Action, attention, and temporal binding

Liyu Cao

Department of Psychology and Behavioural Sciences, Zhejiang University, 310058, Hangzhou, China

Correspondence to: Dr. Liyu Cao, Department of Psychology and Behavioural Sciences, Zhejiang University, Yuhangtang Road 866, 310058, Hangzhou, China

Keywords: outcome binding; temporal binding; attention; action; sense of agency

Abstract

Outcome binding has long been understood as an illusion of timing perception, in which an action-effect is perceived as occurring earlier than it actually does. This illusion has been heavily investigated over the past two decades with regards to the mechanisms behind and potential applications. Here we present evidence in favour of understanding outcome binding as a spatial attentional effect, at least in part. In a series of 3 experiments, it was shown that an action-effect was preceded by a predictive attention shift in the classic Libet clock paradigm. The magnitude of attention shift predicted the size of outcome binding. When the attention shift was controlled for, binding also disappeared. Our study also calls for a reassessment of results obtained from the clock-like method in mental chronometry dating back to Wundt, as attention may well be a critical confounding factor in the interpretation of the results from these studies.

Introduction

Our own actions influence perception profoundly (Rolfs & Schweitzer, 2022). For example, we can barely tickle ourselves, whereas the same tactile stimulation produced by someone else or by an object can be pretty ticklish (Blakemore et al., 1998; Weiskrantz et al., 1971). In the domain of temporal perception, Haggard and colleagues reported an illusion of temporal attraction between an action and a slightly delayed sensory event (i.e. action binding: the action being perceived as occurring later, and outcome binding: the sensory event being perceived as occurring earlier; together known as temporal binding or intentional binding) (Haggard et al., 2002; Wolpe et al., 2013). Over the two past decades, the temporal binding effect has attracted cross-disciplinary attention with regards to its cognitive/neural mechanisms and the potential applications especially with its widespread use as an implicit measure of sense of agency (Antusch et al., 2021; Buehner & Humphreys, 2009; Dogge et al., 2012; Haggard, 2017; Kirsch et al., 2019; Legaspi & Toyoizumi, 2019; Moore & Obhi, 2012). Action also has a tight link to attention. Specifically, action modulates the distribution of attention according to the action goal (Deubel & Schneider, 1996; Rolfs et al., 2011; Tipper et al., 1998). For example, when reaching towards a target, visual attention was shown to be drawn towards the target before the onset of the reaching action (Baldauf & Deubel, 2010; Deubel et al., 1998; Eimer et al., 2006; Rolfs et al., 2013). The action-related attention modulation may have far reaching implications for the understanding of action-related perceptual changes such as temporal binding (see also Haggard and Cole (2007)).

We hypothesise that action-related attention modulation contributes to the measure of, or even cause temporal binding, at least in tasks where there is an identifiable visuospatial component. Temporal binding is usually measured using the Libet clock method (Libet et al., 1983). During the testing, participants watched a clock face with a rapid rotating clock hand and indicated the time of a sound onset by reporting where the clock hand was positioned at sound onset (e.g. the clock hand was at 12 o'clock position when the sound was played). In the instrumental condition (Action Sound condition, AS), the sound was triggered by an action with a short delay (e.g. the sound was played after a keypress with a 250 ms delay). In the baseline condition (Sound Only condition, SO), the sound was controlled by the computer. A robust finding is that the reported clock hand position is earlier in time in the AS condition than in the SO condition, known as outcome binding. This effect is the focus of the current study. The other component of temporal binding (i.e. action binding) is much smaller in size than outcome binding and will be further treated in the discussion. Suppose that the clock hand is positioned at 12 o'clock when the sound is played on one particular trial (Figure 1a). In the AS condition, the sound-triggering action is made when the clock hand is at around the 10 o'clock position, if the action-effect interval is 250 ms and the clock hand rotates with 1800 ms per revolution as done in the current study. Because of the action-related attention modulation, attention may have been activated before the onset of the sound. Different from a target reaching task, in which the hand has to move from a starting position to the location of a target, the keypress in the Libet clock method can be successfully performed rather easily without considering the information of a target. However, the time reporting task requires the identification of the clock hand position around the time of the keypress. Conceivably, the task goal (i.e. reporting the clock hand position) may be associated with an attention shift towards the clock rim similar to the attention shift to the target in a target reaching task. This attention shift should occur before the sound onset in the AS condition due to the preceding keypress, which is predictive of the imminent position reporting task. In the SO condition, since the onset time of the sound is not known in advance, attention may accumulate towards the clock rim only after the sound is played. Therefore, a difference in the spatiotemporal distribution of attention is predicted between the AS and SO conditions, with the attention in the AS condition leading the attention in the SO condition. Since the clock hand position

