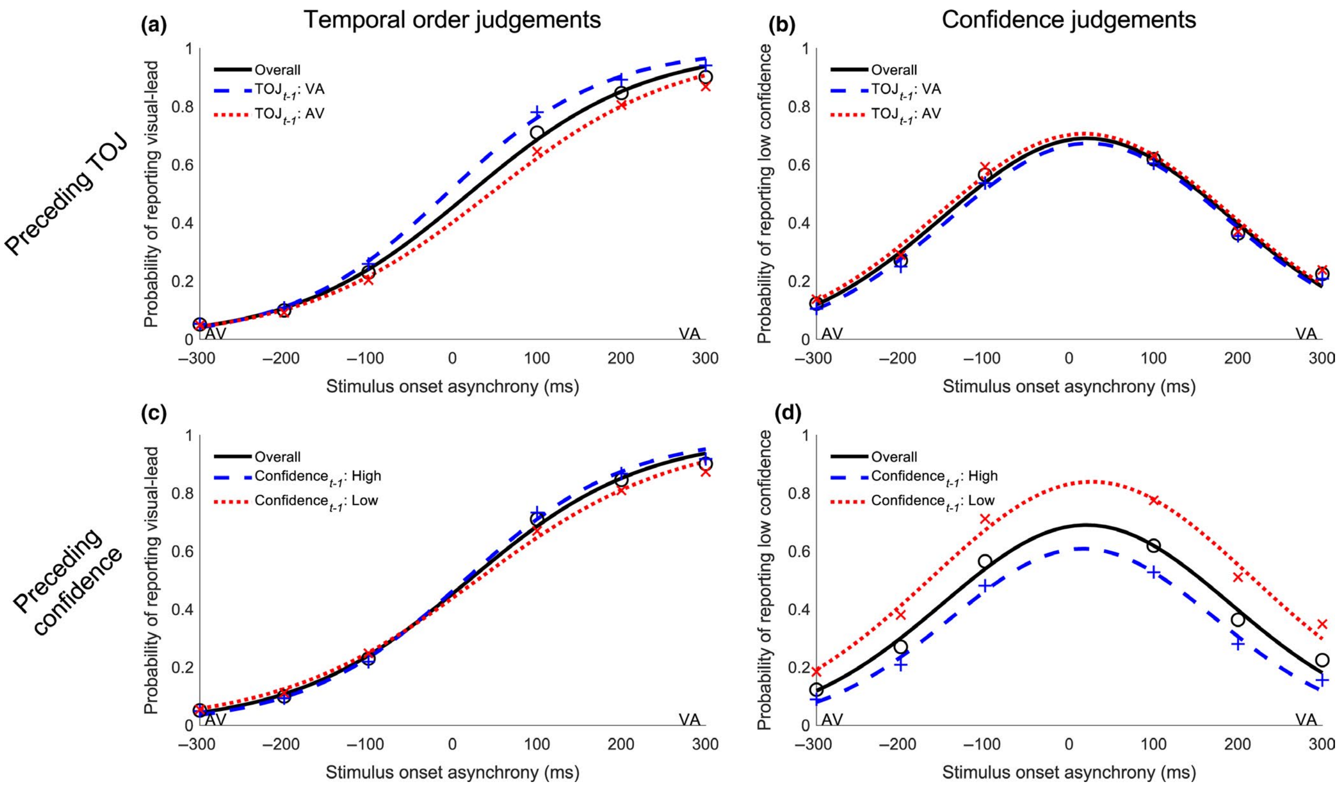

FIGURE 5 Participants' judgements as a function of preceding judgements. In order to investigate choice-repetition biases, we split participants' temporal order and confidence judgements by the preceding temporal order and confidence judgement. In each panel, three data series are plotted: black, solid line, circle markers) overall response distributions, regardless of the choice in the preceding trial; blue, dashed line, + markers) judgements from trials following a visual-lead judgement, or high confidence; red, dotted line, x markers) judgements from trials following an audio-lead judgement, or low confidence. Top row) Participants' TOJs and confidence judgements as a function of their preceding TOJ. (a) We found an effect of the preceding TOJ on subsequent TOJs. When the participants' preceding TOJ was a visual-lead (TOJ$_{t-1}$: VA), participants were more likely to report a visual-lead again. The same is true for TOJs following an audio-lead judgement. (b) However, we found no evidence that participants' previous TOJ biased confidence judgements, despite evidence that the preceding SOA influenced confidence (Figure 4b). This suggests that the observed rapid recalibration of low-confidence is not due to a simple response bias. Bottom row) Participants' temporal order and confidence judgements as a function of the preceding confidence judgement. Note that the preceding confidence judgement is the button-press immediately before these order judgements are made. (c) We found no evidence that the preceding confidence report influenced subsequent TOJs. This suggests that the choice repetition bias is not a simple response bias, but relates specifically to the judgement (as opposed to the motor behaviour required to report that judgement). (d) Participants' confidence on the preceding trial, however, did influence their subsequent confidence. Specifically, when participants reported low-confidence on some trial, they were more likely to do so again on the subsequent trial. Note that this bias leads only to a vertical shift in the fitted distributions (panel d), and does not lead to a lateral shift as it can for TOJs (panel a) [Colour figure can be viewed at wileyonlinelibrary.com]

