Illusion of illusion: A spatial-attentional mechanism underlies action-related distortions of time perception

Liyu Cao[1,2]

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

[2]State Key Laboratory of Brain-Machine Intelligence, Zhejiang University, 311121, Hangzhou, China

Correspondence to: Dr. Liyu Cao, Department of Psychology and Behavioural Sciences, Zhejiang University, 310058, Hangzhou, China, email: liyu.cao@zju.edu.cn

Abstract

Temporal binding has been understood as an illusion of timing perception. When an action triggers an outcome (e.g. a sound) after a brief delay, the action is perceived to occur later than if the outcome does not occur, and the outcome is perceived to occur earlier than a similar outcome not caused by an action. We show here that an attention mechanism can explain the seeming illusion of timing perception. In one method, participants watch a rotating clock hand and report event times by noting the clock hand position when the event occurs. We find that visual attention is critically involved in shaping event time reports made in this way. This occurs because action and outcome events result in shifts of attention around the clock rim, biasing the perceived the location of the clock hand. Using a probe detection task to measure attention, we further show a difference in attention distribution between a single-event condition (sound only or action only) and a two-event agency condition (action plus sound). Participants accordingly report that they perceive the same event (the sound or the action) at different times in the two conditions: spatial attentional shifts masquerading as temporal binding. Studies that use time perception as an implicit marker of voluntary agency should first discount the artefactual changes in perceived event timing that actually reflect differences in spatial attention. The study also has important implications for related results in mental chronometry obtained with the clock-like method since Wundt, as attention may well be a critical confounding factor in the interpretation of these studies.

Keywords: Action; Attention; Temporal binding; Intentional binding; Sense of agency

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, known as temporal binding (or intentional binding). Temporal binding is comprised of action binding (i.e. the action being perceived as occurring later) and outcome binding (i.e. the sensory event being perceived as occurring earlier), with outcome binding having a much bigger effect size (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).

However, the impact of visuospatial attention has been heavily overlooked when considering the measurement of timing perception with the widely used Libet clock method (Libet et al., 1983; Wundt, 1874). In the standard testing procedure (e.g. Haggard et al., 2002), an event was presented (e.g. a keypress or a sound) while participants watched a clock face with a rapid rotating clock hand. Participants should indicate the event time by reporting where the clock hand was positioned at the event onset. This essentially transforms a timing task to a spatial localisation task. It is well-know that attention plays an important role in spatial localisation (Fortenbaugh & Robertson, 2011; Tse et al., 2011; Visser & Enns, 2001; Zhou et al., 2017). When the location of an object is ambiguous, participants tend to localise the object to the place where attention is directed to (Adam et al., 2008; Binda et al., 2009;

Kirsch, 2015). This is particularly relevant to the temporal binding measured with Libet clock method, as the fast rotating clock hand makes the location report a challenging task (Haggard & Cole, 2007).

Outcome binding is usually obtained by comparing an Action Sound (AS) condition and a Sound Only (SO) condition. In both conditions, participants report where the clock hand pointed to at the time of sound play. The reported time is earlier in the AS condition than in the SO condition (Figure 1a). Suppose that the clock hand is positioned at 12 o'clock when

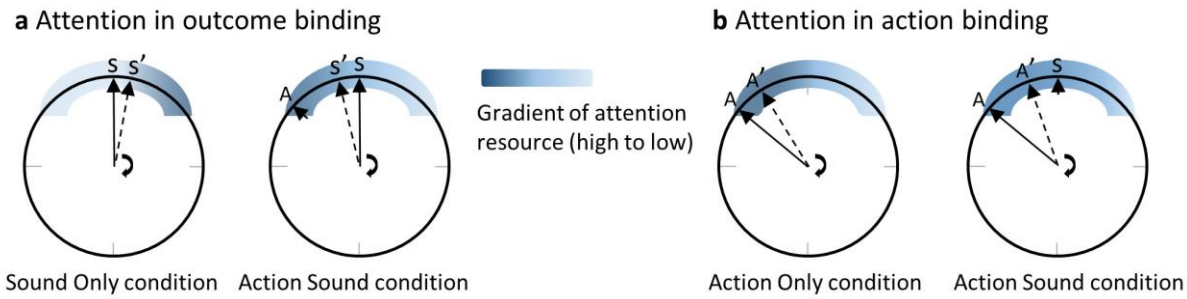


Figure 1 The attention hypothesis of temporal binding. (a) Attention in outcome binding. The distribution of attention around the clock rim, at the time close to the event requiring a timing report, receives modulation from both action and action outcome. When the sound time is reported, attention increases only after the onset of the sound in the sound only condition. In the action sound condition, attention is activated prior to the sound onset due to action. The difference in the attention distribution between the two conditions can lead to the difference in the reported clock hand position at the time of sound onset (i.e. outcome binding). (b) Attention in action binding. When the action time is reported, the sound in the action sound condition is an extra cue for attention activation compared to the action only condition. Therefore, there is more attention in the action sound condition than the action only condition at the clock hand positions after the sound play time, leading to a later reported clock hand position in the action sound time (i.e. action binding). Please refer to the text for detailed information. A stands for the actual clock hand position when the keypress is made, and A' is the reported A from participants. S stands for the actual clock hand position when the sound is played, and S' is the reported S from participants.

a sound is played (Figure 1a). In the SO condition, the sound is controlled by the computer. 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, the level of attention resource is low before the sound play but high after. In the AS condition, the sound is triggered by an action, which is executed when the clock hand is at around 10 o'clock position (given a sound delay of 250 ms and a revolution period of the clock hand for 1800 ms as employed in the current

study). Previous research has provided ample evidence that 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 keypress in the Libet clock method can be successfully performed rather easily without considering the information of a target. However, the timing report 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 a target in a target reaching task. This attention shift should result in a high level of attention resource before the sound onset in the AS condition (Figure 1a). The predicted attention difference between AS and SO conditions bears a striking resemblance to the timing report difference between the two conditions. Due to the tight link between attention and spatial localisation as noted earlier, an earlier attention activation in the AS condition compared to the SO condition can lead to an earlier reported clock hand position (i.e. outcome binding).