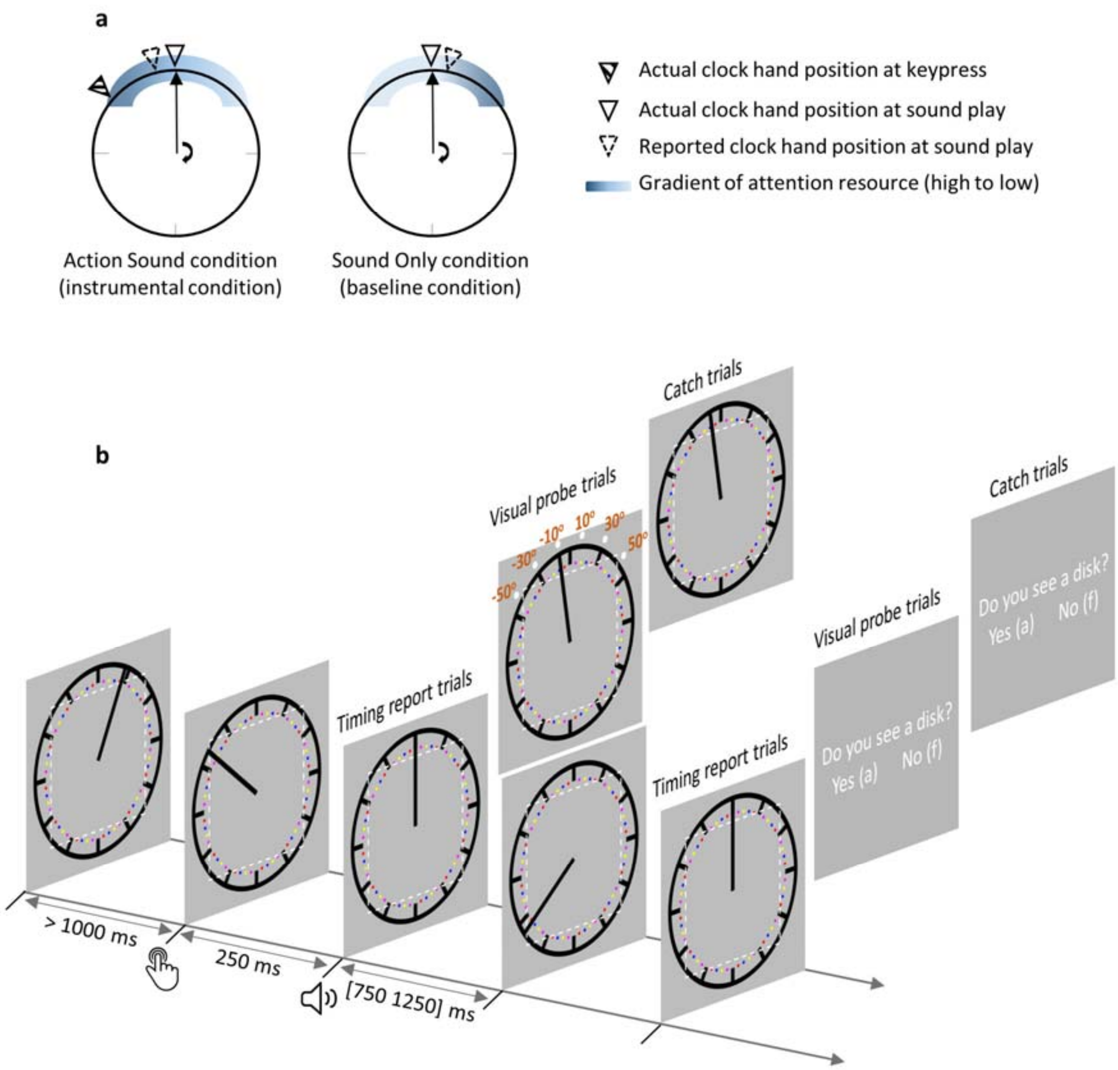

Figure 1 Experiment idea and design. (a) The hypothesis. In the AS condition, action activates attention before the onset of the delayed action effect. In the SO condition, attention increases only after the onset of the sound. (b) Trial structure in the AS condition. Each trial started with the clock hand rotating from a random angle. After at least 1000 ms, a voluntary keypress triggered a 250 ms delayed sound. The clock hand continued rotating for another random period between 750 and 1250 ms after the sound play. In timing report trials, participants should move the clock hand back to its position at sound play. In visual probe trials and catch trials, participants reported if a visual probe was detected. A visual probe was presented in each visual probe trial at one of 6 possible locations relative to the clock hand

position at the time of keypress. In the example here, the keypress was made when the clock hand pointed to the -50^o position. The 6 possible probe locations were: -50^o (0 ms delay from the keypress), -30^o (100 ms delay), -10^o (200 ms delay), 10^o (300 ms delay), 30^o (400 ms delay), and 50^o (500 ms delay). The imaginary dotted white square (not shown during the testing) illustrates the eye movement control area in Experiment 2 (eyes moving out of this area would lead to a trial abortion). In the SO condition, everything was the same expect that no keypress was required. In the VS condition, the keypress was replaced by a passive vibrotactile stimulation.

reporting task is quite demanding due to its fast-rotating speed, it is conceivable that the reported clock hand position can receive strong influence from attention. That is, the more attended spatial location will be more likely reported as the clock hand position. Strikingly, the predicted spatiotemporal difference in attention corresponds very well to the temporal binding effect (i.e. the reported clock hand position in the AS condition leads the reported clock hand position in the SO condition).

In the current study, we tested the attention hypothesis of temporal binding. In a series of 3 experiments, we showed the action led to an attention activation before the onset of the sensory outcome, which corresponded very well to the temporal binding effect (Experiment 1). When the eye movements were controlled, the magnitude of attention activation predicted the size of temporal binding (Experiment 2). Finally, when the attention activation was controlled for, the temporal binding effect disappeared (Experiment 3). Our results indicated that attention is fundamentally involved in temporal binding.

Results

Experiment 1

Participants reported the onset time of a sound using the Libet clock method (Figure 1b). During the experiment, participants were asked to watch the centre of the visual presentation, but the eye movements were not controlled. A clear outcome binding effect was confirmed ($t(12) = -8.84$, $p < 0.001$, $dz = -2.45$; Figure 2a,j). The reported time was earlier when the sound was triggered by participants through a keypress (AS condition; $M = -76.92$ ms, *95% CI* = [-103.67 -50.18] ms) than when the sound presentation was controlled by computer (SO condition; $M = 35.00$ ms, *95% CI* = [7.99 62.01] ms). Crucially, the pattern of attention distribution was drastically different between the two conditions (Figure 2b). In the AS condition, the detection performance in the visual probe detection task was high at the time of keypress, gradually increased, and peaked just before the sound presentation. After the sound presentation, the detection performance underwent a sharp decrease. In the SO condition, the detection performance was low before the sound presentation, but high after the sound presentation. The pattern difference between the two conditions was confirmed by a significant interaction effect in a two-way (condition: AS vs. SO; probe location: -50°, -30°, -10°, 10°, 30°, and 50°) within-participants ANOVA comparing the detection performance ($F(5,60) = 8.94$, $p < 0.001$, $\eta_p^2 = 0.43$). The ANOVA also revealed a significant main effect of probe location ($F(5,60) = 6.97$, $p = 0.001$, $\eta_p^2 = 0.37$). The main effect of condition was not significant ($F(1,12) = 3.30$, $p = 0.094$, $\eta_p^2 = 0.22$).

