Consciousness and Cognition 134 (2025) 103909

Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/yccog

Full Length Article

Is the temporal binding effect in the Libet clock-task based in spatial working memory? A correlational and a dual-task approach

Markus Siebertz [*], Petra Jansen

Faculty of Human Sciences, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany

ARTICLE INFO

Keywords:
Intentional binding
Sense of agency
Libet clock
Working memory

ABSTRACT

Intentional binding research attributes the shift of clock hand positions in the Libet clock task to sense of agency-related processes. We investigated the alternative theory that this binding is based on spatial working memory processes. In a correlational design in Experiment 1, 104 young healthy adults performed the Libet clock task, a second version of this task eliminating the continuous movement of the clock, and a spatial and visual 2-back task. The only evidence for the investigated theory was a correlation between outcome binding and the spatial 2-back task. In an experimental within-participant dual-task design in Experiment 2, 94 young, healthy adults did the Libet clock task in the retention intervals of spatial and visual memory tasks. We could not find conclusive evidence for an effect of spatial memory load on binding. Our results suggest that binding in the Libet clock task is not rooted in spatial working memory processes.

1. Introduction

When intentionally performing actions, we typically recognize ourselves as the authors of these actions and we can attribute our actions' consequences on our environment to be caused by us. This ability is called the sense of agency (Moore & Obhi, 2012). It plays an important role in everyday functioning and social interaction. Alterations in the sense of agency have been investigated in various psychiatric disorders such as schizophrenia (Di Plinio et al., 2019; Graham-Schmidt et al., 2016; Martin, 2013; Moore et al., 2013; Voss et al., 2010). One can ask for subjective agency ratings to measure how much agency someone perceives over their actions. An investigation of the mechanisms that underlie the sense of agency, however, additionally requires objective measures.

1.1. Temporal binding and the Libet clock task

One objective measure was proposed by Haggard et al. (2002). They employed a task using the Libet clock (Libet et al., 1983), a clock face with a clock hand moving fast at around 2.5 s per revolution. In this Libet clock task, participants typically report the time when they pressed a button or when they heard a tone by referencing the clock hand position during that event. They do this in baseline conditions during which only one event occurs, their button press or an automatically played tone, and in contingent conditions where their button press triggers the tone. The consistent result is that participants systematically report later clock hand positions for their

* Corresponding author at: Markus Siebertz, Faculty of Human Sciences, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany.

E-mail address: markus1.siebertz@ur.de (M. Siebertz).

https://doi.org/10.1016/j.concog.2025.103909
Received 2 February 2025; Received in revised form 22 June 2025; Accepted 17 July 2025
Available online 29 July 2025
1053-8100/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

button press and earlier positions for the tone in the contingent compared to the baseline conditions. This is interpreted as a contraction of time between the action and the resulting outcome in the participant's perception compared to the temporal perception of the isolated events. As voluntary actions and outcomes caused by these actions are bound together in time, this phenomenon has been called intentional binding.

Besides the Libet clock task, several different procedures have been conceived to measure intentional binding. One method uses a stimulus stream, a sequence of letters, instead of the clock with the task to remember the letter that was presented during the action or the outcome (Cavazzana et al., 2014). Other methods deviate more from the Libet clock task and, among others, consist of direct action-outcome interval estimation (Engbert et al., 2007), interval reproduction (Humphreys & Buehner, 2010), stimulus anticipation (Buehner & Humphreys, 2009), and psychophysical interval comparison (Nolden et al., 2012). However, the Libet clock task remains the most frequently used method to measure intentional binding. In a previous study, we found evidence that the binding effects in the Libet clock task and in interval estimation are based on different processes (Siebertz & Jansen, 2022). This emphasizes the question of the respective processes and indicates that they must be investigated separately.

Compared to other methods, such as interval estimation, the Libet clock task allows the investigation of binding of actions and outcomes separately by comparing the respective baseline and contingent conditions when the target event is the button press or the tone. Outcome binding is consistently shown to be the larger of the two effects. In the contingent condition, the action is reported to have occurred around 10–20 ms later than in the baseline condition, with the outcome around 40–90 ms earlier (e.g., Haggard et al., 2002; Schwarz & Weller, 2022; Siebertz & Jansen, 2022; Tonn et al., 2021). Both kinds of binding show good reliability in the Libet clock task, meaning they are measured with little random influences and are relatively consistent within individual participants at least for the duration of a single testing session. Despite this, they have repeatedly been found to be uncorrelated (Siebertz & Jansen, 2022; Tonn et al., 2021). This suggests that the two binding effects are based on distinct processes. As one of the reviewers pointed out, it is important to keep in mind that the Libet clock task measures action and outcome binding in separate trials. This might limit to what extend the two can be set in relation to each other.

Previous research looked at potential mechanisms behind the binding effects. In the search for the underlying mechanisms behind the binding phenomenon it is essential to keep in mind that multiple mechanisms could be at play simultaneously with varying degrees of impact depending on situational factors. As Moore and Obhi (2012) point out, slowing of an internal clock or recalibration of perceived stimulus onsets are possible explanations.

Wenke and Haggard (2009) used simultaneity judgments of cutaneous stimuli during action-outcome intervals. Comparing discrimination performance early and late in the action-outcome interval, they found support for an initial slowing and subsequent compensatory acceleration of an internal clock. They assumed this as a basis of temporal binding. A stronger deceleration should show in stronger binding of the action and outcome compared to a less pronounced deceleration. This would suggest a positive correlation between action and outcome binding. Previous studies support neither a positive nor a negative correlation (Siebertz & Jansen, 2022; Tonn et al., 2021).

Lush et al. (2019) investigated Bayesian cue combination as a mechanism underlying the binding effect. According to this explanation, the timing of action or outcome is used as auxiliary information when judging the timing of the other event. The resulting judgment is then the average timing of both events weighted by their temporal perceptual precision. Lush et al. found evidence for lower action binding the more precise the action is perceived temporally relative to the outcome but not for higher outcome binding. While perceptual recalibration, in general, could occur separately for actions and outcomes, Bayesian cue integration predicts a negative correlation between action and outcome binding. One increases while the other decreases when the perceptual precision ratio shifts in one or the other direction. As mentioned before, a correlation between action and outcome binding is not supported by previous studies.

If binding was ensured to measure the extent to which actions are processed as being self-elicited, the cognitive basis behind it could be an afterthought. However, the validity of binding as an indicator of agency has been questioned. Kirsch et al. (2019) found comparable binding when participants pressed the button themselves and when a motor moved the button and the participants' fingers. Omitting the role of intention, the terms causal binding or even more neutral temporal binding have been suggested (Buehner & Humphreys, 2009). Schwarz and Weller (2022) compared the usual contingent and baseline conditions of the Libet clock task described above with conditions where the button press or the tone was preceded or followed by a color change of the clock hand. Participants' attention was directed towards this event by the additional task of remembering the changed color. The result was a binding effect for actions and outcomes. The reported clock hand position during the button press or tone was systematically shifted towards the clock hand position during the color change. This shows that binding can occur without intention and even causation. This does not eliminate intention as a driving factor that can cause or influence binding possibly via attentional processes. It does also not eliminate the possibility that binding phenomena in perception are used as cues to infer agency. To facilitate the investigation of these questions it would be helpful to pinpoint the processes in which the binding phenomena are based perceptually and cognitively. Future research could then study binding's dependence on intention more precisely.

1.2. Working memory

We propose an alternative explanation that binding in the Libet clock task is rooted in interference processes in spatial working memory. At its core the Libet clock task is an effort of spatial working memory. A certain position or orientation of the clock hand indicated by the occurrence of an event, the action or the outcome, must be remembered until the response phase at the end of the trial. During encoding and the retention interval, this relevant spatial stimulus must be shielded from interference by all the irrelevant clock hand positions perceived before and after the target position as the clock hand constantly rotates. In the contingent conditions compared to the baseline, one additional position is highlighted by the respective irrelevant event. This position may be encoded involuntarily into spatial short-term memory. When the action is the target event, the later position during the outcome interferes retroactively. When the outcome is the target event the earlier position during the action interferes proactively. This interference shifts the reported target position towards the interfering position.

Spatial working memory is a modality-specific subset of processes that allow the transient retention and manipulation of spatial information. According to the multi-component working memory model (Baddeley, 2000), the central executive directs limited attentional resources toward specialized subsystems. Verbal memory content is stored in the phonological loop and visual information in the visuospatial sketchpad. At the same time, the episodic buffer constitutes an interface between the two and additional information from long-term memory. The visuospatial sketchpad can be further divided into distinct working memory processes for objects defined by their visual features and for spatial information (Hartley et al., 2001; Klauer & Zhao, 2004; Smith et al., 1995). These processes share central attentional resources (Vergauwe et al., 2009).

An essential function of spatial working memory is storing or maintaining spatial information. According to sensorimotor recruitment models, this happens through the continuous activation of systems involved in the perception of said information and attention-based rehearsal, meaning the sequential assignment of attentional resources to the maintained location (D'Esposito & Postle, 2015). Another function is updating obsolete information. To prevent interference, information no longer relevant is removed from limited short-term storage and replaced by new information (Ecker et al., 2014; Miyake et al., 2000). Both processes are necessary to accurately report the clock hand position in the Libet clock task. The position during the target event must be maintained and shielded from following perceived positions while previously encoded positions from the preceding or the current trial must be updated.