The attention hypothesis may also be extended to action binding (Figure 1b). Action binding is obtained by comparing an AS condition and an Action Only (AO) condition. The reported keypress time is later in the AS condition than in the AO condition. The only difference between the two conditions is that no sound is played after a keypress in the AO condition. The sound in the AS condition may have the effect as an extra source for attention attraction, resulting in a higher level of attention resource after the keypress in the AS condition than in the AO condition, which can lead to a later reported clock hand position at the keypress time in the AS condition (i.e. action binding).

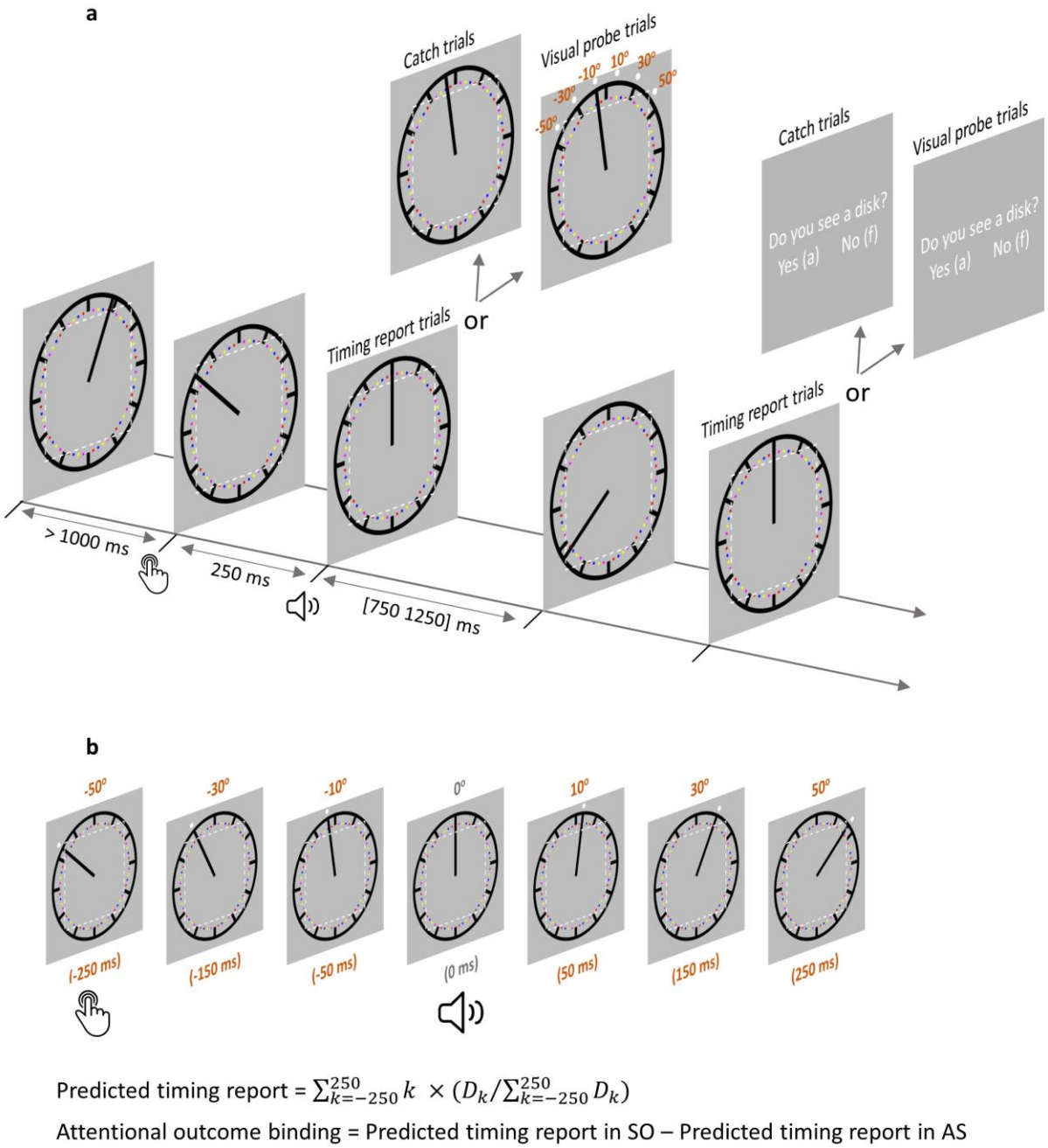
In a series of 4 experiments, we demonstrated distinct patterns of attention modulation induced by action and sensory stimulation using the Libet clock method. Attention had a strong and robust impact on the measure of timing perception, and as a result, on temporal binding, suggesting a fundamental involvement of attention in temporal binding.

Results

An attention distribution shift in outcome binding

Participants reported the onset time of a sound using the Libet clock method (Figure 2). During the experiment, participants were asked to watch the centre of the visual presentation. A clear outcome binding effect was confirmed in Experiment 1 ($t(12) = -8.84$, $p < 0.001$, $dz = -2.45$, one-tailed paired t-test; Figure 3a,c; no eye movements control). 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 was controlled by computer (SO condition; *M* = 35.00 ms, *95% CI* = [7.99 62.01] ms). Crucially, the pattern of visuospatial attention distribution around the time of the sound (the event being judged), operationalised as the probe detection rate, was drastically different between the two conditions (Figure 3b). In the SO condition, the detection rate was low before the sound play, but high after the sound play. In the AS condition, the detection rate was high at the time of keypress, gradually increased, and peaked just before the sound play. After the sound play, the detection rate underwent a sharp decrease. The attention distribution difference 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 rate ($F(5,60) = 8.94$, $p < 0.001$, $\eta_p^2 = 0.43$). 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 ($F(1,12) = 3.30$, $p = 0.094$, $\eta_p^2 = 0.22$).