We next tested if the difference in visual attention distribution between the two conditions would correlate with the outcome binding effect. An attention difference index was calculated for each participant. To calculate the attention difference index, an attention shift score was first calculated separately for each condition through subtracting the detection performance before the sound presentation from the detection performance after the sound presentation. The attention difference index was obtained by subtracting the attention shift score in the AS condition from the SO condition. Accordingly, a large attention difference

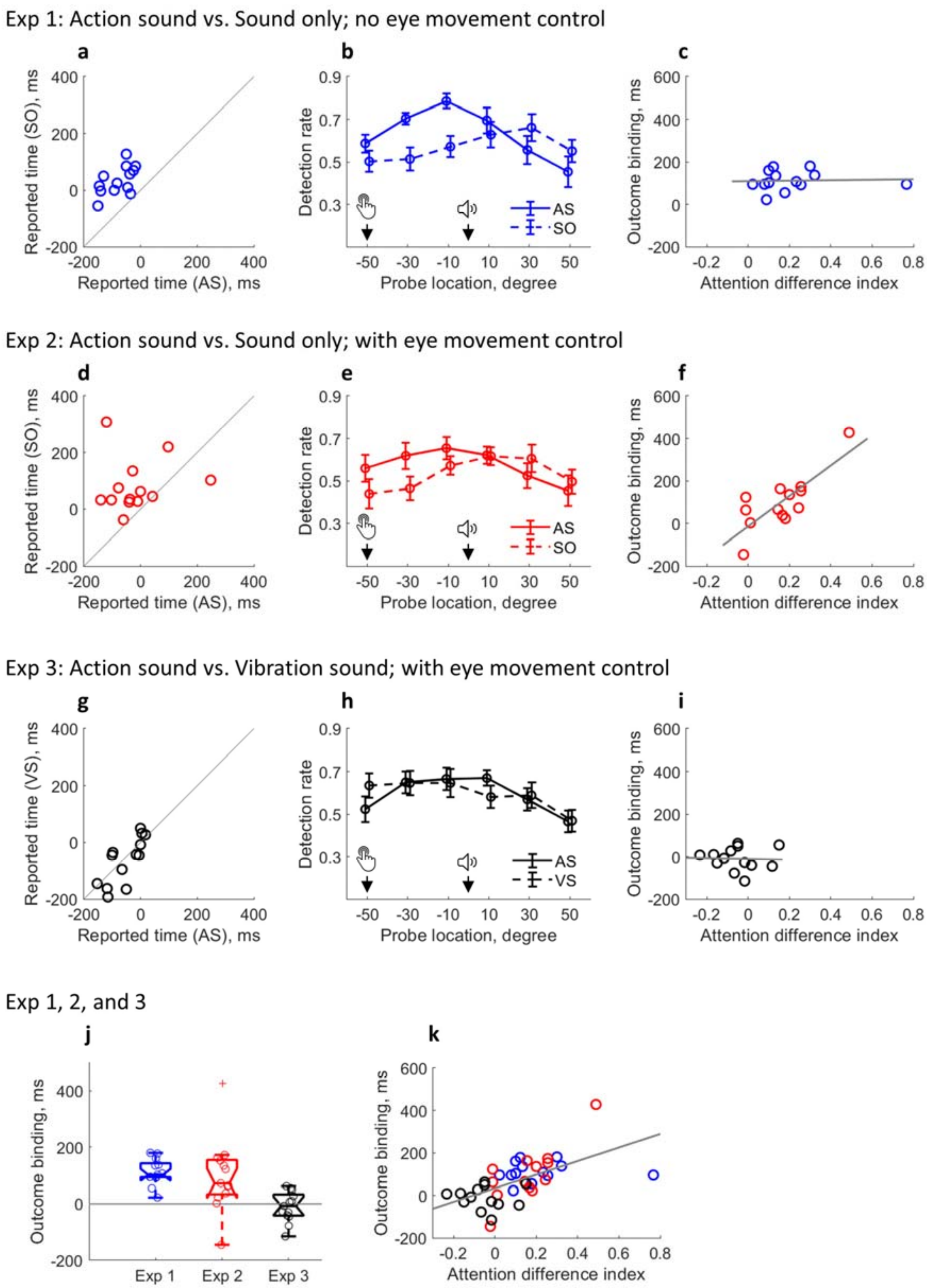

Figure 2 Results from 3 experiments. (a) Outcome binding in Experiment 1 (no eye movement control; each circle represents a participant). (b) Detection rate of visual probes as

a function of condition and probe location in Experiment 1. Attention was activated by the keypress in AS condition and by the sound in the SO condition (note that there was no keypress in the SO condition). Bars represent ± 1 standard error. (c) Scatter plot showing the relationship between attention difference and outcome binding in Experiment 1. No significant correlation was found. (d-f), similar to (a-c), but with the results from Experiment 2. When eye movements were controlled, the pattern of attention modulation by action was replicated (e), and the magnitude of attention difference predicted the size of outcome binding (f). (g-i), similar to (a-c), but with the results from Experiment 3. In Experiment 3, a VS condition replaced the SO condition in Experiment 1 and 2. No significant difference in the reported time or the attention distribution was found between AS and VS conditions, demonstrating the causal role of attention in outcome binding. (j) Box-plot of outcome binding effects from 3 experiments. Each circle represents a participant. Outliers are represented by a plus sign. (k) Pooling all the data from 3 experiments, a significant correlation between the attention difference index and the size of outcome binding was found. AS: action sound condition. SO: sound only condition. VS: vibration sound condition.