but not the preceding confidence judgement (Figure 5b; $t_{46} = 0.01$, $p = .989$, $d < 0.01$). This is consistent with participants' being biased to repeat their most recent TOJ (and not their most recent button-press). In Roseboom's (2019) data, we found a strong correlation between the rapid temporal recalibration and response bias effects, in both TOJ and temporal magnitude judgement tasks (ps < .05). Similarly, we found that participants' choice-repetition bias (change in PSS as a function of preceding TOJ) was strongly correlated with their rapid recalibration effect ($r_{43} = .76$, $p < .001$).

Using the same analytic approach, we then quantified the effect of participants' preceding responses on their decisional confidence. Figure 5 shows participants' probability of reporting low confidence (right column), as a function of their preceding TOJ and confidence. A similar set of results emerges for confidence judgements; we found no evidence that participants' confidence was changed by their preceding TOJ (in terms of amplitude, mean and standard deviation; Figure 5b; ps > .05), but clear evidence that confidence varied as a function of their preceding confidence (Figure 5d). Paired t tests and Cohen's d indicate participants were slightly more likely to report low confidence if they did so on the previous trial ($t_{45} = 2.41$, $p = .020$, $d = 0.18$), and produce slightly broader low-confidence distributions ($t_{45} = 1.82$, $p = .076$, $d = 0.19$), with weak/mixed evidence for a change in PSS ($t_{45} = 1.45$, $p = .153$, $d = 0.31$). Similar to the bias observed for TOJs, this indicates a bias to repeat their preceding confidence judgement.

5 | EXPERIMENT 1—DISCUSSION

We aimed to investigate the effect of stimulus order on subsequent timing judgements and associated confidence. If temporal order perception rapidly recalibrates in line with synchrony perception (e.g. Van der Burg, Orchard-Mills, & Alais, 2015; Van der Burg et al., 2013), similar SOAs in consecutive trials would appear more synchronous, shifting the TOJ PSS towards the preceding SOA (a *negative* aftereffect) and lowering participants' confidence (as they now seem less ordered). Alternatively, if the effect reported by Roseboom (2019) reflects genuine rapid recalibration of temporal order perception, participants' TOJ PSS should shift away from the preceding SOA (a *positive* aftereffect), and confidence ought to be increased for judgements of similar SOAs in consecutive trials (as they seem less synchronous, or *more* ordered). We found no evidence of rapid recalibration of temporal order perception, consistent with neither hypothesis, despite using a much larger sample size than most, and comparable trial numbers.

However, we found evidence that confidence judgements varied from trial to trial, consistent with the effect predicted by *negative* rapid recalibration (Simon et al., 2017; Van der Burg et al., 2013, 2015). Participants were generally less confident in their TOJs if similar SOAs occurred in consecutive trials, plausibly because they seemed more synchronous. This suggests that temporal order perception *negatively* rapidly recalibrated. Why then was this effect not detected in the TOJs? We propose that *negative* rapid recalibration of temporal order perception was obfuscated in TOJs by a *positive* response bias. Specifically, a bias to repeat TOJs in consecutive trials. We found clear evidence of this bias in terms of both temporal order and confidence judgements (see Figure 5) and found that the magnitude of this bias in TOJs was positively correlated with individuals' rapid recalibration effect. Even a small bias of this sort could produce trial-order effects in a TOJ task (see Figure 2) that are erroneously interpreted as rapid recalibration, but have no basis in timing perception. It is therefore perhaps surprising that we did not find a *positive* recalibration-like effect here due to the choice-repetition bias alone (as our simulations would indicate is extremely likely). We therefore reason that the null result reported here is due to the two opposing effects (i.e. *negative* rapid recalibration of temporal order perception and *positive* choice-repetition bias) nullifying one another. Of course, with two effects independently varying in magnitude and valence (e.g. choice-repetition or choice-alternation bias), any outcome is plausible.