To quantify the influence of attention on outcome binding, we calculated the predicted timing report based on the attention distribution pattern for each condition. An attentional outcome binding was then obtained by taking the difference of the predicted timing report between SO and AS conditions, as similarly done for the actual outcome binding effect. There was a significant attentional outcome binding effect (M = 34.06 ms, $t(12) = 3.12$, $p = 0.004$, one-tailed one-sample t-test; Figure 3c).



Predicted timing report = $\sum_{k=-250}^{250} k \times (D_k / \sum_{k=-250}^{250} D_k)$

Attentional outcome binding = Predicted timing report in SO – Predicted timing report in AS

Figure 2 Trial structure in the AS condition. (a) 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 (all 6 probe locations illustrated) and catch trials, participants reported if a visual probe was detected. 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. (b) A visual probe was presented in each visual probe trial at one of 6 possible locations, which also corresponds to 6 different time points. The position where the clock hand pointed to at the time of sound play was defined as 0° position (0 ms). -50° position (-250 ms) corresponds to the location where the clock hand pointed to when a keypress was made in the AS condition. There were 3 probe locations before 0° and 3 probe locations after 0°. Note that the visual probe was made salient only for the purpose of illustration. SO: sound only condition; AS: action sound condition.

The above findings were replicated in Experiment 2 with a new group of 15 participants. Experiment 2 was the same as Experiment 1 except that strict eye movements control was applied. During the testing, the eye fixation of participants never exceeded the central area of the clock face (ensured via an eye tracking device). Again, a clear outcome binding effect was confirmed ($t(12) = -2.75$, $p = 0.009$, $dz = -0.76$, one-tailed paired t-test; Figure 3d,f). 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. In the SO condition, attention was low before the sound play, but high after. In the AS condition, attention was high at the time of keypress, gradually increased to its peak just before the sound play, and started to decrease after the sound play (Figure 3e). 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 rate 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, η_p^2 = 0.08). The attentional outcome binding effect was again significantly larger than 0 (M = 26.95 ms, $t(12)$ = 3.20, p = 0.004, one-tailed one-sample t-test; Figure 3f).

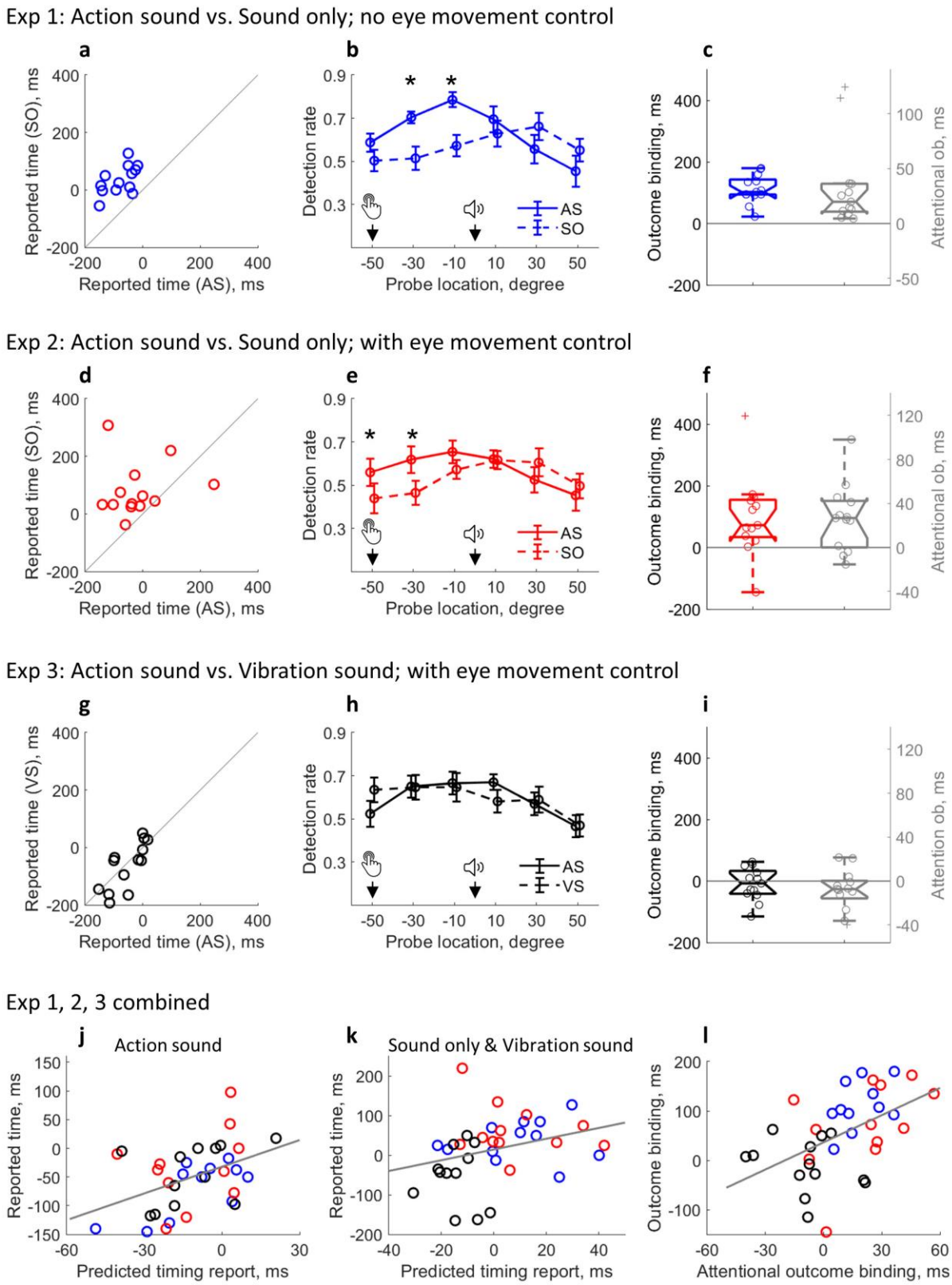


Figure 3 Changed attention and its relevance to outcome. (a-c) Results from Experiment 1. a) Individual reported sound time in each condition; b) Detection rate of visual probes as a