index corresponds to a state of high attention resource before the sound presentation in the AS condition and high attention resource after the sound presentation in the SO condition. The outcome binding effect for each participant was the difference in the reported time between the two conditions (SO - AS). Across participants, no significant correlation was found between attention difference and outcome binding ($r(11) = 0.23$, $p = 0.228$; Spearman rank correlation; Figure 2c). No outliers were identified in the correlation analysis. Eye movements might be a critical factor modulating attention deployment in the task. Although participants were asked to fixate the centre of the visual display, no measurements to taken to implement this requirement. In the next experiment, fixation on the central area of the visual display was ensured with eye tracking.

Experiment 2

Experiment 2 was a replication of Experiment 1 with a new group of 15 participants. The only change in Experiment 2 was that the eye movements were strictly controlled using an eye tracking device. Throughout a trial, participants should always fixate the central area of the visual presentation. Trials with failed fixation were repeated until success. Again, a clear outcome binding effect was confirmed ($t(12) = -2.75$, $p = 0.009$, $dz = -0.76$; Figure 2d,j). The reported time was earlier in the AS condition ($M = -17.50$ ms, *95% CI* = [-73.33 38.33] ms) than in the SO condition ($M = 81.73$ ms, *95% CI* = [31.48 131.98] ms). The pattern of visual detection performance was almost identical to that found in Experiment 1. Attention was high at the time of keypress, gradually increased to its peak just before the sound presentation, and started to decrease after the sound presentation in the AS condition. Attention was low before the sound presentation, but high after the sound presentation in the SO condition (Figure 2e). This demonstrates the robustness of the attention modulation induced by action in the AS condition (and by the sound in the SO condition). The two-way (condition: AS vs. SO; probe location: -50°, -30°, -10°, 10°, 30°, and 50°) within-participants ANOVA comparing the detection performance confirmed a significant interaction effect ($F(5,60) = 5.29$, $p = 0.002$, $\eta_p^2 = 0.31$). The ANOVA also revealed a significant main effect of probe location ($F(5,60) = 6.06$, $p = 0.002$, $\eta_p^2 = 0.34$) and a non-significant main effect of condition ($F(1,12) = 1.03$, $p = 0.331$, $\eta_p^2 = 0.08$).

When the eye movements were strictly controlled, strikingly, a significant correlation between the attention difference index and outcome binding was found ($r(11) = 0.69$, $p = 0.006$; Spearman rank correlation). The participant with an outcome binding effect over 400 ms (Figure 2f) was identified as a statistical outlier in the correlation analysis using the box-plot rule, although this participant failed to be included as an outlier in the data prepossessing stage when looking at the reported time alone. The correlation was still significant after excluding this participant ($r(10) = 0.60$, $p = 0.021$; Spearman rank correlation). The positive correlation was in the direction in favour of our prediction, demonstrating that a large difference in attention distribution between AS and SO was associated with a strong outcome binding effect.

Experiment 3

Experiment 1 and 2 demonstrated the attention effect in outcome binding measure and the correlation between the attention effect and outcome binding. Experiment 3 sought to find evidence that attention is the driving force of outcoming binding (i.e. causal evidence). To this end, the SO condition was replayed by a VS condition, in which a vibrotactile stimulation was applied to the keypressing finger 250 ms before the sound play. The vibrotactile stimulation had the same timing as the keypress in the AS condition. It should have a similar effect on attention to the keypress in the AS condition, as it is a signal predictive of the sound onset and the time reporting task. If this is the case, outcome binding should vanish when comparing VS and AS conditions.

Another 15 participants were recruited for Experiment 3. Experiment 3 was similar to Experiment 2, but a VS condition replaced the SO condition. Indeed, no outcome binding effect was found between VS and AS conditions ($t(12) = 0.68$, $p = 0.509$, $dz = 0.19$; AS condition: $M = -53.46$ ms, *95% CI* = [-84.82 -22.10] ms; VS condition: $M = -63.46$ ms, *95% CI* = [-107.80 -19.12] ms; Figure 2g,j). The size of outcome binding in Experiment 3 was smaller than in Experiment 1 ($t(24) = -6.29$, $p < 0.001$, $ds = -2.47$) and Experiment 2 ($t(24) = -2.80$, $p = 0.005$, $ds = -1.10$). The pattern to attention distribution also appeared similar between VS and AS conditions. In both conditions, attention was high before the sound play, but low after (Figure 2h). The two-way (condition: AS vs. VS; probe location: -50°, -30°, -10°, 10°, 30°, and 50°) within-participants ANOVA comparing the detection performance only revealed a significant main effect of probe location ($F(5,60) = 9.04$, $p < 0.001$, η_p^2 =

0.43). The interaction effect ($F(5,60) = 2.16$, $p = 0.102$, $\eta_p^2 = 0.15$) and the main effect of condition ($F(1,12) = 0.01$, $p = 0.922$, $\eta_p^2 = 0.001$) were not significant. Therefore, when the attention difference between conditions disappeared, outcome binding also disappeared, demonstrating the causal role of attention in outcome binding.