In our second experiment, we aimed to dissociate participants' preceding TOJ from the preceding timing relationship. In doing so, we disrupted the relationship between the previous judgement and the previous SOA present in Experiment 1, and Roseboom's (2019) TOJ and temporal magnitude judgement tasks. Consistent with the results of Van der Burg et al. (2013, 2015), Simon et al. (2017) and our own confidence results here, we predicted that mitigating this choice-repetition bias effect would allow us to detect *negative* rapid recalibration of TOJs.

6 | EXPERIMENT 2—METHODS

We recruited 30 participants to complete an orientation judgement task interleaved with an audiovisual TOJ task. Participants gave written informed consent and were remunerated \$20/hour for their time. This study was approved by the ethics committee of the School of Human Movement and Nutrition Sciences at the University of Queensland (project number 2018002472). Participants reported having normal or corrected-to-normal hearing and vision and were naïve to the purposes of the experiment.

On every trial, participants were asked to fixate a red dot, subtending 0.25 DVA, in the centre of the screen. They were then presented with an oriented Gabor patch (12.5 DVA diameter) in a Gaussian aperture (standard deviation of 1.56 DVA) and an auditory tone (440 Hz). Stimuli were 10 ms in duration (corresponding to one frame at 100 Hz). The Gabor was oriented at one of six pre-defined orientations relative to vertical (±4, ±2 and ± 1 degrees). Presentation of the audio and visual stimuli was asynchronous, with stimuli presented at one of six SOAs (±300, ±200, and ± 100 ms). Visual stimuli were presented on a BENQ XL2720-B monitor at 100 Hz, using an NVIDIA GeForce GTX 1060 (6GB) graphics card. Auditory stimuli were presented via headphones, using the PsychPortAudio functions in MATLAB.

On the first trial, participants were asked to report which of the two stimuli had occurred first using the left and right mouse buttons. On the second trial, participants were asked to report whether the Gabor patch was oriented clockwise or counter-clockwise relative to vertical. These tasks alternated throughout the experimental blocks (i.e. participants never made two successive judgements of the same type). A text prompt appeared on-screen to remind participants of the response-mapping for each trial. Response mappings were constant for both judgements throughout the experiment and across participants. We took this approach to avoid making the task too difficult for participants and to make analysis of the bias itself simpler, though other studies have used random response mappings between trials (e.g. Pape et al., 2017; Zhang & Alais, 2019).

Each experimental block (241 trials each) took approximately 15 minutes to complete, and participants took a short break between blocks. We constructed a trial order for each block, and for each participant, which approximated orthogonal sampling of SOAs in TOJ trials relative to both orientation and SOA in orientation judgement trials. As a result, physical temporal order was approximately random with respect to physical orientation. With three blocks of 240 analysed trials (the first trial from each block was excluded from analyses), we yielded 60 TOJs for each SOA and 60 orientation judgements for each angle. When split by preceding-SOA, we yielded ~30 TOJs per SOA in each of the preceding SOA distributions.

Participants were given written and verbal task instructions and completed a demonstration version of the task. This demo version of the task was identical to the experimental version, except that it was shorter (25 trials), and participants were given explicit feedback after each trial. We included this demo version as we reasoned that interleaving two psychophysical tasks here would be even more difficult than the single task used in Experiment 1. Despite this demo block, 10 participants were removed from analyses as their responses indicated extremely poor performance, plausibly due to misunderstanding the task instructions or difficulty alternating between judgements.

Since the custom stimulus-presentation device used in Experiment 1 was limited to flashes and beeps, we used Psychtoolbox in MATLAB to conduct this experiment. One major concern using Psychtoolbox for timing experiments is the precision and accuracy of when stimuli are presented. On frames where the visual stimuli were drawn, we included a small white disc in the bottom left corner of the screen. When this frame was presented, the disc illuminated the lens of a photodiode attached to the corner of the computer monitor, providing a reliable timestamp for the onset of the visual stimulus. The audio output from the testing PC was split, with one connection going to headphones worn by the participant and one to an independent microcontroller running custom software. The output of the photodiode and PC audio were recorded by the microcontroller, which determined when the stimuli had onset, returning time stamps to the PC via USB. Stimulus timing was accurate to within ± 8 ms, and no trials were excluded on these grounds.