Theoretically, the clock hand position during the irrelevant event is especially suitable to interfere with the one during the target event after the latter has been encoded. Similarity-based interference by superposition leads to a mixing of activation in the representational space between two stimuli held in working memory (Oberauer & Lin, 2017). According to the interference model of Oberauer and Lin, this representational overlap requires sufficient similarity between the interfering stimuli relative to the precision with which they are perceived. The clock hand positions during action and outcome in the Libet clock task with the often-used period of around 2500 ms and an offset of 250 ms are just around 36° apart. The difference can be seen between the second and third frame of Fig. 1 in the Method section of Experiment 1.

Another possible point of interaction between the two positions is during the encoding and consolidation phase of transferring the perceived target position into short-term memory. Klauer and Zhao (2004) describe that a sensory activation pattern is stored in a rapidly fading short-term memory during perception. It must be consolidated into short-term memory using central resources to keep the representation available. The duration of this process is stated to be around 500 ms. During this time frame, into which the usual action-outcome delay in the Libet clock task falls, the representation of the target position might be especially susceptible to interference by another position highlighted by an attention-grabbing event. When the clock hand position during the action is the target, the irrelevant position during the outcome is perceived while the former is still being encoded into short-term memory. When the position during the outcome is the target, attentional resources might still be bound by the consolidation of the position during the action. This fits the observation, that binding tends to decrease when the action-outcome delay moves from around 250 ms towards

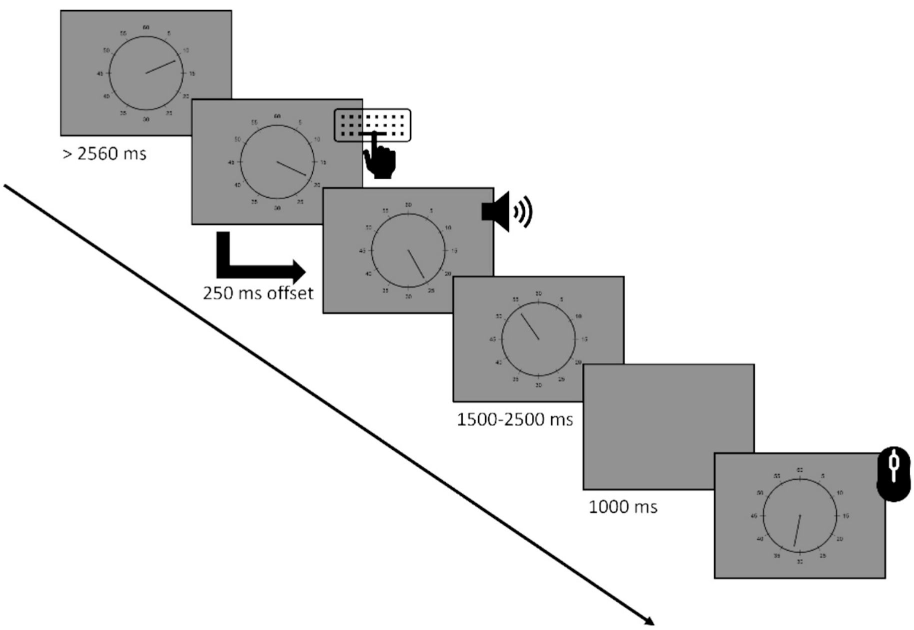

Fig. 1. The sequence of events in a contingent trial of the Libet clock task. Depending on the target event, participants reinstate the clock hand position during the key press or the tone using the mouse wheel at the end of the trial.

500 ms and beyond (Ruess et al., 2017).

A similar argument can be drawn from the interference model of Oberauer and Lin (2017) who assign two functions to the so-called focus of attention in working memory. One is to hold a high-precision representation of a stimulus, and the other is to shield this representation from interference. They argue that the first function needs the focus of attention in working memory and in perception to coincide during encoding. Otherwise, the high resolution of the representation cannot be recovered retrospectively. A disparity of attentional focusses through the irrelevant event and associated clock hand position during encoding could be a potential mechanism behind the temporal binding effect. This also connects to the findings of Schwarz and Weller (2022) described above, that attention-grabbing events other than the action or the outcome can cause binding in the Libet clock task.

Two properties of the binding phenomenon could be explained by a working memory interference account. First, different processes behind action and outcome binding (e.g., updating and maintaining or pro- and retroactive interference) could cause the difference in magnitude. Action binding has consistently been shown to be the smaller of the two effects. Second, if the two kinds of binding are based in separate processes, the missing correlation between the two would be accounted for.

Interference in spatial working memory could explain the binding effect in the Libet clock task. To investigate this, we report data from two experiments, in which we use correlational (Experiment 1) and experimental (Experiment 2) approaches to look for a relationship between spatial working memory and binding in the Libet clock task.

2. Experiment 1

The preregistration can be found on OSF (https://osf.io/6prgd). Data, R analysis code and PsychoPy experiment code for Experiment 1 can be found in its OSF repository (https://doi.org/10.17605/OSF.IO/2NM34).

Experiment 1 aimed to investigate whether binding effects in the Libet clock task could result from processes in spatial working memory. To this end, we tested the following hypotheses. First, we expect to replicate the typical binding effects in the Libet clock task.

H1: Action and outcome binding occur in the Libet-clock task.

If binding effects result from processes in spatial working memory, they should also occur when the remembered positions associated with the action and the outcome are not sequenced on a linear, perfectly time-dependent dimension, such as the path of the rotating clock hand.

H2: Action and outcome binding occur in a spatially random task.

If we find binding effects in a spatially random task, we must investigate whether these effects have the same basis as those in the Libet clock task. If they have, they should correlate across participants. Because action and outcome binding are not correlated, we test this separately for the two binding effects.

H3: Action binding in the Libet-clock task correlates positively with action binding in the new, spatially random task.

H4: Outcome binding in the Libet-clock task correlates positively with outcome binding in a spatially random task.

If interference between the clock hand positions in spatial working memory leads to the binding effects, having better spatial working memory, especially being better at maintaining and updating, should reduce binding. The clock hand positions during action and outcome could be kept from interfering with each other more effectively.

H5: Binding in both tasks correlates negatively with spatial working memory performance.

To eliminate more general attentional processes as the cause of binding effects, we expect a correlation with spatial working memory but not with the working memory of other modalities such as visual working memory.

H6: Binding in both tasks does not correlate with visual working memory performance.

2.1. Method

2.1.1. Design

Experiment 1 uses a correlational design to investigate the theory that temporal binding in the Libet clock task originates from spatial memory processes. All participants complete four tasks: The Libet clock task, a 2-back task with spatial stimuli, a 2-back task with visual stimuli, and a task resembling the Libet clock task. The latter replaces the clock face with a circle in which dots appear simultaneously with the action and outcome and whose positions must later be recalled.

2.1.2. Participants

To achieve a sufficient sample size, we used adaptive sampling with hypothesis-relevant Bayes factors below 1/3 or above 3 as a stopping criterion. To set a viable endpoint for sampling if Bayes factors fail to indicate conclusive evidence, we conducted a power analysis in G*Power (version 3.1.9.7, Faul et al., 2009). For our correlational hypotheses, to achieve a power of at least 95 % to detect differences from zero for correlations of 0.3 or −0.3 at a significance level of 5 % in a one-sided test requires 111 participants. We arbitrarily decided that the effect size of 0.3 would be the smallest effect of interest due to general conventions. To account for exclusions based on our previous experience, we increased this sample size by 10 % resulting in 122 participants. As preregistered, we first inspected the data at 50 participants and then after every 10 new participants.

Because some Bayes factors failed to indicate conclusive evidence at all steps, we sampled all 122 participants. Of those, seven reported current neurological or psychiatric disorders. One reported impaired and uncorrected vision and another impaired and uncorrected hearing. Due to these preregistered criteria, nine participants were excluded from further analyses. Following the preregistered work-flow, nine other participants were excluded for having judgment errors 3 *SD* above or below the condition sample mean in at least one condition of the two binding tasks. This left a final sample of 104 participants, 50 reported their gender as female, 54 as male and none as non-binary. The mean age was 22.9 years (SD = 3.4). All participants were healthy, young sports science students at our faculty, gave written informed consent, and received course credit for participation.

The study was executed in accordance with the declaration of Helsinki for the guidelines of ethical considerations. Ethical approval for this study was not required in accordance with the conditions outlined by the German Research Society (DFG) where research that carries no additional risk beyond daily activities does not require a research ethics board's approval. We communicated all considerations necessary to assess the question of ethical legitimacy of the study.

2.1.3. Material

Participants worked through the tasks and the demographic questionnaire in PsychoPy (version 2023.1.1, Peirce et al., 2019) on a Lenovo Thinkpad T15 wearing noise-cancelling headphones. Responses were given using an external keyboard and a mouse with a mouse wheel. The demographic questionnaire queried age, gender, and the exclusion criteria of current psychiatric or neurological disorders and uncorrected, impaired vision and hearing.