The correlation between the attention difference index and outcome binding was not significant in Experiment 3 ($r(11) = -0.15$, $p = 0.692$; Spearman rank correlation; Figure 2i). No outliers were identified. One likely reason for the lack of a significant correlation as found in Experiment 2 was the reduced variability in the attention difference index due to the effort of trying to match the attention distribution between conditions in Experiment 3. However, a significant correlation between the attention difference index and outcome binding was again found when the data from 3 experiments were combined ($r(37) = 0.67$, $p < 0.001$; Spearman rank correlation; Figure 2k). The participant with an outcome binding effect over 400 ms was again identified as a statistical outlier using the box-plot rule. However, the correlation was still significant after excluding this participant ($r(36) = 0.64$, $p < 0.001$; Spearman rank correlation).

Discussion

The current study investigated the attention distribution in the temporal binding measurement using the classic Libet clock method. In 3 experiments, it was demonstrated that action induced a predictive attention activation before the onset of the delayed sensory outcome. Reactive attention activation was observed when the sensory stimulus was presented alone (i.e. without a preceding action). The attention distribution pattern additionally predicted the size of the temporal binding effect. When the predictive attention activation induced by action was controlled for, temporal binding also disappeared.

The results pose a strong challenge to the long-held view that temporal binding reflects an illusion in timing perception (Haggard et al., 2002; Tanaka et al., 2019; Wen & Imamizu, 2022). Classically, these studies use a chronometric display, such as a clock, to measure the time at which the event occurs. We show that the difference between conditions in event perception may be confounded by a difference between conditions in the location of spatial attention on the clock display. Because reporting the event time using the clock hand position is a quite demanding task (clock hand rotates rapidly), the location received more spatial attention conceivably will also be given higher weights when reporting the clock hand position retrospectively. Thus, the difference in attention between conditions can lead to a difference in reported clock times. Our results indeed showed a strong correlation between these two measures. More importantly, the difference in reported clock times disappeared as a consequence of experimentally matching the attention distribution between conditions. Previous research also showed that temporal binding declines when the temporal predictability of action and effect is controlled for (Kirsch et al., 2019). Conceivably, such control allows predictive attention deployment to operate in the same manner in different conditions. Therefore, phenomena such as temporal binding effects may in fact be an indirect consequence of an interaction between spatial attention and the chronometric methods used to measure time perception, rather than a genuine difference in time perception.

However, extra care should be taken before concluding that timing perception is not involved in temporal binding at all. First, action binding has not been tested under the attention explanation. Action binding is calculated as the difference in the reported action time between the AS condition and an action only condition (i.e. actions without the sensory outcome as presented in the AS condition). One idea is that the sensory outcome in the AS condition presents itself as an additional source of attention attraction. Therefore, compared to the action only condition without a sensory outcome, the spatial distribution of attention after the sensory outcome is stronger in the AS condition, which corresponds to the delay of the reported time in the AS condition (i.e. action binding). Second, the temporal binding effect has also been demonstrated using paradigms in which visual attention does not seem to critically involved, including but not limited to the interval estimation paradigm (Engbert et al., 2007; Humphreys & Buehner, 2009) and the auditory timer paradigm (Martinez et al., 2018; Muth et al., 2021). Whether attention could explain the temporal binding effect reported there awaits further investigations.

The present study also has important implications to the clock method in mental chronometry. Wundt probably was the first to use the clock with a fast rotating hand for the scientific study of mental chronometry (Wundt, 1874). The method is now also known as the Libet clock method, as Libet popularised this method in his famous study on free will (Libet et al., 1983). What is clear is that the time reporting results from this method are quite variable among individuals and strongly depend on the method details such as the speed of the clock hand (Ivanof et al., 2022; Miller et al., 2010; Pockett & Miller, 2007; Sanford, 1974; Seifried et al., 2010; Wundt, 1874; Yabe & Goodale, 2015). However, the results are quite often used to draw conclusions about the temporal properties of mental processing. In fact, all these results have always involved attention to spatial location on a clock-like device in order to report a perceived time. The current study suggests that any difference between conditions in perceived time from this method may be in fact due to differences in spatial attention. We should therefore investigate where such differences in spatial attention might arise in order to better understand the results produced by this method.

To sum up, the current study provided novel and important insights into the understanding of temporal binding. It showed that the temporal binding effect measured with the Libet clock method may be better understood as an attention effect. It also illustrates the importance of disentangling attention and perception in action-effect research, although the two are quite often intricately linked between themselves and to action.

Material and Methods

Experiment 1

Participants

15 participants (9 females; mean age = 21.3, *SD* = 1.9) were recruited from a local participant pool. All participants have normal or corrected-to-normal vision. Written informed consent was obtained prior to experiment, and participants were debriefed and received monetary payment after the experiment. The experiment was conducted in accordance with the Declaration of Helsinki (2013) and was approved by the Ethics Committee of Department of Psychology and Behavioural Sciences, Zhejiang University (ethics application number: [2022]003).

Stimuli, Task and Procedure

The experiment consisted of three parts: threshold testing, AS condition, and SO condition, in this order. The average testing duration was about 110 minutes.