function of condition and probe location. 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. Asterisks indicate significant differences between two conditions at the indicated probe location (false discovery rate adjusted over 6 comparisons); c) Box-plots of outcome binding and attentional outcome binding (i.e. outcome binding predicted from the attention distribution pattern). (d-f) Results from Experiment 2 with eye movements control. (g-i) Results from Experiment 3 with eye movements control. An VS condition replaced the SO condition. (j) Scatter plot of the correlation between attention distribution index and the reported time in AS condition, with data from Experiment 1-3 combined. (k) Same as j but showing the relationship in the SO and VS condition. (l) Scatter plot of the correlation between attention difference index and outcome binding. AS: action sound condition; SO: sound only condition; VS: vibration sound condition.

The attention distribution shift is functionally relevant to outcome binding

Experiment 1 and 2 demonstrated the existence of an attention effect in outcome binding. Experiment 3 sought to demonstrate a functional relevance of the attention effect in outcome binding. To this end, the SO condition was replaced with a VS condition, in which a vibrotactile stimulation was applied to the finger 250 ms before the sound play (no keypress). 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 timing report task. If this is the case, the binding pattern 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 binding was found between VS and AS conditions ($t(12) = 0.68$, $p = 0.509$, $dz = 0.19$, two-tailed paired-sample t-test; 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 3g,i). The size of binding in Experiment 3 (as calculated in Experiment 1 and 2) was smaller than in Experiment 1 ($t(24) = -6.29$, $p < 0.001$, $ds = -2.47$, one-tailed unpaired t-test) and Experiment 2 ($t(24) = -2.80$, $p = 0.005$, $ds = -1.10$, one-tailed unpaired t-test). 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 3h). The two-way (condition: AS vs. VS; probe location: -50°, -30°, -10°, 10°, 30°, and 50°) within-participants ANOVA comparing the detection rate only revealed a significant main effect of probe location ($F(5,60) = 9.04$, $p < 0.001$, $\eta_p^2 = 0.43$). 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. Furthermore, the attentional outcome binding effect was not statistically different from 0 ($M = -8.13$ ms, $t(12) = -1.59$, $p = 0.137$, two-tailed one-sample t-test; Figure 3i), but significantly smaller than the attentional outcome binding in Experiment 1 ($t(24) = -3.50$, $p < 0.001$, $ds = -1.37$, one-tailed unpaired t-test) and Experiment 2 ($t(24) = -3.56$, $p < 0.001$, $ds = -1.40$, one-tailed unpaired t-test). Therefore, it seems to indicate that the outcome binding disappeared when attention was controlled for.

Combining the data from the 3 experiments, we also found a tight link between the timing report predicted from the attention distribution and the actual reported time. In the AS condition, a significant positive correlation was found between the two variables ($r(32) = 0.41$, $p = 0.009$, one-tailed Spearman correlation, 5 outliers removed, Figure 3j; without removing outliers: $r(37) = 0.42$, $p = 0.004$, one-tailed Spearman correlation). The same correlation was also found in the SO condition (here we combined the SO condition from Experiment 1 & 2 and the VS condition from Experiment 3) ($r(34) = 0.44$, $p = 0.004$, one-tailed Spearman correlation, 3 outliers removed, Figure 3k; without removing outliers: $r(37) = 0.48$, $p < 0.001$, one-tailed Spearman correlation).

Besides the correlation between attention distribution and the actual reported time in a single condition, the outcome binding calculated according to the attention distribution difference between SO and conditions (attentional outcome binding) also predicted the actual outcome binding effect. Across participants, a significant correlation between the attentional outcome binding and the actual outcome binding was found ($r(34) = 0.56$, $p < 0.001$, one-tailed

Spearman correlation, 3 outliers removed, Figure 3l; without removing outliers: $r(37) = 0.61$, $p < 0.001$, one-tailed Spearman correlation). It is interesting to note that a linear regression analysis taking the actual outcome binding as the response variable and the attentional outcome binding as the predictor gave an intercept of 36.44 ms ($t(34) = 2.76$, $p = 0.009$, *95% CI* = [9.61 63.26], 3 outliers removed; without removing outliers: intercept 39.07 ms, $t(37) = 2.59$, $p = 0.014$, *95% CI* = [8.53 69.61]). This 36.44 ms could be understood as the outcome binding effect remaining after the attention measure in the current study was controlled for.

A similar attention mechanism in action binding

Attention might influence action binding because the action and the sound in the AS (action sound) condition both can be salient cues for attention orientation. Whereas in the AO (action only) condition, no sound was played after the action. Therefore, it is conceivable that the attention resource right after the time of sound play should be higher in the AS condition than in the AO condition, resulting in a larger attention distribution index in the AS condition than in the AO condition (Figure 1b). Experiment 4 tested the idea of attention involvement in action binding by measuring action binding and attention distribution at the same time as similarly done with outcome binding (Figure 4a). The action binding effect was confirmed ($t(22) = -3.45$, $p = 0.001$, $dz = -0.72$, one-tailed paired t-test; Figure 4b,d). The reported keypress time was later in the AS condition ($M = 102.28$ ms, *95% CI* = [86.06 118.50] ms) than in the SO condition ($M = 78.37$ ms, *95% CI* = [58.91 97.83] ms). The general pattern of attention distribution looks different to that found in the outcome binding task. Attention was high right after the keypress in both AS and AO conditions and then underwent a sharp drop (Figure 4c). This is probably due to the fact that the task here was to report the keypress time rather than the sound time. Therefore, attention dropped quickly after the keypress execution. Importantly, A two-way (condition: AS vs. AO; probe location: -50°, -40°, -30°, -20°, -10°, 0°, 10°, 20°, 30°, 40° and 50°) within-participants ANOVA comparing the detection rate