2.1.3.1. Libet clock task. The Libet clock task is identical to one of our previous studies (Siebertz & Jansen, 2022, Experiment 1) in which we adapted it from Kirsch et al. (2019), and goes back to the study of Haggard et al. (2002). Participants observed a clock face with a clock hand rotating at 2560 ms per revolution after starting at a random position. In the contingent conditions the participants' action, a key press, was followed by an outcome, a sine tone (1000 Hz, 100 ms duration) with a delay of 250 ms. Fig. 1 shows the sequence of events for a contingent trial. Their task was to press the space bar at a time of their chosing after the first revolution which triggered the tone. Beforehand, they were instructed to remember the clock hand position during their key press or during the tone. The clock hand continued to rotate after the target event for 1500 to 2500 ms before the clock disappeared for 1000 ms. The clock hand reappeared randomly positioned 45° to 60° before or after its position during the target event. Participants used the mouse wheel to turn the clock hand to the position they remembered.

In the two baseline conditions, participants either performed a key press that did not trigger a tone or remained passive, and the tone was played automatically 2810 ms to 4810 ms after the start of the trial. After the clock disappeared for 1000 ms, participants reinstated the clock hand position during the target event. The baseline condition with the tone as the target event mirrored its contingent counterpart with 2560 ms for the first revolution, 250 ms action-outcome delay and 2000 ms jitter. The jitter is introduced to make the baseline condition more similar to the contingent one where participants press the button at variable times after trial onset. The order of the four conditions was counterbalanced across participants. There were 28 trials per condition, and participants did three unanalyzed training trials in the same condition order before the actual task started.

Before calculating binding scores, trials were excluded if the judgment error was at least 3 SD above or below the condition mean of the indivdual participant. We then calculated temporal binding scores for the action as the mean judgment errors of the contingent minus the baseline condition. The binding score for the outcome is calculated the opposite way, leading to higher, more positive scores to indicate more binding for action and outcome. For comparability to the spatially random binding task described below, we report binding magnitude in proportion to the actual spatial distance between the clock hand positions during the two events instead of the usual time shift metric. The 250 ms action-outcome delay corresponding to 35.16° of rotation equals the proportion of the actual temporal distance.

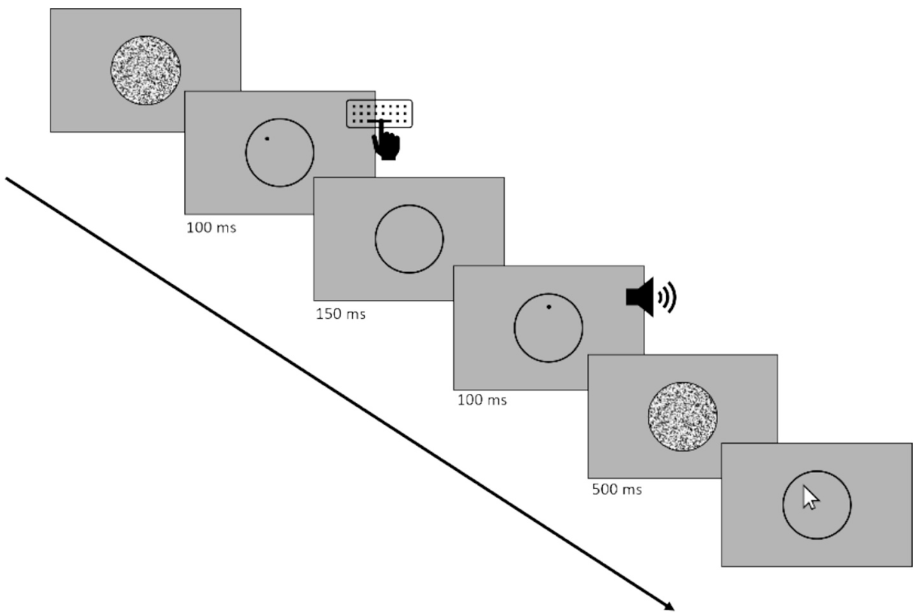

Fig. 2. The sequence of events in a contingent trial of the spatially random task.

We calculated reliability scores for action and outcome binding by calculating the split-half reliability. We did so by following Parsons et al. (2019) suggestions. First, for each participant the trials per condition were split randomly into two halves and from the halves two action and two outcome binding scores were calculated. Then the two scores for action binding were correlated across the participants as were the two scores for outcome binding. To account for the fact that each score was calculated from only half of the trials, we used the Spearman-Brown correction formula on the correlations. This was repeated 5000 times with different permutations of splitting the trials in halves. Averaged across the permutations, Spearman-Brown attenuation corrrected split-half reliability as the correlation between binding scores for the two halves was $r_{split-half}$ =.86 (95 % CI = [.81, 0.90]) for action binding and $r_{split-half}$ = 0.98 (95 % CI = [.98, 0.99]) for outcome binding. These values are satisfactory. Reliability scores represent the proportion of variance in a measurement that is not caused by random measurement error. Random measurement error attenuates the observable correlation between two variables. Reliability scores should generally be kept in mind as a continuous indicator but as a rule of thumb we would interpret results of measurements with reliabilities below r = 0.6 with caution on a group level.

2.1.3.2. Spatially random binding task. The Libet clock task confounds the temporal and spatial distance between action and outcome as they correlate perfectly. Through the constant rotation speed, a certain temporal delay between action and outcome corresponds to a fixed rotation angle in the clockwise direction. To disentangle the temporal and the spatial delay, we designed a task miming the Libet clock task. Instead of positions of a clock hand, participants observe pairs of consecutively appearing dots in a circle.

The task contained the same four conditions as the Libet clock task, two contingent and two baseline conditions in which participants had to remember the position of a dot appearing simultaneously with their key press or a thereby triggered tone. Fig. 2 shows an exemplary trial. Participants saw a circle in which a randomly changing greyscale visual noise mask flickered at 60 Hz. The mask's purpose was to introduce visual noise comparable to the motion of the clock hand in the Libet clock task. In the contingent conditions, participants pressed the space bar at a time of their choice, immediately letting a dot appear at a random position inside the circle for 100 ms. After a delay of another 150 ms a sine tone (1000 Hz, 100 ms duration) was played leading to 250 ms of action-outcome delay like in the Libet clock task. Simultaneously with the tone a second dot appeared for 100 ms. The distance between the two dots was equal to the distance between the position of the tip of the clock hand during the key press and the tone in the Libet clock task. The direction in which the second dot appeared relative to the first one was random. Immediately after the disappearance of the second dot, the visual noise mask was displayed for 500 ms to prevent afterimages from highlighting the dot positions and introducing visual noise. Then, the mouse cursor appeared, and participants clicked where they recalled the position of the dot appearing with the current target event.

In the baseline condition with the key press as the target event, no second dot appeared, and no tone was played. Participants did not press the space bar in the baseline condition with the target event tone. The first dot appeared automatically after a random delay of 750 ms to 1750 ms after trial onset. There were 28 trials and three unanalyzed training trials per condition. Participants did the latter before the actual task in the same condition order. Condition order was counterbalanced across participants.

The judgment error was calculated by orthogonally projecting the position that participants clicked at onto the line going through both actual dot positions. The judgment error was the distance of this projection to the actual target dot position. The 28 pairs of dot positions were generated randomly once at the start of the task and then reused in random order for all conditions to control for the positioning of the dots relative to the bordering circle. Judgment errors in the baseline conditions were calculated in the same way as in the contingent conditions using the same coordinates of a given pair of dots even though only one dot is presented per trial. This allows for a direct comparison of the shift of the perceived target position in the specific direction of the other dot as it is presented in the contingent condition. As with the Libet clock task, the difference between the judgment errors in the baseline and contingent conditions for each target event yielded the action and outcome binding scores. Higher, more positive values would indicate higher binding for both, which would mean that participants remembered the dot position systematically shifted towards the other dot in the contingent compared to the baseline condition. As with the Libet clock task, trials were excluded if the judgment error was at least 3 SD

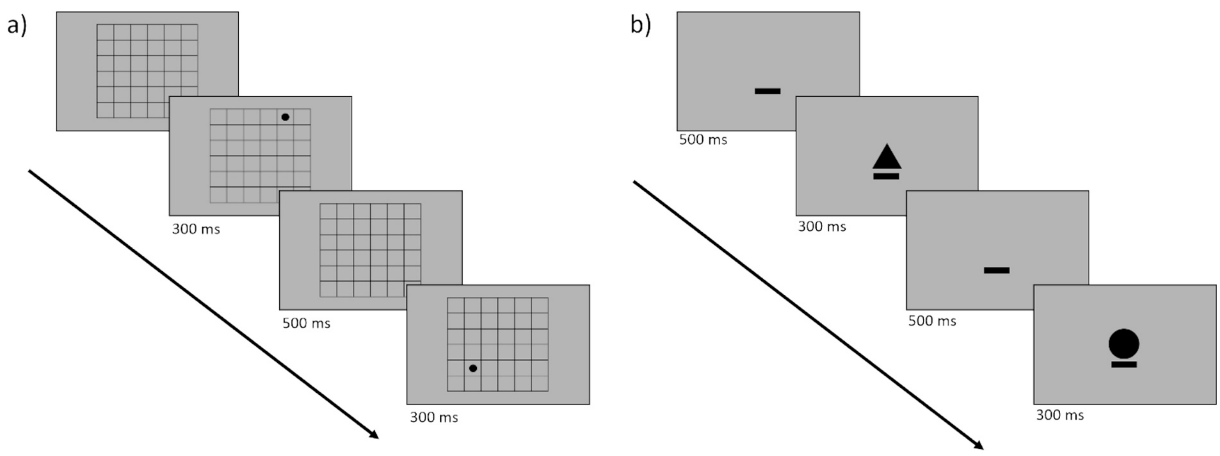

Fig. 3. The sequence of events of the spatial (a) and visual (b) 2-back tasks.

above or below the condition mean of the individual participant.