6.1 | Data analysis

The analytic method used here was identical to that used in Experiment 1. We split participants' TOJs into two data sets based on the temporal order of stimuli presented in the preceding trial. Of course, the preceding trial in this experiment was actually an orientation judgement, and participants were aware that they could ignore the auditory stimulus. We computed participants' PSS and JND for both data sets, and the difference in their PSS' to compute rapid recalibration. To quantify any response bias, we also completed this process with TOJs split by the preceding orientation judgement and preceding TOJ. Participants were excluded from analyses if any of their TOJ model parameters exceeded the sample mean by more than two standard deviations. This task was considerably more difficult for participants, given it required alternating between two challenging discriminations, unfortunately resulting in 10 of 30 participants being excluded from analyses. As in Experiment 1 and our preliminary analyses, we have used frequentist statistics with an arbitrary significance criterion of $\alpha = .05$, and estimates of effect size (Cohen's d). Our analysis code and raw data files have also been made available online (Rapid recalibration of temporal order judgements; https://doi.org/10.17605/OSF.IO/CNQ4V).

7 | EXPERIMENT 2—RESULTS

Before considering rapid recalibration or response biases, we first computed participants' sensitivity and accuracy to temporal order and orientation. Overall, participants exhibited a slight bias to perceive synchrony for audio-leads ($M_{PSS} = -52.37$ ms, $SD = 84.44$ ms), and similar JNDs to Experiment 1 ($M_{JND} = 143.85$ ms, $SD = 47.57$ ms). Group-average response distributions are presented in Figure 6 (below).

Paired t tests indicated that the temporal order of stimuli on the preceding trial systematically changed participants' perception of synchrony ($t_{19} = 2.42$, $p = .026$, $d = 0.23$), consistent with the effect found in SJ tasks (Simon et al., 2017; Van der Burg et al., 2013, 2015). The average PSS for trials following an audio-lead was −62.51 ms, and −43.52 ms for trials following a visual-lead. The magnitude of this effect (~19 ms) is