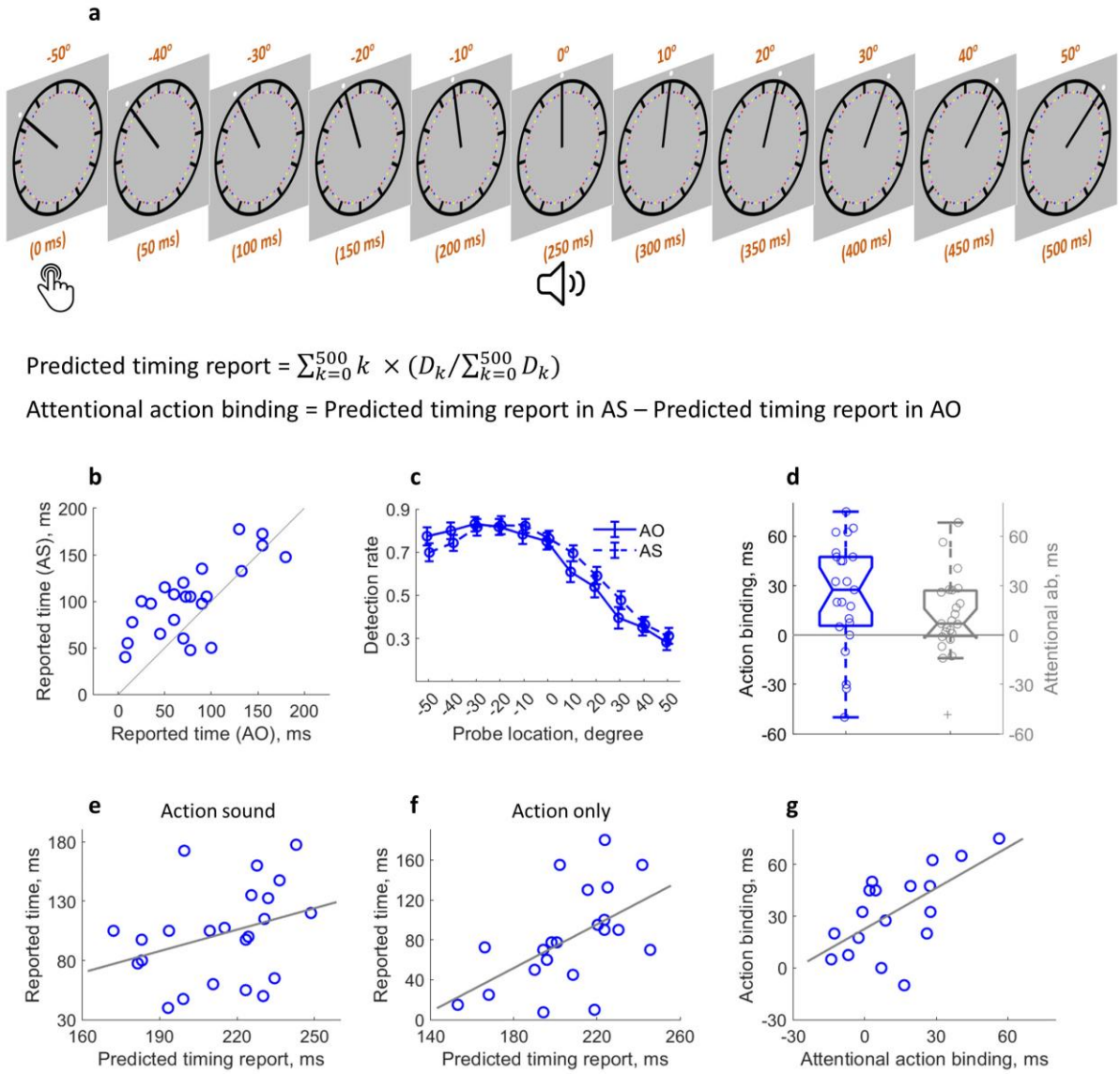


Figure 4 Attention involvement in action binding. (a) Attention was measured at 11 locations (time points) following the keypress. (b) Individual reported keypress time in each condition. (c) Detection rate of visual probes as a function of condition and probe location. Bars represent ± 1 standard error. (d) Box-plots of action binding and attentional action binding (i.e. action binding predicted from the attention distribution pattern). (e) Scatter plot of the correlation between attention distribution index and the reported time in AS condition. (f) Same as e but showing the relationship in the AO condition. (l) Scatter plot of the correlation between attention difference index and action binding. AS: action sound condition; AO: action only condition.

revealed a significant interaction effect ($F(10,220) = 2.93$, $p = 0.026$, $\eta_p^2 = 0.12$). The main effect of probe location was unsurprisingly significant ($F(10, 220) = 79.56$, $p < 0.001$, $\eta_p^2 = 0.78$), and the main effect of condition was not significant ($F(1,22) = 1.66$, $p = 0.210$, $\eta_p^2 = 0.07$). We also calculated the attentional action binding based on the attention distribution

pattern as similarly done with the attentional outcome binding. There was a significant attentional action binding (M = 11.74 ms, $t(22) = 2.31$, $p = 0.015$, one-tailed one-sample t-test; Figure 4d).