Using the permutation-based split-half approach for calculating reliability described for the Libet clock task, the average spearman–brown attenuation corrected split-half reliability was $r_{\text{split-half}} = .65$ (95 % CI = [.54, 0.74]) for action binding and $r_{\text{split-half}} = 0.74$ (95 % CI = [.66, 0.81]) for outcome binding.

2.1.3.3. Spatial 2-Back task. To measure spatial working memory performance, we used a 2-back task in which participants had to remember the positions of dots appearing sequentially in a 6-by-6 grid (see Fig. 3a). The dots stayed visible for 300 ms with an ISI of 500 ms. Participants were instructed to press the space bar when a dot appeared in the same field as the next-to-last dot. Per block participants saw a sequence of 25 dots, five of which were 2-back repetitions. Participants started with two training blocks with a stimulus duration of 500 ms. If participants reacted in these blocks, the sequence stopped, and feedback was displayed until the participant continued the sequence via key press. Two further training blocks followed, which were identical to the following six actual task blocks with a stimulus duration of 300 ms and no feedback about the correctness of their response given. Training blocks were not analyzed. Participants could react to a stimulus while it was visible and during the ISI until the next stimulus was presented. To provide them with feedback that their key press was registered in time, the outer border of the grid changed color to a lighter grey for 100 ms. Sequences were equal and presented in the same order for all participants to eliminate sequence order as a source of variance.

For our analyses, we calculated *d*' as an informative index for working memory performance (Haatveit et al., 2010). Across the six blocks, participants saw 150 stimuli of which 30 were targets. We calculated the *z*-transformed hit rate and false alarm rate for each participant and subtracted the latter from the former to derive d'. Hit rates of 1 and false alarm rates of 0 were replaced with $1 - 1/(2^* n_{targets})$ and $1/(2^* n_{non-targets})$ respectively. Again, we used the permutation-based split-half approach to calculate reliability scores for the *d*'-index. Splitting targets and non-targets into halves separately using 5000 permutations and calculating d' for each half, the average spearman–brown corrected inter-half correlation was $r_{\text{split-half}} = .67$ (95 % CI = [.58, 0.75]).

2.1.3.4. Visual 2-Back task. The 2-back task measuring visual working memory was identical to the one for spatial working memory apart from the stimulus material. Instead of dots in a grid, participants had to remember geometric shapes presented at the center of the screen (see Fig. 3b). There were 5 shapes: A circle, a square, a triangle, a star, and a pentagon. To provide feedback for key press registration, a black bar was displayed below the position where the shapes appeared. Like the grid border in the spatial 2-back task, it changed color to a lighter grey for 100 ms upon key press.

We calculated the *d*'-index as described for the spatial working memory task. The same permutation-based split-half approach resulted in an average spearman–brown corrected inter-half correlation of $r_{\text{split-half}} = .57$ (95 % CI = [.45, 0.69]) for *d*'.

2.1.4. Procedure

First, participants read the study information and gave written informed consent. They started the testing session with the spatial and visual 2-back tasks. The task order was counterbalanced across participants. After that, participants did the two binding tasks. Again, task order was counterbalanced across participants. Finally, they answered the demographic questionnaire. Participants were encouraged to ask questions at any time and proceed through the testing session at their own pace. Testing sessions typically lasted for one to one and a half hours.

2.1.5. Statistical analysis

We used Bayes factors for population-level inference to investigate the absence of an effect in Hypothesis 6 and to use adaptive sampling. For Hypotheses 1 and 2, expecting binding above zero in both binding tasks, we calculated Bayes factors by comparing the null model (H0) with the alternative (H1). H0 was modeled as a normal distribution with a mean of zero and a *SD* equal to the observed *SD* of the binding in question. H1 was modeled as a half-normal distribution with mode 0 and *SD* equal to the expected effect. Using data from Siebertz and Jansen (2022), this was 13 ms for action binding and 74 ms for outcome binding. These values correspond to 5.2 % and 29.6 % of the actual delay of 250 ms. Therefore, in proportion of the actual event position distance, the *SD* of the H1 model was set to 0.052 and 0.296 respectively.

For Hypotheses 3 and 4, we calculated Bayes factors using the Fisher's *z*-transformation of the Pearson correlation coefficient. The H0-model was a normal distribution with mean zero and $SD = 1/\sqrt{df - 1}$. The alternative H1 was modeled as a half-normal distribution with mode 0 and *SD* equal to the expected correlation effect size. Following Dienes (2019), this was set to half of a rough sensible maximum. Using reliability estimates from Siebertz and Jansen (2022) and accounting for reliability-related correlation attenuation, we calculated the maximum observable correlation of binding between the two tasks using the formula: $r_{expected} = r_{theoretical}{}^* \sqrt{rel(x)^* rel(y)}$, with r_{expected} being the observable correlation attenuated by imperfect reliability, $r_{\text{theoretical}}$ being the true correlation between the variables and *rel*(x) and *rel*(y) being the reliabilities of the variables. Assuming that binding in the two tasks is caused by the same mechanisms, the $r_{\text{theoretical}}$ was set to 1. Reliability was assumed to be 0.84 for action binding and 0.96 for outcome binding. Using the formula above, these were simultaneously the maximum expected correlations. Accordingly, the *SD* of the half-normal H1 model for the between-task correlation of action binding was set to $Fisher'sz(.84)/2 = 0.612$ and the *SD* for outcome binding to $Fisher'sz(.96)/2 = 0.973$.

Hypotheses 5 and 6 regard the correlation between working memory performance and binding. The models for H0 and H1 were constructed as described for Hypotheses 3 and 4. To investigate even small correlations, we compared the null model that assumes the correlation is zero with an alternative that assumes the correlation is −0.3. As it is computationally equivalent, the observed correlation's sign was flipped and the H1 half-normal was modeled with an *SD* of $Fisher'sz(.30) = 0.310$.

We calculated robustness regions to check the dependence of our results on our prior specifications. For mean comparisons, Bayes factors were calculated for prior effect sizes from 0.001 to five times the actual specified effect size. The interval in which the same qualitative result regarding the Bayes factor (<1/3, >3 or between) held true is reported as the robustness region. Bayes factors were calculated for SDs corresponding to $r = 0.001$ ($SD = \text{Fisher's } z(r) = 0.001$) to $r = 0.999$ ($SD = \text{Fisher's } z(r) = 3.800$) for correlations.

2.2. Results

2.2.1. Action and outcome binding in the Libet clock task

Fig. 4a shows all binding scores. Average action binding in the Libet clock task was 7.1 % of the event position distance ($SD = 13.4$ %, corresponding to a temporal shift of 17.8 ms, $SD = 33.5$ ms). The Bayes factor shows overwhelming evidence for action binding being above zero ($t(103) = 5.44$, $p < 0.001$, $B_{HN(0,\ 0.052)} \gg 100$, $RR_{B>3}$ [0.003, >0.26]). Mean outcome binding was 36.3 % of event position distance ($SD = 31.0$ %, 90.8 ms and $SD = 33.5$ ms in temporal shift). Again, the Bayes factor shows overwhelming evidence for outcome binding above zero ($t(103) = 11.92$, $p < 0.001$, $B_{HN(0,\ 0.296)} \gg 100$, $RR_{B>3}$ [0.003, >1.48]). These results replicate the typical binding effect in the Libet clock task an confirm Hypothesis 1. Robustness regions indicate virtually no dependence on prior effect size. The correlation between action and outcome binding in the Libet clock task was $r = 0.03$ (95 % CI = [−0.17, 0.22]).