In the threshold testing session, the luminance threshold of the visual probe used in AS and SO conditions was obtained using a 2-down-1-up staircase procedure (Levitt, 1971). Experiment stimuli were presented on a grey background (RGB value: [128 128 128]; used throughout the experiment). The staircase procedure for obtaining the luminance threshold was run in two parallel lines, with one line having the luminance intensity starting at 128 (RGB value: [128 128 128]; intensity increase line) and the other starting at 200 (RGB value: [200 200 200]; intensity decrease line). Each trial picked a random line until 15 reversals was obtained for each line. In each trial, participants watched a clock face (diameter: 2.7 degrees of visual angle, dva) with a rapid clockwise rotating hand (1800 ms per revolution). The clock hand started rotating from a random angle, and participants were asked to make a keypress (‘k’ on a standard QWERTY keyboard with the right index finger) no earlier than 1 second from the trial start. A trial would be aborted with a visual warning signal and repeated if it was made before 1 second. After the keypress, a visual probe (a disk with a diameter of

0.1 dva) and a sound (1000 Hz tone, 50 ms long, 5 ms rise/fall envelop, comfortable volume level) were both presented with a delay of 250 ms. The visual probe was presented for 30 ms outside the clock rim (distance to the clock centre: 1.5 dva) but aligned to the position of the clock hand at the onset of the visual probe. After the visual probe presentation, the clock hand continued rotating for a random period between 750 and 1250 ms. Participants were then asked if a visual probe was detected.

In the AS condition, a visual detection task was employed together with the outcome binding measurement using the Libet clock method (Libet et al., 1983). Participants were asked to report the time of a sound play (timing report trials) or to report if a visual probe was detected (a visual probe was presented in visual probe trials to assess the distribution of attention; no visual probe was presented in catch trials, for assessing the false alarm rate). In each trial, the clock hand started rotating from a random angle (Figure 1b). Participants were asked to make a keypress (‘k’ on a standard QWERTY keyboard with the right index finger) at their own decision. However, they were told that no strategies should be used to plan the keypress time (e.g. making a keypress when the clock hand was at 3 o’clock position) and that the keypress should not be made within 1 second from the start of the trial. If a keypress was made within 1 second from the start of the trial, a visual warning signal was displayed, and the trial was repeated. After the keypress, a sound was played via the headphones (1000 Hz tone, 50 ms long, 5 ms rise/fall envelop, comfortable volume level) with a 250 ms delay. After the sound, the clock hand continued rotating for a random period between 750 and 1250 ms. In visual probe trials, the threshold titrated visual probe was presented for 30 ms. There were 6 possibilities for the onset time of the visual probe, which was defined in relation to the delay from the keypress: 0 ms, 100 ms, 200 ms, 300 ms, 400 ms, or 500 ms delay. The visual probe was presented outside the clock rim (distance to the clock centre: 1.5 dva) but aligned to the position of the clock hand at the onset of the visual probe. For example, if a keypress was made when the clock hand was at -50^o (Figure 1b; polar coordinates with 12 o'clock position being 0^o, 3 o'clock position being 90^o, and 9 o'clock position being -90^o), the 6 possibilities of visual probe location/timing were -50^o (0 ms delay), -30^o (100 ms delay), -10^o (200 ms delay), 10^o (300 ms delay), 30^o (400 ms delay), and 50^o (500 ms delay). In this case, the sound was played when the clock hand was at 12 o'clock position (0^o, 250 ms delay). In catch trials and timing report trials, no visual probe was presented. At the end of the trial, participants should indicate if a visual probe was detected in both visual probe trials and catch trials. In timing report trials, the clock hand always stopped at the 12 o'clock position. Participants should move the clock hand to its position at the time of the sound, using the left hand (pressing 'a' and 's' to move the clock hand counter-clockwise by 10^o and 1^o, respectively; pressing 'd' and 'f' to move the clock hand clockwise by 1^o and 10^o, respectively). There were 50 catch trials, 50 timing report trials, and 180 visual probe trials (30 trials for each visual probe location/timing). Trials were presented in a random order. 5 trials of practice were given before the formal testing. Since there were more trials asking for visual probe detection than trials asking for timing report, the participants were told in the beginning of the experiment that they should perform the task as if the timing report were required in each trial, with the aim of ensuring a good quality in the timing report.

The SO condition was identical to the AS condition except that the sound play was controlled by computer (i.e. no keypress was required to trigger the sound). The timing information of the keypress and the stimulus presentation from the AS condition was recorded for its full replication in the SO condition. For this reason, the SO condition always followed the AS condition.

The stimuli were presented on a liquid crystal display screen (refresh rate: 100 Hz; 24 inch screen size). Stimulus generation and presentation was controlled by Psychtoolbox-3 (Kleiner et al., 2007) using Matlab (The MathWorks Inc., USA). The sound was presented using a pair of headphones (Beyerdynamic DT 770 pro, 32 OHM, Germany). The experiment was performed in a well-lit, soundproof testing booth.

Experiment 2

A new group of 15 participants were recruited for Experiment 2 (8 females; mean age = 21.5, *SD* = 2.4). Experiment 2 was the same as Experiment 1 except that a strict eye movement control was employed. In Experiment 1, participants were asked to fixate the centre of the clock face but no measures were taken to enforce this requirement. In Experiment 2, the movements of the right eye were monitored at 1000 Hz with an eye tracking device (Eyelink Portable Duo, SR Research Ltd, Canada). In all the three parts of testing (threshold testing, AS condition, and SO condition), participants fixated the centre of the clock face. If the right eye was out of a square area subtending 2.0 dva from the centre of the clock face at any time from the start of a trial to the point when the clock hand stopped rotating, the trial would be aborted with a visual warning signal and repeated after. Because of this strict eye movement control, trial repetitions were more often, and the average testing duration of Experiment 2 increased to about 180 minutes.

Experiment 3

A new group of 15 participants were recruited for Experiment 3 (10 females; mean age = 20.9, *SD* = 2.7). Experiment 3 has the following 3 changes as compared to Experiment 2. First, the SO condition was replaced by a Vibration Sound condition (VS). In the VS condition, the right index finger (the keypressing finger) received a mild and short vibrotactile stimulation (two impulses in 10 ms) from a miniature electromagnetic solenoid-type stimulator (Dancer Design, UK) 250 ms before the sound play. The onset of the vibrotactile stimulation aligned with the keypress time in the AS condition. Second, the total number of trials in each condition reduced to 170 (20 catch trials, 30 timing report trials, and 120 visual probe trials with 20 trials for each of the 6 visual probe locations). Third, the visual fixation area was a circle centred on the clock centre (diameter was the same as the side length of the square in Experiment 2). The average testing duration was about 110 minutes.