For the relationship between the timing report predicted from the attention distribution and the actual reported keypress time in each condition, significant correlations were found in both AS ($r(21) = 0.44$, $p = 0.018$, one-tailed Spearman correlation, Figure 4e; no outlier detected) and AO ($r(20) = 0.62$, $p = 0.001$, one-tailed Spearman correlation, 1 outlier removed, Figure 4f; without removing outliers: $r(21) = 0.64$, $p < 0.001$, one-tailed Spearman correlation) conditions. Furthermore, the attentional action binding significantly correlated with the actual action binding ($r(16) = 0.61$, $p = 0.004$, one-tailed Spearman correlation, 5 outliers removed, Figure 4g; without removing outliers: $r(21) = 0.55$, $p = 0.003$, one-tailed Spearman correlation). A linear regression analysis taking actual action binding as the response variable and the attentional action binding as the predictor gave an intercept of 22.74 ms ($t(16) = 4.14$, $p < 0.001$, *95% CI* = [11.10 34.38], 5 outliers removed; without removing outliers: intercept 15.93 ms, $t(21) = 2.33$, $p = 0.030$, *95% CI* = [1.68 30.17]). This 22.74 ms could be understood as the action binding effect remaining after the attention measure in the current study was controlled for.

Discussion

The current study investigated the visuospatial attention distribution in the temporal binding effect measured with the classic Libet clock method. In 4 complementary experiments, it was demonstrated that the visuospatial attention had a direct link to the event timing report, with the reported time corresponding to the location on the clock where attention was paid to. Action and sensory input both can activate the attention concentration to the clock, resulting in a dynamic attention distribution pattern around the event of which a timing report was required. Critically, the difference in the attention distribution pattern activated by action and

sensory input may explain the temporal binding effect. Temporal binding consists of two effects in opposite directions, i.e. a shift of outcome time towards the past in outcome binding, and a shift of action time towards the future in action binding. The attention account provides a unified explanation for both outcome binding and action binding in terms of an attention bias.

It seems thus an immediate question is what temporal binding really means. The canonical view of temporal binding is that it reflects an illusion in timing perception (Haggard et al., 2002; Tanaka et al., 2019; Wen & Imamizu, 2022). That is, an action is perceived to occur later when it is followed by a sensory outcome than when not (action binding), and a sensory input is perceived to occur earlier when it follows an action than when not (outcome binding). In contrast, the data from the current study illustrate the fact that when the event timing perception is measured with the Libet clock method, there are differences in the contrasting conditions with regards to how the timing perception is read out from the measurement tool (i.e. the Libet clock). Therefore, the bias in the way of reading measurements poses a challenge to a clear understanding of the variable being measured (i.e. timing perception). It is unlikely that a changed timing perception leads to a changed attention distribution, at least in outcome binding. In the outcome binding test, the attention activation following the keypress in the AS condition was before the sound onset, i.e. before the event that requires timing perception (Figure 3b,e). Since timing perception is a constructive process which integrates available information from both before and after the event of interest (Cao et al., 2020; Haggard et al., 2002; Klaffehn et al., 2021; Moore & Haggard, 2008; Takahata et al., 2012), there is a possibility that a common process before the action that triggers both a change in attention and a change in timing perception. One candidate for this common process may be the prediction of the action outcome. However, this idea needs further evidence.

The linear regression analysis though did show that there was a residual outcome binding of 36 ms and a residual action binding of 23 ms when the attention bias in clock hand position reporting was controlled for. Both remainders were statistically significant. This seems to indicate that there is a timing perception component in temporal binding. However, the impact of attention may have not been fully captured in the current study. In the current testing paradigm, attention may be measured as a three-dimension variable (two dimensions of space and one dimension of time). At any time point, attention measures may be taken at all possible locations. In the current study, attention was only measured at one location for each time point. Furthermore, the attention at the action preparation stage may already have an influence on the timing report. Due to the unpredictability of the exact action time, measuring attention before the action is actually executed is an experimental challenge. This might explain why the size of predicted outcome binding and action binding effects from attention distribution pattern does no match the size of actual binding effects, even though strong and reliable predictive power of attention on the reported event time in all single conditions and on both outcome binding and action binding can be demonstrated. However, experiment 3 provided evidence that outcome binding was not statistically significant when the attention difference was experimentally controlled for between contrasting conditions. Previous research also showed that temporal binding declined when the temporal predictability of action and effect was controlled for (Kirsch et al., 2019). Conceivably, such control allows the attention deployment to operate in a similar manner in different conditions. Therefore, it is an open question about the cognitive underpinnings of temporal binding at the moment.

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 timing repor 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 (or at least confounded by) differences in spatial attention. We should therefore investigate when and how such differences in spatial attention might arise in detail in order to better understand the results produced by this method.

Some limitations of the current study need to be born in mind. First, a full picture of the spatiotemporal visual attention distribution dynamic is still waiting to be revealed in the timing report task with the Libet clock method. In the current study, attention was only measured at a limited time points and locations. A full picture of the attention distribution pattern could help achieve a complete understanding of the impact of attention on the timing report in the Libet clock method (Hon, 2022; Schwarz & Weller, 2023). In addition, a within-participants design paired with an attention intervention paradigm could help determine the impact of attention on temporal binding at the individual level. Second, the Libet clock method may not be the optimal tool for resolving the issue of a timing perception component in temporal binding. Even if attention could fully account for the temporal binding measured with the Libet clock, it could also mean that the timing perception component could not be easily measured in this way (Stetson et al., 2006). The timing perception of the attraction between an action and the action outcome may exist like a Gestalt. However, the Gestalt is

lost when measuring its constituting elements (the perceived time of the action or the outcome). Indeed, temporal binding has also been demonstrated using paradigms in which visual attention does not seem to be 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). Of course, whether attention in other modalities or other cognitive processes could explain the temporal binding effect reported there is an intriguing question for further investigations.

To sum up, the current study provided novel and important insights into the understanding of temporal binding and timing perception in general measured with the Libet clock method. Attention is critically involved in temporal binding, at least when the measurement process involves attention. Our study 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.