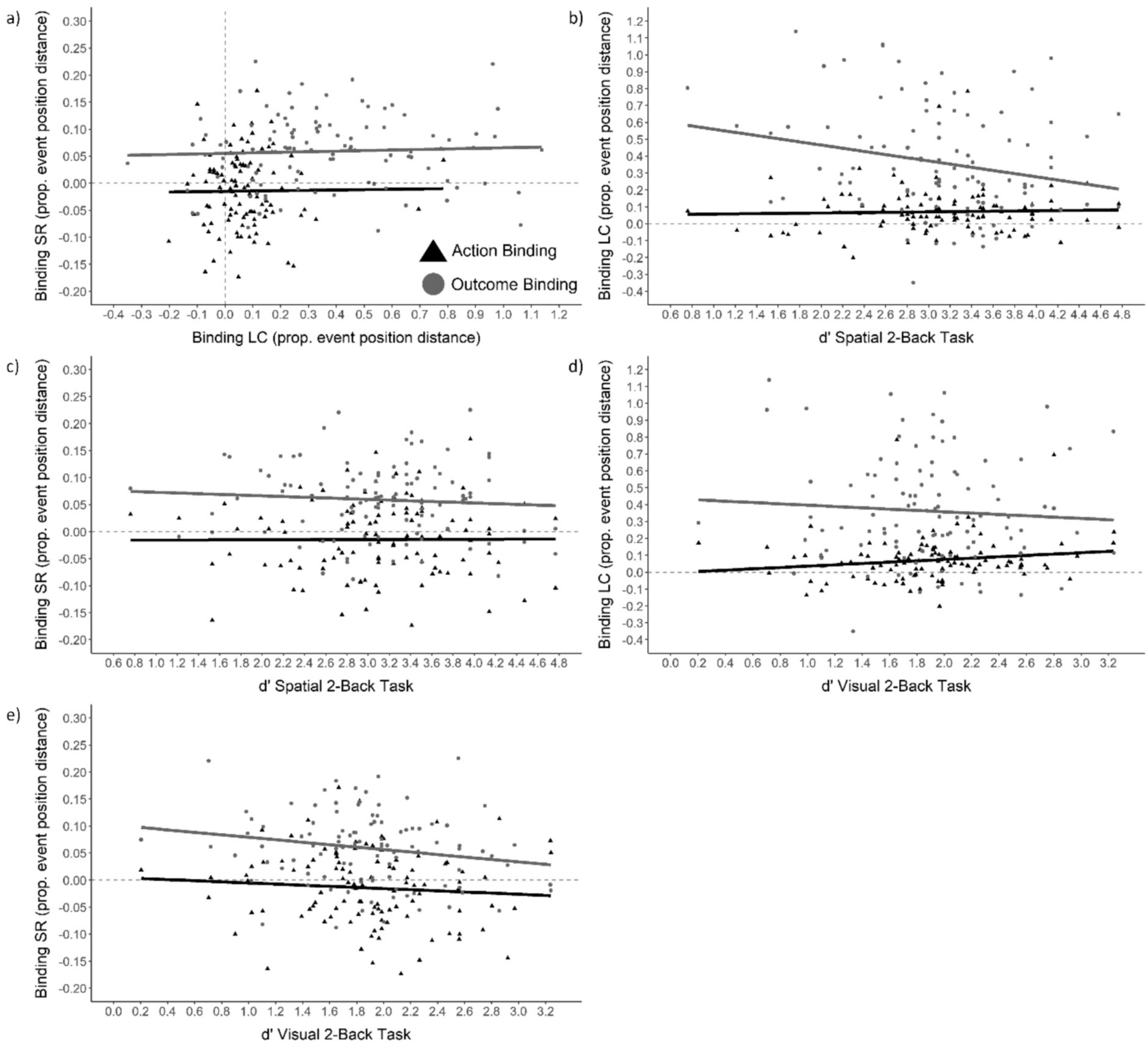

Fig. 4. Scatter plots showing the relationships between binding in the Libet clock task and the spatially random binding task (a), spatial working memory performance and binding in the Libet clock task (b) and the spatially random binding task (c), and visual working memory performance and binding in the Libet clock task (d) and the spatially random binding task (e). Data regarding action binding is shown in black triangles, and regarding outcome binding is in grey circles. Regression lines show the results of simple linear regression.

2.2.2. Action and outcome binding in the spatially random task

Mean action binding in the spatially random task was −1.4 % of event position distance (*SD* = 7.9 %). The Bayes factor shows strong evidence for action binding not being above zero ($t(103)$ = -2.14, $p = 0.983$, $B_{HN(0,\ 0.052)} = 0.04$, $RR_{B<1/3}$ [0.006, >0.26]). On average, outcome binding was 5.9 % of event position distance (*SD* = 6.6 %). The Bayes factor indicates overwhelming evidence for outcome binding above zero ($t(103) = 9.08$, $p < 0.001$, $B_{HN(0,\ 0.296)} \gg 100$, $RR_{B<3}$ [0.003, >1.48]). This confirms the Hypothesis that binding effects occur in a spatially random task only for outcome binding and rejects it for action binding. Large robustness regions indicate virtually no dependence on prior effect size. The correlation between action and outcome binding in the spatially random task was r = 0.07 (95 % CI = [−0.12, 0.26]).

2.2.3. Relationship of binding between the binding tasks

Fig. 4a shows the relationship of binding scores between the two binding tasks. The correlation between action binding scores was 0.01 (95 % CI = [−0.18, 0.20]). The Bayes factor indicates moderate evidence for the absence of a positive correlation ($t(102) = 0.12$, $p = 0.451$, $B_{HN(0,\ 0.611)} = 0.18$, $RR_{B<1/3}$ [0.316, 1.323]). These robustness region limits correspond to assumed rough maximum correlations of 0.56 to 0.99. This result rejects Hypothesis 3. Compared to the assumed rough maximum correlation of 0.84, the robustness region indicates little dependence on prior effect size on the lower end.

The correlation between outcome binding scores was 0.05 (95 % CI = [−0.15, 0.24]). Again, the Bayes factor indicates moderate evidence for the absence of a positive correlation ($t(102) = 0.48$, $p = 0.314$, $B_{HN(0,\ 0.973)} = 0.16$, $RR_{B<1/3}$ [0.444, 1.323]). The robustness region limits correspond to rough maximum correlations of 0.71 to 0.99. Hypothesis 4 is rejected. Compared to the assumed rough maximum effect size of 0.96, the robustness region indicates some dependence on the prior effect size on the lower end.

2.2.4. Relationship of binding and working memory performance

The average *d*' in the spatial 2-back task was 3.1 (*SD* = 0.7, range = [0.8, 4.8]). Average *d*' in the visual 2-back task was 1.9 (*SD* = 0.5, range = [0.2, 3.2]). The correlation between spatial and visual working memory performance was $r = 0.31$ (95 % CI = [.13, 0.48]). Table 1 shows the results of our analyses regarding the correlation between binding and spatial and visual working memory performance.

2.2.4.1. Relationship between binding and spatial working memory performance. The correlation between action binding in the Libet clock task and spatial working memory performance was 0.04 (95 % CI = [-0.23, 0.16], black triangles in Fig. 4b). The Bayes factor indicates moderate evidence against the correlation being below zero ($t(102) = 0.36$, $p = 0.641$, $B_{HN(0,\ 0.310)} = 0.24$, $RR_{B<1/3}$ [0.107, 1.323]). Robustness region limits correspond to rough maximum correlations of −0.21 to −0.99.

The correlation between outcome binding in the Libet clock task and spatial working memory performance was −0.23 (95 % CI = [−0.40, −0.03], grey circles in Fig. 4b). The Bayes factor shows moderate evidence in favor of the correlation being below zero ($t(102) = 2.34$, $p = 0.011$, $B_{HN(0,-0.310)} = 6.76$, $RR_{B>3}$ [0.030, 0.454]). Robustness region limits correspond to rough maximum correlations of −0.06 to −0.72.

The correlation between action binding in the spatially random task and spatial working memory performance was 0.01 (95 % CI = [−0.19, 0.20], black triangles in Fig. 4c). The Bayes factor shows moderate evidence against the correlation being below zero ($t(102) = 0.05$, $p = 0.520$, $B_{HN(0,\ 0.310)} = 0.29$, $RR_{B<1/3}$ [0.138, 1.323]). Robustness region limits correspond to rough maximum correlations of −0.27 to −0.99.

The correlation between outcome binding in the spatially random task and spatial working memory performance was −0.07 (95 % CI = [−0.26, 0.12], grey circles in Fig. 4c). The Bayes factor shows inconclusive evidence regarding the correlation being below zero ($t(102)$ = -0.75, $p = 0.228$, $B_{HN(0,\ 0.310)} = 0.60$, $RR_{1/3<B<3}$ [0.000, 0.295]). Robustness region limits correspond to rough maximum correlations of 0.00 to −0.53. Prior effect sizes below −0.53 lead to the Bayes factor showing moderate evidence against the correlation being below zero.

2.2.4.2. Relationship between binding and visual working memory performance. Action binding in the Libet clock task correlated with

Table 1
Correlations between binding scores in the Libet clock and spatially random tasks with spatial and visual working memory performance in the 2-back task. Bayes factors over 1 favor a negative correlation versus no correlation.