Data analysis

The same data analysis procedure was applied to the 3 experiments (the VS condition in Experiment 3 was treated as the SO condition in Experiment 1 and 2 in the data analysis). All the findings reported in the results section were based on the remaining 13 participants from each experiment (2 participants were excluded for each experiment, see below).

For the timing report trials, the reported time in each trial was calculated as the difference between the reported position of the clock hand and the actual position of the clock hand at the time of sound onset. This difference in spatial location was converted to a temporal judgment error, based on the clock hand rotation speed of 1800 ms per revolution. The median of the reported time in each condition (50 trials in Experiment 1 and 2; 30 trials in Experiment 3) was used for subsequent processing. Individuals with extreme values of either the median reported time or the standard deviation of reported time across trials were excluded using the median absolute deviation from median (MAD–median) rule: let p be the individual value and P be the individual values from the whole sample. If $|p - \text{median}(P)| \times 0.6745 > 3 \times$ MAD–median, this value is an outlier (Leys et al., 2013). This procedure excluded 1 participant from Experiment 1 and 2, 2 participants from Experiment 3. Individual outcome binding effect was calculated as the difference in the reported time between SO and AS conditions (SO – AS; Experiment 1 and 2) or the difference between VS and AS (VS – AS; Experiment 3). Group level outcome binding effect was evaluated with a one-sided paired t-test comparing the reported time between SO and AS conditions for Experiment 1 and 2 (assuming an outcome binding effect), and a two-sided paired t-test comparing VS and AS conditions for Experiment 3 (assuming the outcome binding effect would be smaller than in Experiment 1 and 2, but not sure to what extent or if the effect could be reversed). Comparisons of the outcome binding effect between experiments were made with one-sided unpaired t-tests.

For the catch trials (without visual probe presentation), the false alarm rate was calculated as the ratio of trials reporting a visual probe was detected. One more participant from Experiment 1 and 2 was excluded due to extremely high false alarm rate (0.63 and 0.95). For the remaining participants, the average false alarm rate was 0.05 (*SD* = 0.08) in Experiment 1, 0.02 (*SD* = 0.02) in Experiment 2, and 0.06 (*SD* = 0.07) in Experiment 3.

For the visual probe trials (with visual probe presentation), the detection rate was calculated for each of the 6 probe locations as the ratio of trials where a visual probe was detected, relative to the total number of visual probe trials. This was used as a measure of visual attention. Since the clock hand position at the keypress time in each trial was random, it was realigned to the -50^{o} position for data processing so that the clock hand position at the keypress time in each trial was -50^{o}. After the realignment, visual probes presented with delays (relative to the keypress time) of 0 ms was at -50^{o}, with a delay of 100 ms was at -30^{o}, and so on. The attention distribution was compared between conditions using a two-way (condition: AS vs. SO in Experiment 1 and 2, AS vs. VS in Experiment 3; probe location: -50^{o}, -30^{o}, -10^{o}, 10^{o}, 30^{o}, and 50^{o}) within-participants ANOVA. For each condition, an attention shift score was calculated as the difference between the average detection rate after the sound play (i.e. locations 10^{o}, 30^{o}, and 50^{o}) and the average detection rate before the sound play (i.e. locations -50^{o}, -30^{o}, and -10^{o}). An attention difference index was then obtained by subtracting the attention shift score in the AS condition from the SO condition (Experiment 1 and 2), or from the VS condition (Experiment 3).

To assess the relationship between action-related attention modulation and outcome binding, a cross-participants correlation analysis (Spearman's rank correlation) was performed between the attention difference index and the outcome binding effect. Since a positive correlation was predicted, the right-tailed *p* value was reported. Bivariate outliers in the correlation analysis were detected using the box-plot rule (Pernet et al., 2013).

Declarations of interest: none.

Open Practices Statement

The study reported in this article was not preregistered. The data and analysis code are available from https://figshare.com/s/d10a00c3e76ebf1cfedf, and will be uploaded to Figshare for unrestricted access upon the acceptance of the paper.

Acknowledgements

We would like to thank Patrick Haggard and Wilfried Kunde for helpful comments on an earlier draft. LC was supported by the National Natural Science Foundation of China (grant number: 32271078).

References

Antusch, S., Custers, R., Marien, H., & Aarts, H. (2021). Intentional action and limitation of personal autonomy. Do restrictions of action selection decrease the sense of agency? *Consciousness and Cognition*, *88*. https://doi.org/10.1016/j.concog.2021.103076

Baldauf, D., & Deubel, H. (2010). Attentional landscapes in reaching and grasping. *Vision Research*, *50*(11), 999-1013. https://doi.org/10.1016/j.visres.2010.02.008

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

Buehner, M. J., & Humphreys, G. R. (2009). Causal Binding of Actions to Their Effects. *Psychological Science*, *20*(10), 1221-1228. https://doi.org/10.1111/j.1467-9280.2009.02435.x

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827-1837. https://doi.org/Doi 10.1016/0042-6989(95)00294-4

Deubel, H., Schneider, W. X., & Paprotta, I. (1998). Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition*, *5*(1-2), 81-107. https://doi.org/Doi 10.1080/713756776

Dogge, M., Schaap, M., Custers, R., Wegner, D. M., & Aarts, H. (2012). When moving without volition: Implied self-causation enhances binding strength between involuntary actions and effects. *Consciousness and Cognition*, *21*(1), 501-506. https://doi.org/10.1016/j.concog.2011.10.014