Materials and Methods

Experiment 1

Participants

15 participants (9 females; mean age = 21.3, *SD* = 1.9) were recruited from a local participant pool. Assuming an effect size of 0.89 (note that this is the effect size of outcome binding and action binding combined as reported in a meta-analysis, see Tanaka et al. (2019). Since outcome binding has larger effect size than action binding, the real effect size of outcome binding should be larger than this), 13 participants (2 excluded in the data analysis, see below) lead to a statistical power of 0.92 (alpha = 0.05; one-tailed) (Faul et al., 2007). 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.

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 2a). 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 right outside the clock rim (distance to the clock centre: 1.5 dva) close to the top of the clock hand. The location where the clock hand pointed to at the time of sound play was defined as 0° (time 0). Accordingly, the clock hand would point to 10° 50 ms after the sound play, and -10° 50 ms before the sound play, given its rotation speed of 1800 ms per revolution. Visual probes were presented at 6 possible locations (corresponding to 6 time points): -50° (-250 ms; this is when the keypress was made in the AS condition), -30° (-150 ms), -10° (-50 ms), 10° (50 ms), 30° (150 ms), and 50° (250 ms) (Figure 2b). 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 after the random period succeeding the sound. This

is to ensure that the stop position of the clock hand does not introduce any bias in the judgment. 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° and 1°, respectively; pressing ‘d’ and ‘f’ to move the clock hand clockwise by 1° and 10°, 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. Intermixing all trial types provides the advantage of simultaneous measure of timing report and attention distribution. The validity of the testing paradigm can be directly assessed through checking the existence of the classic outcome binding effect. 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. 5 trials of practice were given before the formal testing.

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 of 2.0 dva centring on 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.

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, i.e. 2.0 dva).

Experiment 4

A new group of 30 participants were recruited for the test of attention in action binding (16 females; mean age = 21.2, *SD* = 1.9). Two previous studies from our own group using a

similar set-up reported an effect size of 0.79 (n = 52) (Cao et al., 2021) and 0.62 (n = 42) (Cao et al., 2020). Assuming an effect size of 0.70 here, 23 participants (7 excluded in the data analysis, see below) lead to a statistical power of 0.95 (alpha = 0.05; one-tailed) (Faul et al., 2007).

Action binding and attention were measured in the same experiment using a similar set-up as in the outcome binding. In the action sound (AS) condition, participants made a voluntary keypress, which was followed by a 250 ms delayed sound. In timing report trials, the keypress time was reported as the clock hand position at the time of keypress. In visual probe trials, participants should report if a visual probe was detected during the testing. Visual probes were presented at 11 possible locations (corresponding to 11 time points): -50° (0 ms; this is when the keypress was made), -40° (50 ms), -30° (100 ms), -20° (150 ms), -10° (200 ms), 0° (250 ms; this is when the sound was played), 10° (300 ms), 20° (350 ms), 30° (400 ms), 40° (450 ms), and 50° (500 ms) (Figure 4a). Here the keypress time was defined as time 0 as the event of interest in action binding is the keypress. The location definition followed the routine set in the outcome binding experiments. In the action only (AO) condition, the keypress was not followed by an auditory outcome. Everything else was the same as the AS condition. There were 480 trials in each condition (100 catch trials, 50 timing report trials, and 330 visual probe trials with 30 trials for each of the 11 visual probe locations). The order of trials was randomised.

The order of the AS and the AO condition was counterbalanced between participants as there is a clear impact of the testing order on the size of action binding (Cao et al., 2021). The threshold of the visual probe was obtained prior to the action binding measure using a 2-down-1-up staircase procedure (Levitt, 1971). No eye movements control was applied.

Data analysis

The same data analysis procedure was applied to the 4 experiments.

The false alarm rate in the catch trials (calculated as the ratio of trials reporting a visual probe was detected) was used for participant exclusion. Four participants were excluded due to extremely high false alarm rate (1 from Experiment 1: 0.63; 1 from Experiment 2: 0.95; 2 from Experiment 4: 0.69 and 0.89). For the remaining participants, the average false alarm rate was 0.04 *(SD* = 0.08) in Experiment 1, 0.02 *(SD* = 0.02) in Experiment 2, 0.07 *(SD* = 0.09) in Experiment 3, and 0.04 *(SD* = 0.03) in Experiment 4.

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. 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 standard deviation and the median of the reported time in each condition (50 trials in each condition for Experiment 1, 2, and 4; 30 trials in each condition for Experiment 3) was used for subject exclusion. Individuals with extreme values of either the standard deviation or the median of reported time 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 further excluded 1 participant from Experiment 1 and 2, 2 participants from Experiment 3, and 5 participants from Experiment 4. The remaining participants were included in the formal analysis (13 in Experiments 1-3, 23 in Experiment 4).

Individual outcome binding effect was calculated as the difference in the median reported sound time between SO and AS conditions (SO – AS; Experiment 1 and 2) or the difference between VS and AS (VS – AS; Experiment 3). Individual action binding effect as calculated as the difference in the median reported keypress time between AS and AO conditions (AS – AO). Group level outcome binding effect was evaluated with a one-tailed paired t-test comparing the reported sound time between SO and AS conditions for Experiment 1 and 2

(assuming an outcome binding effect), and a two-tailed 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-tailed unpaired t-tests. Group level action binding effect was evaluated with a one-tailed paired t-test comparing the reported keypress time between AO and AS conditions (assuming an action binding effect).