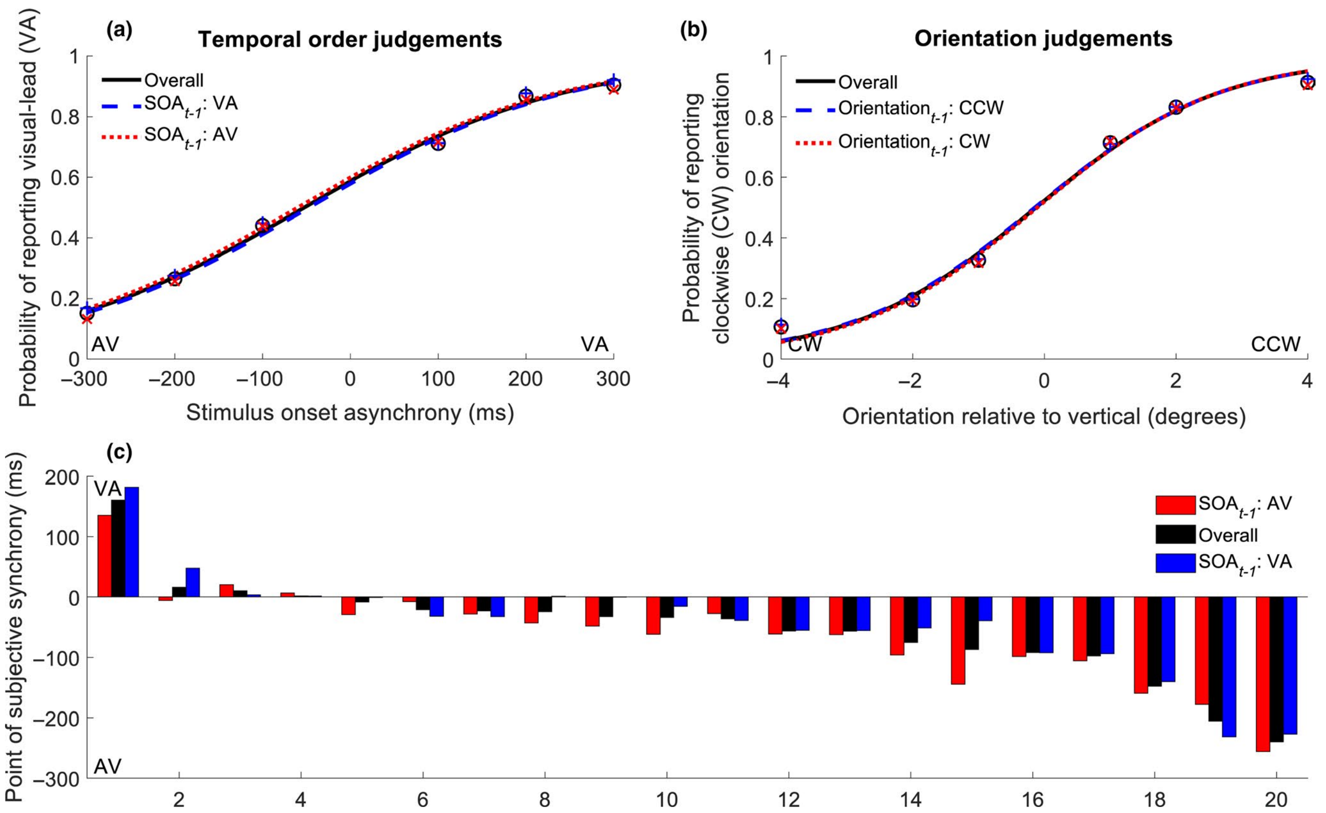

FIGURE 6 Temporal order and orientation judgements, collated across participants. Three data series are plotted: black, solid line, circle markers) overall response distributions, regardless of the timing relationship or orientation in the preceding trial; blue, dashed line, + markers) judgements from trials following a visuallead or counter-clockwise rotation; red, dotted line, x markers) judgements from trials following an audiolead or clockwise rotation. Note that panel C presents these data series in the order: red, black, blue; black ought to be between red and blue for all participants. (a) Participants' TOJs were generally sensitive to the timing of stimuli. Unlike in Experiment One, we found evidence that temporal order judgements negatively rapidly recalibrated. Note that the distributions provided here do not illustrate the within-participant differences we were testing, but provide the clearest depiction of the overall response distributions, and mimic the figure style of Van der Burg et al. (2013). (b) Participants' orientation judgements, both overall and as functions of the preceding orientation. We found no evidence of a trial-order effect on orientation perception, despite the well-replicated findings of Fischer and Whitney (2014). This is perhaps unsurprising given the substantial differences in the range of orientations presented; we used only an 8 degree range of orientations, whereas Fischer and Whitney used a 120 degree range in their first experiment. (c) Participants' PSS as a function of the temporal order of stimuli on the preceding trial, ordered by overall PSS. Most participants' PSS for trials following audio-leads was closer to the audio-lead SOAs than for trials following visual-leads. This negative rapid recalibration is consistent with our hypotheses, and results from synchrony judgement paradigms (e.g., Van der Burg et al., 2013) [Colour figure can be viewed at wileyonlinelibrary.com]

consistent with that found by Van der Burg et al. (2013) in their SJ task (~20 ms). However, Cohen's *d* indicates this is a small/weak effect; re-analysis of data collected by Roseboom (2019) reveals rapid recalibration of synchrony judgements on the scale of approximately 20 ms, but an effect size (*d*) of 0.62. As predicted, we found no evidence that stimulus order on the preceding trial changed participants' sensitivity to temporal order ($t_{19} = -0.37$, $p = .712$). While it did not reach our arbitrary significance criterion ($\alpha = 0.05$), participants' sensitivity to temporal order was correlated quite strongly with the magnitude of their rapid recalibration effect ($r_{18} = .40$, $p = .084$), consistent with predictions and the results of Van der Burg et al. (2013).

As a manipulation check, we then split participants' TOJs based on their preceding orientation judgement (*trial−1*) and TOJ (*trial−2*). Paired *t* tests revealed no evidence that participants' TOJs were biased by their most recent button-press (i.e. orientation judgement; $t_{19} = 1.18$, $p = .254$, $d = 0.21$), replicating the result of this analysis in Experiment 1, where we found no evidence that TOJs were biased by the preceding confidence judgement. Similarly, we found no evidence that participants' TOJs were biased by their most recent TOJ ($t_{19} = 0.87$, $p = .396$, $d = 0.16$). It is noteworthy that these effect size statistics (*d*) are only slightly less than that found for rapid recalibration in this paradigm, again suggesting that rapid recalibration is a subtle effect. Nevertheless, these results indicate our paradigm successfully disrupted the confounding correlation between the previous TOJ and the previous SOA in Experiment One and Roseboom (2019).

8 | GENERAL DISCUSSION

In this study, we aimed to resolve the key issue of whether *positive* rapid recalibration of temporal order perception reflects a choice-repetition bias, unrelated to timing perception. Previous studies found that participants' timing judgements are biased by the stimulus order in the preceding trial. In the SJ task, participants tend to perceive repeated stimulus order in consecutive trials as increasingly synchronous (e.g. Simon et al., 2017; Van der Burg et al., 2013). In the TOJ task (and magnitude judgement task, which requires temporal order be reported via an SOA estimate), the opposite result was reported (Roseboom, 2019). Further analysis of Roseboom's data (see Figure 1) indicated a bias to repeat consecutive TOJs could account for the reported effect. In essence, the previous SOA and previous TOJ were confounded.