			95 % CI					RR		
	Binding type	r	LL	UL	$t(102)$	$p_{r<0}$	$B_{HN(0,\ 0.310)}$	Qual. res.	LL	UL
SWM	Action LC	0.04	−0.23	0.16	0.36	0.641	0.24	$B < 1/3$	0.107	1.323
	Outcome LC	−0.23	−0.40	−0.03	−2.34	0.011	6.76	$B > 1/3$	0.030	0.454
	Action SR	0.01	−0.19	0.20	0.05	0.520	0.29	$B < 1/3$	0.138	1.323
	Outcome SR	−0.07	−0.26	0.12	−0.75	0.228	0.60	$1/3 < B < 3$	0.000	0.295
VWM	Action LC	0.16	−0.03	0.34	1.65	0.949	0.12	$B < 1/3$	0.055	1.323
	Outcome LC	−0.07	−0.26	0.13	−0.70	0.244	0.57	$1/3 < B < 3$	0.000	0.275
	Action SR	0.08	−0.11	0.27	−0.83	0.203	0.66	$1/3 < B < 3$	0.000	0.324
	Outcome SR	−0.19	−0.37	0.01	−1.92	0.029	3.01	$B > 1/3$	0.050	0.155

Note. RR: Robustness region, SWM: Spatial working memory performance, VWM: Visual working memory performance, LC: Libet clock task, SR: spatially random task, Qual. res. = qualitative result.

visual working memory performance at $r = 0.16$ (95 % CI = [−0.03, 0.34], black triangles in Fig. 4d). This yields substantial evidence against the correlation being below zero ($t(102) = 1.65$, $p = 0.949$, $B_{HN(0,\ 0.310)} = 0.12$, $RR_{B<1/3}$ [0.055, 1.323]). The robustness region indicates the same qualitative results for prior correlation effect sizes between 0.11 and 0.99.

Outcome binding in the Libet clock task showed a correlation of $r = -0.07$ (95 % CI = [−0.26, 0.13], grey circles in Fig. 4d) with visual working memory performance. Evidence for the correlation being below zero is inconclusive ($t(102) = -0.70$, $p = 0.244$, $B_{HN(0,\ 0.310)} = 0.57$, $RR_{1/3<B<3}$ [0.000, 0.275]). The robustness region corresponds to prior correlation effect sizes between 0.00 and 0.50, with prior effect sizes above 0.50 leading to conclusive evidence against the correlation being below zero.

Action binding in the spatially random task was correlated with visual working memory performance with an r of 0.08 (95 % CI = [−0.11, 0.27], black triangles in Fig. 4e). Evidence for a correlation below zero is inconclusive ($t(102) = -0.83$, $p = 0.203$, $B_{HN(0,\ 0.310)} = 0.66$, $RR_{1/3<B<3}$ [0.000, 0.324]). Robustness region limits correspond to prior correlation effect sizes of 0.00 and 0.57 with larger prior effect sizes leading to evidence against a correlation below zero.

Outcome binding in the spatially random task correlates with visual working memory performance at $r = -0.19$ (95 % CI = [−0.37 0.01], grey circles in Fig. 4e). This results in moderate evidence in favor of the correlation being below zero ($t(102) = -1.92$, $p = 0.029$, $B_{HN(0,\ 0.310)} = 3.01$, $RR_{B>1/3}$ [0.050, 0.155]). The robustness region limits translate to correlation prior effect sizes of 0.10 and 0.30, with prior effect sizes below and above leading to inconclusive evidence.

2.3. Preliminary discussion

Experiment 1 aimed to investigate whether binding in the Libet clock task is rooted in spatial working memory processes. We compared binding in the Libet clock task with binding in a similar task where action- and outcome-related event locations are not in a linear and predictable spatial relationship. We also investigated whether binding in the two tasks is reduced with better spatial working memory performance but not with visual working memory performance.

We were able to replicate the typical binding effects in the Libet clock task confirming Hypothesis 1. Binding magnitudes are in the usual magnitude and ratio with mean action binding of 18 ms and mean outcome binding of 91 ms. Reliability is good for both and better for outcome binding confirming previous results (Siebertz & Jansen, 2022).

Average action binding in the spatially random task is not significantly above zero and does not correlate with action binding in the Libet clock task. Outcome binding in the spatially random task is significantly above zero and of similar magnitude as action binding in the Libet clock task. However, despite finding a systematic shift in the reported outcome-related position towards the action-related one, it is not correlated to outcome binding in the Libet clock task. This confirms Hypothesis 2 that binding occurs in the spatially random task, but only for outcome binding. Hypotheses 3 and 4 regarding correlations between the two binding tasks are rejected. Reliability is lower than in the Libet clock task at 0.65 and 0.74 for action and outcome binding respectively. This is especially interesting for action binding. Despite the sample average being close to zero, individuals consistently biased their judgments.

Even though we found evidence for binding in the spatially random task, it seems to result from different processes than binding in the Libet clock task. Binding in the Libet clock task may depend on the specific judgment modality of the clock face. Another explanation could be that the visual noise introduced in the spatially random task did not lead to similar perceptual uncertainty compared to the movement of the clock hand. This could especially be the case, as there was no visual noise between action and outcome whereas the clock hand moves between action and outcome in the Libet clock task. This explanation fits the higher variance of action and outcome binding in the Libet clock task compared to the spatially random task. The spatially random task is a first attempt at trying to untangle the continuous visual and temporal dimensions in the Libet clock task. A central challenge with this is creating a task design that contains a comparable amount of perceptual uncertainty to the Libet clock task. The parameters we used are the presentation duration of the dots and the visual noise mask. Future research should investigate the influence of different levels of uncertainty on binding scores. One reviewer correctly pointed out another difference between the two binding tasks. In the active conditions of the spatially random task the first dot position is a consequence of the button press while the clock hand position in the Libet clock task is simply co-occurring with the button press. Both tasks share that a position is associated with an action, but the spatially random task introduces a causal component to this relationship. This could cause different processes to occur during the sequence of events and be an explanation for the lack of correlation between the two binding tasks. Another point raised by one of the reviewers is that the metric of the spatially random task, that is the projection of the reported position on the imagined line between the two dots, might lack validity if the deviation perpendicular to that line is large compared to the deviation along the line. However, further analysis showed this not to be the case.

To measure spatial and visual working memory performance we used 2-back tasks. We expected binding to correlate negatively with spatial working memory performance (Hypothesis 5). This could only be confirmed for outcome binding in the Libet clock task. Hypothesis 5 is rejected for action binding in both tasks, and evidence for the relationship between outcome binding in the spatially random task and spatial working memory performance is inconclusive. Hypothesis 6 states that binding does not correlate with visual working memory performance. This could be shown for action binding in the Libet clock task. Evidence remains inconclusive for outcome binding in the Libet clock task and action binding in the spatially random task. Outcome binding in the latter does seem to decline with increasing visual working memory performance.

As action and outcome binding in the Libet clock task are uncorrelated and, therefore, probably stem from different processes, it is not surprising that only one binding type correlates with spatial working memory performance. If one would speculate, our results might be explained by outcome binding in the Libet clock task stemming from interference in spatial working memory. Better spatial working memory abilities would then lead to a more effective shielding of the relevant outcome-related position from the preceding action-related position via updating of the latter. Arguably, the 2-back task consists of a higher updating to maintaining ratio than for example a 3-back task (Ren et al., 2023). A spatial working memory task relying more on maintaining relative to updating might show a relationship to action binding. As one reviewer pointed out, it is important to remember that we calculated multiple correlations and that the one between spatial working memory performance and outcome binding in the Libet clock task could always just be due to chance. Additionally, the correlational design does not allow for causal interpretations. If the correlation should indeed not be due to chance, one explanation for the correlation could be that binding stems from spatial working memory interference but it could also be caused by a common dependence on a third variable. Limiting the scope of our results is that we chose our statistical models to represent a certain magnitude of correlation which is $r = 0.3$. The half-normal distribution we used broadens this to a range of magnitudes with a greater weight on smaller values (Dienes, 2019). This seems reasonable but if one would expect larger correlations the models should be adapted appropriately. Another point is that, even though we adapted the 2-back memory tasks as they are often used in research, their reliability is rather low. This attenuates any correlation we might observe and highlights the need to inspect the reliabilities of our tasks.

Our results on visual working memory performance do not allow us to completely exclude more general attentional mechanisms as an alternative explanation. As the evidence remains inconclusive, especially for outcome binding in the Libet clock task, we cannot rule out a relationship with visual working memory processes comparable to that with spatial working memory.

3. Experiment 2

In Experiment 1 we took a correlational approach. To test the theory that temporal binding results from processes in spatial memory more thoroughly, we expand this with an experimental approach in Experiment 2. Our theory postulates that binding in the Libet clock task occurs because the clock hand positions during action and outcome interfere with each other as spatial memory content. Accurate recall of the target clock hand position should rely on the use of limited attentional resources to shield it from interference. Based on sensorimotor recruitment models such as attention-based rehearsal (D'Esposito & Postle, 2015) we assume that loading spatial memory binds those resources. This should impair the ability to prevent interference by the irrelevant clock hand position and increase the binding effect. If temporal binding in the Libet clock task originates from spatial memory processes and not more general attentional processes, loading other memory modalities like visual memory for objects should not increase binding. These considerations lead us to the following hypotheses. Hypothesis 1 is visualized in Fig. 5a and Hypothesis 2 in Fig. 5b.

H1: In the Libet clock task embedded in the spatial memory task, temporal binding increases the more items have to be remembered.

H2: In the Libet clock task embedded in the visual object-related memory task, the number of items does not increase temporal binding.

The preregistration can be found on OSF (https://osf.io/vxagf). Data, R analysis code and PsychoPy experiment code for Experiment 2 can be found in its OSF repository (https://doi.org/10.17605/OSF.IO/5FK7Z).