For the visual probe trials (with visual probe presentation), the detection rate was calculated for each probe location as the ratio of trials with the probe detected. This was used as a measure of visual attention. The potential attention difference between conditions was evaluated with a two-way (condition and probe location) within-participants ANOVA. For each condition, we also calculated the predicted timing report based on the attention distribution pattern for each individual. The assumption is that the actual reported clock hand position should correspond to the location where attention is directed to. A full cover of the spatiotemporal attention distribution pattern may be necessary for an accurate prediction of the actual reported location. The limited (yet highly relevant) attention sampling points in the current study should at least provide a proportionate prediction result. That is, the predicted value may not be exactly as the actual one, but should at least be proportionate to the actual value (e.g. 60% of the actual value). The predicted timing report in the outcome binding experiment (Experiment 1-3) was calculated as: Predicted timing report = $\sum_{k=-250}^{250} k \times (D_k / \sum_{k=-250}^{250} D_k)$, where k is the 6 attention sampling points (i.e. -250, -150, -50, 50, 150, 250 ms; sound play time is 0 ms), D_k is the individual detection rate at time point k. The predicted timing report in the action binding experiment (Experiment 4) was calculated as: Predicted timing report = $\sum_{k=0}^{500} k \times (D_k / \sum_{k=0}^{500} D_k)$, where k is the 11 attention sampling points (i.e. 0, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 ms; keypress time is 0 ms). A

one-tailed Spearman's rank correlation analysis was performed between the predicted timing report and the actual reported time across participants in each single condition, since a positive correlation was predicted. Predicted temporal binding effects based on the attention distribution (attentional temporal binding) were calculated using the predicted timing report in single conditions (SO - AS or VO - AS for attentional outcome binding; AS - AO for attentional action binding). Since positive values were predicted for attentional outcome binding (Experiment 1,2) and attentional action binding (Experiment 4), one-tailed one-sample t-test was used (comparing against 0). In Experiment 3, the attentional outcome binding may be positive or negative (two-tailed one-sample t-test), but was predicted to be smaller than in Experiment 1 and 2 (one-tailed unpaired t-test). As already stated, the attentional temporal binding effect was predicted to be at least proportionate to the actual temporal binding effect (one-tailed Spearman's rank correlation). In all correlation analyses, bivariate outliers in correlation analyses were detected using the box-plot rule and then excluded (Pernet et al., 2013). However, results without outlier exclusion were also reported for reference. In addition, a linear regression analysis was performed by taking the actual temporal binding as the response variable and the attentional temporal binding the predictor. This could help assess the adequacy of the attention measure in the current study.

Compliance and Ethics

We declare no conflict of conflict of interest. We have conformed with the Helsinki Declaration of 1975 (as revised in 2008) concerning Human and Animal Rights.

Open Practices Statement

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.

Acknowledgments

We would like to thank Patrick Haggard for the insightful discussions and helpful comments on an earlier draft, and Wilfried Kunde for helpful comments on an earlier draft. This work was supported by the National Natural Science Foundation of China (grant number: 32271078) and the STI 2030—Major Projects (grant number: 2021ZD0200409).

References

Adam, J. J., Davelaar, E. J., van der Gouw, A., & Willems, P. (2008). Evidence for attentional processing in spatial localization. *Psychological Research-Psychologische Forschung*, *72*(4), 433-442. https://doi.org/10.1007/s00426-007-0126-2

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

Binda, P., Cicchini, G. M., Burr, D. C., & Morrone, M. C. (2009). Spatiotemporal Distortions of Visual Perception at the Time of Saccades. *Journal of Neuroscience*, *29*(42), 13147-13157. https://doi.org/10.1523/Jneurosci.3723-09.2009

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/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

Cao, L., Steinborn, M., Kunde, W., & Haendel, B. (2020). Action force modulates action binding: evidence for a multisensory information integration explanation. *Experimental Brain Research*, *238*(9), 2019-2029. https://doi.org/10.1007/s00221-020-05861-4

Cao, L. Y., Steinborn, M. B., & Haendel, B. F. (2021). Delusional thinking and action binding in healthy individuals. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-97977-y

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/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/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

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/Bf03193146

Fortenbaugh, F. C., & Robertson, L. C. (2011). When here becomes there: attentional distribution modulates foveal bias in peripheral localization. *Attention Perception & Psychophysics*, *73*(3), 809-828. https://doi.org/10.3758/s13414-010-0075-5

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

Hon, N. (2022). Attention and expectation likely underlie temporal binding measured using the Libet Clock. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218221132762

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. (2015). Impact of action planning on spatial perception: Attention matters. *Acta Psychologica*, *156*, 22-31. https://doi.org/10.1016/j.actpsy.2015.01.002

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

Klaffehn, A. L., Sellmann, F. B., Kirsch, W., Kunde, W., & Pfister, R. (2021). Temporal binding as multisensory integration: Manipulating perceptual certainty of actions and their effects. *Attention Perception & Psychophysics*, *83*(8), 3135-3145. https://doi.org/10.3758/s13414-021-02314-0

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/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., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, *17*(1), 136-144. https://doi.org/10.1016/j.concog.2006.12.004

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

Schwarz, K. A., & Weller, L. (2023). Distracted to a fault: Attention, actions, and time perception. *Atten Percept Psychophys*, *85*(2), 301-314. https://doi.org/10.3758/s13414-022-02632-x

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

Stetson, C., Xu, C., Montague, P. R., & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron*, *51*(5), 651-659. https://doi.org/10.1016/j.neuron.2006.08.006

Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's not my fault: postdictive modulation of intentional binding by monetary gains and losses. *Plos One*, *7*(12), e53421. https://doi.org/10.1371/journal.pone.0053421

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/10.1098/rstb.1998.0292

Tse, P. U., Whitney, D., Anstis, S., & Cavanagh, P. (2011). Voluntary attention modulates motion-induced mislocalization. *Journal of Vision*, *11*(3). https://doi.org/10.1167/11.3.12

Visser, T. A. W., & Enns, J. T. (2001). The role of attention in temporal integration. *Perception*, *30*(2), 135-145. https://doi.org/DOI 10.1068/p3089

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

Zhou, L., Deng, C. L., Ooi, T. L., & He, Z. J. J. (2017). Attention modulates perception of visual space. *Nature Human Behaviour*, *1*(1). https://doi.org/10.1038/s41562-016-0004