In our first experiment, we aimed to determine whether timing perception had been rapidly recalibrated or whether trial-by-trial changes in TOJs were due to a choice-repetition bias. We found no evidence that TOJs positively *or* negatively rapidly recalibrated. However, we found evidence that participants' confidence *negatively* rapidly recalibrated. This suggests that temporal order perception had also *negatively* rapidly recalibrated, meaning similar SOAs in consecutive trials seemed more synchronous (or less like the repeated temporal order, and therefore judged with lower confidence). We proposed that this effect would have been detectable in TOJs had it not been obfuscated by a *positive* choice-repetition bias.

Experiment 2 tested this prediction by dissociating the previous TOJ from the previous SOA. In this case, participants' most recent TOJ were random with respect to the most recently presented SOA, statistically eliminating the effect of the choice-repetition bias. Under these conditions, we found evidence of *negative* rapid recalibration of TOJs, consistent with the effect reported in SJ tasks. That is, participants perceived similar SOAs in consecutive trials as more synchronous (i.e. less like the repeated temporal order).

Interestingly, participants' overall PSS in Experiment 2 was biased towards audio-leads, relative to participants' PSS in Experiment 1 and in other TOJ studies (e.g. Vroomen et al., 2004). We speculate that this discrepancy may be related to the need for fine visual discriminations in Experiment 2 which are classically absent from other TOJ tasks. This attentional bias has been found to induce an effect called prior entry, whereby attended (in this case, visual) stimuli are typically perceived as having occurred earlier (Vibell, Klinge, Zampini, Spence, & Nobre, 2007; Yates, & Nicholls, 2011). Fortunately, this effect is likely to be uniform with respect to the preceding timing relationship and judgement, so does not limit the inferences we can draw here.

Our finding of *negative* rapid recalibration potentially undermines some of the conclusions drawn by Roseboom (2019) to explain the contradictory effects in TOJ and SJ tasks. Based on his findings, he argued that rapid recalibration likely reflects changes in decision-making processes, and these changes manifest differently for synchrony and temporal order/magnitude judgements. This was a reasonable interpretation for the set of contradictory results and may still be true. However, our findings suggest that more traditional explanations, which assume that rapid recalibration occurs in the sensory domain, remain plausible.

If we assume that timing information is represented by some sensory process that then feeds forward into separate decision-making processes, the most parsimonious explanation for a common effect in all judgements is that the effect is on sensory processing itself. For instance, single-trial adaptation of a population of delay-tuned neurons, like that proposed in Roach, Heron, Whitaker, and McGraw (2010), could account for this effect. Of course, this is at odds with the results of Simon et al. (2017) who investigated rapid recalibration of SJs using EEG. They found that rapid recalibration was indexed by late-evoked and therefore presumably post-sensory, processes. The alternative, then, is that a common rapid recalibration effect across timing judgements reflects a common decisional strategy.

For instance, Roseboom (2019) proposed that an assimilative effect on decision criteria could produce rapid recalibration as observed in the SJ task, in line with a proposal by Yarrow, Jahn, Durant, and Arnold (2011). Roseboom then pointed out how this would not account for his *positive* TOJ rapid recalibration effects, as such a model can produce only *negative* rapid recalibration effects. However, since we found *negative* rapid recalibration of TOJs, we should reconsider this assimilative model. A shift of decision criteria towards the SOA of recent stimulus pairs could account for the similar effects of rapid recalibration on temporal order and synchrony judgements. This is also consistent with the observed changes in participants' confidence reports.

To clarify the point, imagine a participant perceives a flash-lead SOA on some trial, and their decision criterion then shifts slightly towards flash-lead SOAs. If the subsequent trial is a flash-lead SOA, it is less likely to be perceived as such, being closer to the criterion, accounting for the effect observed here. If we assume that confidence is in some way related to the difference in participants' perceived SOA and the decision criterion, the second trial would also be associated with lower confidence, as observed in Experiment 1. This simple decisional model, relying on assimilation of decision criteria, provides neat and supported predictions regarding trial-by-trial changes in temporal order and confidence judgements. However, it is not sufficient to explain other recent findings (e.g. Roseboom, Linares, & Nishida, 2015), making it unlikely that changes in decisional processing solely drive rapid recalibration (or temporal recalibration following prolonged adaptation).

Roseboom et al. (2015) found that prolonged exposure to temporal asynchrony induces changes not only in synchrony perception, but the sensitivity of the observer to timing information itself. For instance, when participants are repeatedly exposed to a particular timing relationship, their sensitivity to other timing relationships is relatively increased. This effect is at odds with simple, purely decisional, accounts of temporal recalibration. It remains to be seen whether such effects (non-uniform distortion of sensitivity around adapted SOAs) are observable on the scale of a single trial. If so, that would provide strong evidence that rapid recalibration occurs at the level of sensory processing.