3.1. Method

3.1.1. Design

Experiment 2 utilized an experimental within-participant design with four factors. Participants did the Libet clock task embedded in the retention interval of a memory task. We manipulated the type of memory stimulus (spatial vs. visual) and the number of items

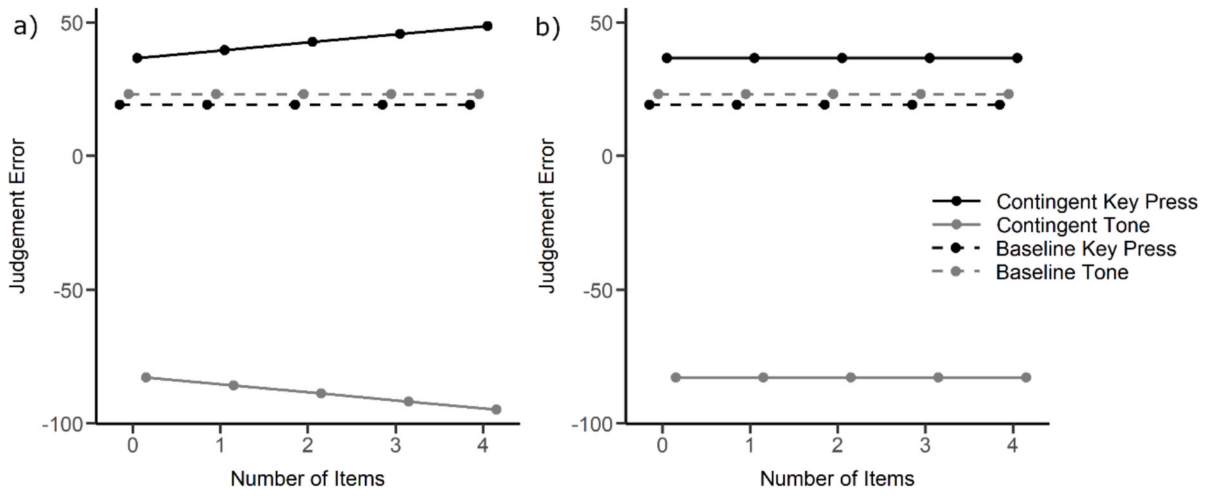

Fig. 5. Visualization of the hypotheses for Experiment 2. Action and outcome binding are expected to increase with spatial memory load (a) but not with visual memory load (b). Binding manifests in the difference between the contingent (solid lines) and the baseline conditions (dashed lines) for the action (target event key press, black lines) and the outcome (target event tone, grey lines).

participants had to remember (0–4). Like Experiment 1 and as usual in the Libet clock task, participants did two baseline and two contingent conditions with the target event being the action and the outcome respectively.

3.1.2. Participants

Our sample size was derived from a simulation-based power analysis (see OSF repository for R code). Data was simulated 5000 times with a SD of 1, a slope for condition of 0.3 and a slope for number of items of 0.05 in the contingent condition and 0.0 in the baseline condition. These effect sizes were chosen to resemble a small effect size with smaller effects being not of interest for our hypotheses. For a linear regression model including condition (contingent vs. baseline), number of items and their interaction, the power analysis resulted in 95 participants to reach a power of 0.90 for the interaction at a significance level of 5 %. Our preregistered analysis using Bayesian mixed models should lead to a higher power. To account for drop-out based on our previous studies of temporal binding, we increased this by 10 %, leading to a sample size of 105 participants.

Six participants reported current neurological or psychiatric disorders, four impaired and uncorrected hearing, and three impaired and uncorrected vision. With overlaps, eleven participants were excluded due to these preregistered criteria leading to a final sample size of 94 participants. 45 participants reported their gender to be female, 49 male and none non-binary. On average, participants were 22.13 years old ($SD = 2.98$). All participants were healthy, young students of sports science at our faculty and received course credit for participation.

The study was executed in accordance with the declaration of Helsinki for the guidelines of ethical considerations. Ethical approval for this study was not required in accordance with the conditions outlined by the German Research Society (DFG) where research that carries no additional risk beyond daily activities does not require a research ethics board's approval. We communicated all considerations necessary to assess the question of ethical legitimacy of the study.

3.1.3. Material

Participants worked through the memory task-embedded Libet-clock task and answered the demographic questionnaire in PsychoPy (version 2023.1.3, Peirce et al., 2019) on a Lenovo Thinkpad T15 wearing noise-cancelling headphones over which they heard the auditory stimuli. They responded using an external keyboard and a mouse with mouse wheel. The demographic questionnaire queried age, gender and the exclusion criteria of current psychiatric or neurological disorders and uncorrected, impaired vision and hearing.

The experiment consisted of two instances of the Libet-clock task, each embedded in the retention interval of a memory task. The two Libet-clock tasks were identical besides the stimulus material of the embedding memory task. Each trial started with the presentation of the memory stimuli immediately followed by one trial of the Libet-clock task. After the presentation and response phases of the Libet-clock task participants had to recognize the presented memory stimuli. All PsychoPy code is available in the OSF repository.

3.1.3.1. Libet-Clock task. The Libet-clock task was identical to the one in Experiment 1, but each trial was embedded in the retention interval of one of two memory tasks. Condition order was again counterbalanced across participants and was the same for both instances of the task. Each condition consisted of 30 trials, 6 for each number of items in the memory task (0–4). Before the presentation phase of the memory task, participants saw a screen informing them of the current Libet-clock condition and target event, how many trials of the current condition they already completed and how many conditions were left. They could start the trial via key press.

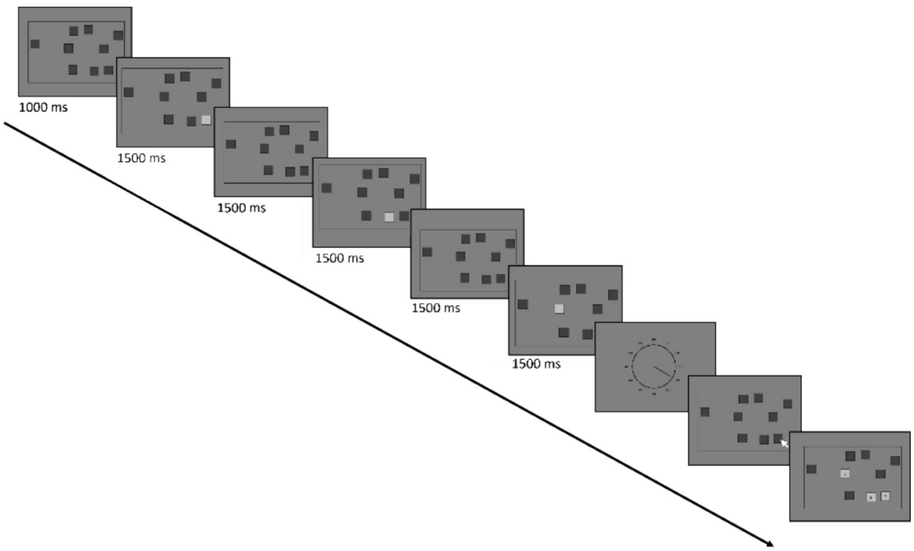

Fig. 6. Exemplary sequence of events for the spatial memory task for a trial with three items.

Before the two instances of the Libet-clock task, participants did three trials per condition without the embedding memory tasks in the same condition order that was used in the actual runs. Training trials were not analyzed.

To examine reliability of temporal binding scores we calculated the correlation between binding scores during the spatial and the visual memory task for action and outcome binding across participants respectively. For action binding this was $r_{\text{test-retest}} = 0.64$ (95 % CI = [.50, 0.74]) and for outcome binding $r_{\text{test-retest}} = 0.81$ (95 % CI = [.72, 0.87]). Like our previous research (Siebertz & Jansen, 2022), we found a higher reliability for outcome binding.

3.1.3.2. Spatial memory task. The spatial memory task was adapted from the Corsi block-tapping test. Fig. 6 shows an exemplary trial. At the start of each trial, participants saw nine dark grey squares on a grey background randomly positioned inside a rectangular border for 1000 ms. Squares could not touch or overlap each other. Corresponding to the current trial's number of items, randomly selected squares changed their color to a light grey for 1500 ms sequentially with 1500 ms interstimulus interval. After the last color change, the squares were visible for another 2000 ms before the presentation phase of the Libet-clock task started. If the current trial's number of items was zero, a white bar appeared below the rectangular border for 1500 ms to inform the participants of this and after another 1000 ms the Libet-clock task started.

After the response phase of the Libet-clock task, the array of squares reappeared, and participants had to select the squares that previously changed their color by clicking on them. Participants could freely select and deselect squares and the button to proceed only appeared when the number of selected squares matched the number of items. For comparability with the visual memory task, the sequence in which squares changed color did not matter and participants were informed of this beforehand. There was no time limit for recognition. For trials with zero items, the recognition phase was skipped. Each of the four conditions of the Libet-clock task consisted of 6 trials per number of items (0–4), totaling in 120 trials of the spatial memory task. The 30 trials per condition were presented in a randomized order.