Eimer, M., Van Velzen, J., Gherri, E., & Press, C. (2006). Manual response preparation and saccade programming are linked to attention shifts: ERP evidence for covert attentional orienting and spatially specific modulations of visual processing. *Brain Research*, *1105*, 7-19. https://doi.org/10.1016/j.brainres.2005.10.060

Engbert, K., Wohlschlager, A., Thomas, R., & Haggard, P. (2007). Agency, subjective time, and other minds. *Journal of Experimental Psychology-Human Perception and Performance*, *33*(6), 1261-1268. https://doi.org/10.1037/0096-1523.33.6.1261

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 197-208. https://doi.org/10.1038/nrn.2017.14

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

Haggard, P., & Cole, J. (2007). Intention, attention and the temporal experience of action. *Consciousness and Cognition*, *16*(2), 211-220. https://doi.org/10.1016/j.concog.2006.07.002

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude Estimation Reveals Temporal Binding at Super-Second Intervals. *Journal of Experimental Psychology-Human Perception and Performance*, *35*(5), 1542-1549. https://doi.org/10.1037/a0014492

Ivanof, B. E., Terhune, D. B., Coyle, D., Gottero, M., & Moore, J. W. (2022). Examining the effect of Libet clock stimulus parameters on temporal binding. *Psychological Research-Psychologische Forschung*, *86*(3), 937-951. https://doi.org/10.1007/s00426-021-01546-x

Kirsch, W., Kunde, W., & Herbort, O. (2019). Intentional Binding Is Unrelated to Action Intention. *Journal of Experimental Psychology-Human Perception and Performance*, *45*(3), 378-385. https://doi.org/10.1037/xhp0000612

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 14-14.

Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications*, *10*(1), 1-11. https://doi.org/10.1038/s41467-019-12170-0

Levitt, H. (1971). Transformed up-down Methods in Psychoacoustics. *Journal of the Acoustical Society of America*, *49*(2), 467-477. https://doi.org/Doi 10.1121/1.1912375

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764-766. https://doi.org/10.1016/j.jesp.2013.03.013

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) - the Unconscious Initiation of a Freely Voluntary Act. *Brain*, *106*(Sep), 623-642. https://doi.org/10.1093/brain/106.3.623

Martinez, P. I. C., Maggioni, E., Hornbaek, K., Obrist, M., & Subramanian, S. (2018). Beyond the Libet Clock: Modality Variants for Agency Measurements. *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems (Chi 2018)*. https://doi.org/10.1145/3173574.3174115

Miller, J., Vieweg, P., Kruize, N., & McLea, B. (2010). Subjective reports of stimulus, response, and decision times in speeded tasks: How accurate are decision time reports? *Consciousness and Cognition*, *19*(4), 1013-1036. https://doi.org/10.1016/j.concog.2010.06.001

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*(1), 546-561. https://doi.org/10.1016/j.concog.2011.12.002

Muth, F. V., Wirth, R., & Kunde, W. (2021). Temporal binding past the Libet clock: testing design factors for an auditory timer. *Behavior Research Methods*, *53*(3), 1322-1341. https://doi.org/10.3758/s13428-020-01474-5

Pernet, C. R., Wilcox, R., & Rousselet, G. A. (2013). Robust correlation analyses: false positive and power validation using a new open source matlab toolbox. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00606

Pockett, S., & Miller, A. (2007). The rotating spot method of timing subjective events. *Consciousness and Cognition*, *16*(2), 241-254. https://doi.org/10.1016/j.concog.2006.09.002

Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature Neuroscience*, *14*(2), 252-U347. https://doi.org/10.1038/nn.2711

Rolfs, M., Lawrence, B. M., & Carrasco, M. (2013). Reach preparation enhances visual performance and appearance. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *368*(1628). https://doi.org/10.1098/rstb.2013.0057

Rolfs, M., & Schweitzer, R. (2022). Coupling perception to action through incidental sensory consequences of motor behaviour. *Nature Reviews Psychology*, *1*, 112–123. https://doi.org/10.1038/s44159-021-00015-x

Sanford, A. J. (1974). Attention bias and the relation of perception lag to simple reaction time. *J Exp Psychol*, *102*(3), 443-446. https://doi.org/10.1037/h0035871

Seifried, T., Ulrich, R., Bausenhart, K. M., Rolke, B., & Osman, A. (2010). Temporal preparation decreases perceptual latency: Evidence from a clock paradigm. *Quarterly Journal of Experimental Psychology*, *63*(12), 2432-2451. https://doi.org/10.1080/17470218.2010.485354

Tanaka, T., Matsumoto, T., Hayashi, S., Takagi, S., & Kawabata, H. (2019). What makes action and outcome temporally close to each other: A systematic review and meta-analysis of temporal binding. *Timing & Time Perception*, *7*(3), 189-218. https://doi.org/10.1163/22134468-20191150

Tipper, S. P., Howard, L. A., & Houghton, G. (1998). Action-based mechanisms of attention. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *353*(1373), 1385-1393. https://doi.org/DOI 10.1098/rstb.1998.0292

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

Wen, W., & Imamizu, H. (2022). The sense of agency in perception, behaviour and human–machine interactions. *Nature Reviews Psychology 1*, 211-222. https://doi.org/10.1038/s44159-022-00030-6

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. https://doi.org/10.1007/s00221-013-3419-2

Wundt, W. M. (1874). *Grundzüge der physiologischen Psychologie*. verlag von Wilhelm Englemann.

Yabe, Y., & Goodale, M. A. (2015). Time Flies When We Intend to Act: Temporal Distortion in a Go/No-Go Task. *Journal of Neuroscience*, *35*(12), 5023-5029. https://doi.org/10.1523/Jneurosci.4386-14.2015