Future research might make substantial progress by leveraging an understanding of rapid recalibration from the already well-established literature on the temporal window of multisensory integration. For instance, it is very well established that if cross-modal stimuli are presented within a given window of time, they will be integrated and be responded to more rapidly (e.g. Diederich & Colonius, 2004; Frens, Van Opstal, & Van der Willigen, 1995; Nozawa, Reuter-Lorenz, & Hughes, 1994). We found that repeated temporal orders on consecutive trials were perceived as more synchronous, and so may therefore be more likely to fall within the temporal window of integration and be responded to more rapidly. Using a reaction-time paradigm might substantiate the marked asymmetry in rapid recalibration reported by Van der Burg et al. (2013), and even exaggerate rapid recalibration. Since the window of integration tends to be larger in reaction-time tasks than temporal order judgement tasks (Diederich & Colonius, 2015; Mégevand, Molholm, Nayak, & Foxe, 2013), and rapid recalibration is greater for those with poorer sensitivity to synchrony (i.e. broader synchrony distributions, and therefore plausibly a broader window of integration; Van der Burg et al., 2013), rapid recalibration ought to be greater in a reaction-time task. Future research might also benefit from our existing knowledge of the brain regions associated with the temporal window of multisensory integration, for instance by using brain stimulation to modulate the breadth of the window (e.g. Cecere, Rees, & Romei, 2015; Zmigrod & Zmigrod, 2015).

One of the broader implications of our work is that a simple behavioural response (like a button-press) reflects a judgement that is contaminated by biases of perception and decision-making, and others have shown that the action itself is subject to its own independent biases (e.g. Diedrichsen, White, Newman, & Lally, 2010; Pape et al., 2017). While this presents methodological challenges for some future work, especially for investigations of serial dependencies in perception, it also presents an opportunity to gain novel insights into the neural computations underlying participants' responses. For instance, Gekas et al. (2019) found that sensory history influences subsequent perception variably at different timescales, and independently of response bias effects. Similarly, Pape et al. (2017) were able to dissociate serial response biases from choice-related biases, and found they trend in opposite directions (i.e. participants were more likely to alternate responses, but repeat their task-related judgements). Carefully controlled neuro-imaging and modelling studies might allow for identification of the neural structures and processes contributing to each of these independent biases, and the integration of these independent biases prior to motor output (e.g. Urai, De Gee, Tsetsos, & Donner, 2019).

9 | CONCLUSION

This study aimed to determine whether rapid recalibration of temporal order perception reflects a generalised bias in decision-making, unrelated to timing perception. Analyses of our own data and data provided by Roseboom (2019) indicated that a simple bias to repeat judgements in consecutive trials could account for the contradictory effects reported in temporal order and synchrony judgement tasks. In our first experiment, we measured participants' confidence in their order judgements and found evidence of *negative* rapid recalibration. In a final experiment that employed a novel design with interleaved tasks, we eliminated participants' bias to repeat their judgements and found evidence that temporal order judgements *negatively* rapidly recalibrate, consistent with synchrony judgements. It remains unclear how rapid recalibration manifests, and at what level of processing; this effect could be accounted for by simple changes in either sensory or decision-making processes. Further research is required to determine the locus of these effects, although attempting to replicate recent findings by Roseboom et al. (2015) in single-trial adaptation would appear to be a good place to start.

ACKNOWLEDGEMENTS

This work was funded by an Australian Research Council Discovery Grant held by Timothy J Carroll and an Australian Government Research Training Program stipend held by Brendan Keane. Our thanks to Reuel Chua, Michael Camargo and Erica Hungerford, who assisted with some data collection.

CONFLICT OF INTERESTS

The authors declare no competing or conflicting interests.

DATA AVAILABILITY STATEMENT

Data collected in experiments 1 and 2 are available at the following https://doi.org/10.17605/osf.io/cnq4v, hosted by the Open Science Framework. The repository also holds code used to analyse the data collected in experiments 1 and 2 and used to re-analyse published data (Roseboom, 2019; Audiovisual synchrony, temporal order and, magnitude judgements; https://doi.org/10.1037/xhp0000591.supp).

AUTHOR CONTRIBUTIONS

BK designed and programmed experiments 1 and 2, collected data for experiments 1 and 2, analysed data for experiments

1 and 2, re-analysed published data and contributed to interpreting these results and writing the manuscript. NSB, NSB, TJC and GW contributed to the interpretation of results from experiments 1 and 2, re-analyses of published data (including simulations conducted by BK) and drafting the manuscript.

ORCID

Brendan Keane https://orcid.org/0000-0002-3761-9689
Nicholas S. Bland https://orcid.org/0000-0001-9610-9972
Natasha Matthews https://orcid.org/0000-0001-5683-7289
Timothy J. Carroll https://orcid.org/0000-0003-0761-1819
Guy Wallis https://orcid.org/0000-0003-1601-1540

REFERENCES

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology*, *25*(2), 231–235.

Diederich, A, & Colonius, H (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*(8), 1388–1404.

Diederich, A, & Colonius, H (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. *Psychological Review*, *122*(2), 232–241.

Diedrichsen, J, White, O, Newman, D, & Lally, N (2010). Use-dependent and error-based learning of motor behaviors. *Journal of Neuroscience*, *30*(15), 5159–5166.

Donohue, SE, Woldorff, MG, & Mitroff, SR (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics*, *72*(4), 1120–1129.

Fischer, J, & Whitney, D (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738.

Frens, MA, Van Opstal, AJ, & Van der Willigen, RF (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics*, *57*(6), 802–816.

Fujisaki, W, Shimojo, S, Kashino, M, & Nishida, SY (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*(7), 773.

Gekas, N, McDermott, KC, & Mamassian, P (2019). Disambiguating serial effects of multiple timescales. *Journal of Vision*, *19*(6), 24–24.

Marinovic, W, Poh, E, de Rugy, A, & Carroll, TJ (2017). Action history influences subsequent movement via two distinct processes. *Elife*, *6*, e26713.

Mégevand, P, Molholm, S, Nayak, A, & Foxe, JJ (2013). Recalibration of the multisensory temporal window of integration results from changing task demands. *PLoS ONE*, *8*(8), e71608.

Nozawa, G, Reuter-Lorenz, PA, & Hughes, HC (1994). Parallel and serial processes in the human oculomotor system: bimodal integration and express saccades. *Biological Cybernetics*, *72*(1), 19–34.

Pape, A. A., Noury, N., & Siegel, M. (2017). Motor actions influence subsequent sensorimotor decisions. *Scientific reports*, *7*(1), 15913.

Roach, NW, Heron, J, Whitaker, D, & McGraw, PV (2010). Asynchrony adaptation reveals neural population code for audio-visual timing. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1710), 1314–1322.

Roseboom, W (2019). Serial dependence in timing perception. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), 100.

Roseboom, W, Linares, D, & Nishida, SY (2015). Sensory adaptation for timing perception. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1805), 20142833.

Simon, DM, Noel, JP, & Wallace, MT (2017). Event related potentials index rapid recalibration to audiovisual temporal asynchrony. *Frontiers in Integrative Neuroscience*, *11*, 8.

Urai, AE, De Gee, JW, Tsetsos, K, & Donner, TH (2019). Choice history biases subsequent evidence accumulation. *ELife*, *8*, e46331.

Van der Burg, E, Alais, D, & Cass, J (2013). Rapid recalibration to audiovisual asynchrony. *Journal of Neuroscience*, *33*(37), 14633–14637.

Van der Burg, E, Orchard-Mills, E, & Alais, D (2015). Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental Brain Research*, *233*(1), 53–59.

Verstynen, T, & Sabes, PN (2011). How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *Journal of Neuroscience*, *31*(27), 10050–10059.

Vibell, J., Klinge, C., Zampini, M., Spence, C., & Nobre, A. C. (2007). Temporal order is coded temporally in the brain: early event-related potential latency shifts underlying prior entry in a cross-modal temporal order judgment task. *Journal of cognitive neuroscience*, *19*(1), 109–120.

Vroomen, J, Keetels, M, De Gelder, B, & Bertelson, P (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*(1), 32–35.

Yarrow, K, Jahn, N, Durant, S, & Arnold, DH (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, *20*(4), 1518–1531.

Yates, M. J., & Nicholls, M. E. (2011). Somatosensory prior entry assessed with temporal order judgments and simultaneity judgments. *Attention, Perception, & Psychophysics*, *73*(5), 1586–1603.

Zhang, H, & Alais, D (2019). Individual difference in serial dependence results from opposite influences of perceptual choices and motor responses. *BioRxiv*. *631309*, 1–29.

Zmigrod, S., & Zmigrod, L. (2015). Zapping the gap: Reducing the multisensory temporal binding window by means of transcranial direct current stimulation (tDCS). *Consciousness and cognition*, *35*, 143–149.

How to cite this article: Keane B, Bland NS, Matthews N, Carroll TJ, Wallis G. Rapid recalibration of temporal order judgements: Response bias accounts for contradictory results. *Eur J Neurosci*. 2020;51: 1697–1710. https://doi.org/10.1111/ejn.14551