3.1.3.3. Visual memory task. The visual memory task was identical to the spatial memory task in timing and sequence of events. Fig. 7 shows an exemplary trial. Instead of squares a sequence of zero to four Chinese characters was presented in the center of the screen for 1500 ms each with an interstimulus interval of 1500 ms. Like in the spatial memory task, one trial of the Libet-clock task followed, after which participants had to select the presented characters from an array of nine characters consisting of the presented ones and distractors. Again, there was no time limit for recognizing the characters. In trials with zero items, a white bar appeared for 1500 ms to avoid participants being confused about missing a character. In these trials, the recognition phase was skipped.

For each condition, a list of 240 preselected characters was randomized. For each trial, besides the trials with zero items, nine characters were extracted from that list, and the targets were chosen at random among them. In this way, characters did only appear once per condition and in random character combinations across conditions to mitigate familiarity effects.

3.1.4. Procedure

Due to time constraints and the length of the experiment, we tested two participants at the same time. They sat on opposite sides of

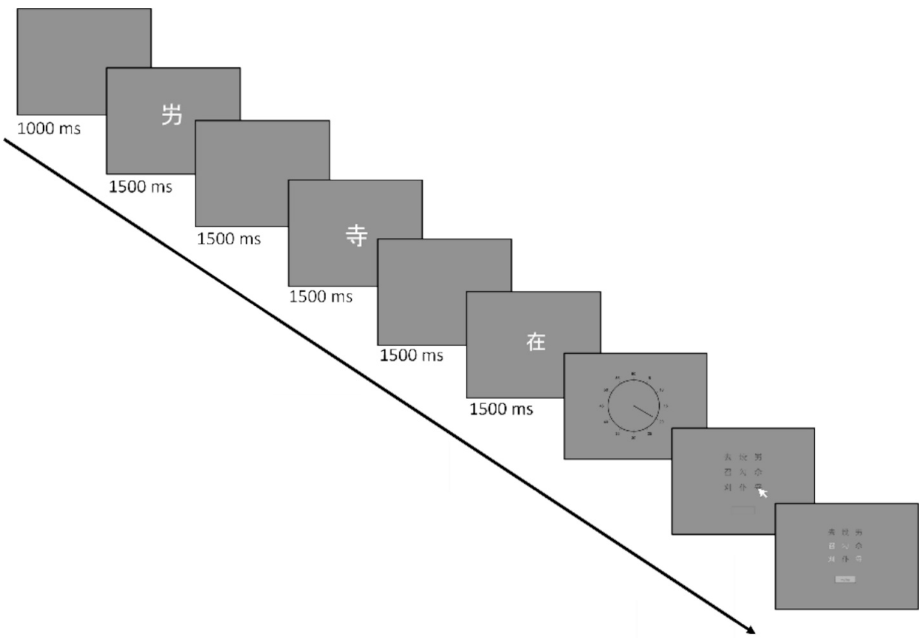

Fig. 7. Exemplary sequence of events for the visual memory task for a trial with three items.

the testing room facing away from each other. Between them was a room divider and they wore noise-cancelling headphones. First, participants read the study information and data protection information before giving digital informed consent. The experimenter informed the participants of the general procedure and the tasks and encouraged them to ask questions. Then they worked through the training trials clearing all ambiguities when they arose. The two instances of the Libet clock task followed and the session ended with the demographic questionnaire. The order of the two instances was counterbalanced across participants. Each instance typically lasted around 45 min and the whole session around two hours. Participants were encouraged to take pauses any time they needed them.

3.1.5. Statistical analysis

We preregistered analyses for Experiment 2 (see the beginning of the Experiment 2 section). We calculated Bayesian linear mixed models using the R package brms (Bürkner, 2017) and choosing non– or weakly informative priors (Bürkner, 2018; Danböck et al., 2023). For ease of interpretation, we separated the preregistered analyses for Hypothesis 1 and 2 and ran separate models for the Libet clock instances embedded in the spatial and the visual memory tasks. We also separated the analysis for action and outcome binding, that is the trials in which participants had to attend the key press and the tone respectively. Instead of calculating temporal binding scores as we did in Experiment 1, we used the judgment errors of the Libet clock task as the outcome and included the predictors of condition (contingent vs. baseline) and number of items (0–4) and their interaction as fixed effects as well as the intercept. The initial maximum model additionally included the random effects for the intercept, condition, number of items, and their interaction.

We fitted this model to our data using the Gaussian, lognormal, skew normal, and student's *t*-distributions recommended for continuous data and identified the most appropriate distribution family with Pareto-*k*s, visual posterior prediction checks, and leave-one-out cross-validation (Gabry et al., 2019; Vehtari et al., 2017). A model was deemed unfitting if more than 5 % of observations showed high Pareto-*k*s (>.7). In direct model comparison via leave-one-out cross-validation if the difference in expected log predictive density between two models was more than two times its SE. For all analyses, student's *t*-distribution was preferred.

Following our preregistration, we reduced our model to optimize power (Matuschek et al., 2017). We reduced the random effects by starting with the correlations between random effects followed by the interaction and the main effects. Model reduction was based on leave-one-out cross-validation. We stopped model reduction when the more complex model showed a better fit which was defined as the difference in expected log predictive density being more than two times its *SE*. The random intercept and the fixed effects were not reduced.

Hypothesis 1 predicts an increase in temporal binding with increasing memory load. In terms of judgment errors, more items in the memory task should lead to a more positive error in the contingent compared to the baseline condition for action binding and a more negative error in the contingent compared to the baseline condition for outcome binding. With our coding of the predictor Condition (baseline as the reference condition), this would show in a positive interaction term of condition and number of items for action binding and a negative interaction term for outcome binding. We interpret statistical significance as zero not being included in the 89 % CI for two-sided and in the 78 % CI for one-sided tests. These intervals have been shown to produce more stable estimates during model calculation than the usual 95 % and 97.5 % CIs and are recommended for Bayesian model fitting (Makowski et al., 2019). As Hypothesis 2 postulates the absence of an effect of substantial magnitude, we interpret Bayes factors to interpret evidence for the null hypothesis versus the alternative with Bayes factors between 1/3 and 3 being inconclusive.

Per participant and per condition trials with judgment errors 3.5 SD below or above the respective mean were excluded as outlier trials.

3.2. Results

3.2.1. Memory task performance

In the spatial memory task across all four conditions of the Libet clock task, participants correctly recognized 225.0 of 240 targets on average ($SD = 12.3$, range = [174, 239]). In the visual memory task, the average number of correctly recognized targets was 214.6 ($SD = 23.9$, range = [112, 240]). As no participants scored below the chance level of 80 targets per memory task, no participant was excluded for subpar memory performance. The total number of correctly recognized targets in the two memory tasks correlated at $r = 0.59$.

Table 2
Bayesian mixed model explaining the judgment error for key presses under spatial memory load with the factor Condition (reference category is the baseline condition) and the continuous predictor Number of Items and their interaction.

			89 %-CI			ESS	
Variable	*B*	*SE*	LL	UL	$\hat{R}$	Bulk	Tail
Fixed Effects							
Intercept	10.62	5.12	2.48	18.80	1.00	3213	5780
Condition	25.10	5.36	16.57	33.60	1.00	5430	9521
n Items	0.17	0.65	−0.86	1.21	1.00	32,876	15,693
Condition: *n* Items	0.15	0.92	−1.31	1.61	1.00	35,357	16,254
Random Effects							
SD(Intercept)	46.83	3.64	41.38	52.94	1.00	5095	8250
SD(Condition)	47.06	3.80	41.37	53.42	1.00	9321	14,007

Note. *n* Items = number of items, LL = lower limit, UL = upper limit.

3.2.2. Action binding

3.2.2.1. Under spatial memory load. The remaining random effects of the model for judgment errors of the key press during the spatial memory task were the intercept and the factor Condition (see Table 2). The one-sided test for the factor Condition shows that action binding occurs under spatial memory load ($B = 25.10$, 78 %-CI = [18.55, 31.61], $PP_{b>0} = 100$ %, $BF_{10} \gg 100$). On average, the key press is reported as occurring 25 ms later in the contingent compared to the baseline condition (Fig. 8a, black lines). Contrary to Hypothesis 1, the interaction of condition and the number of items is not significantly above 0 according to the one-sided test ($B = 0.15$, 78 %-CI = [−0.62, 0.96], $PP_{b>0} = 60$ %, $BF_{10} = 1.51$). The judgment error does not increase more strongly with increasing spatial memory load in the contingent compared to the baseline condition although the Bayes factor indicates inconclusive evidence. The number of items also does not influence the judgment error across both conditions as shown by the absence of a significant main effect.

3.2.2.2. Under visual memory load. The remaining random effects of the model for judgment errors of the key press during the visual memory task were the intercept and the factor Condition (see Table 3). Like with spatial memory load, action binding occurs under visual memory load as well (Fig. 8b, black lines). The one-sided test of the factor Condition is significant ($B = 27.35$, 78 %-CI = [20.80, 33.93], $PP_{b>0} = 100$ %, $BF_{10} \gg 100$). Following Hypothesis 2, the interaction is non-significant ($B = -0.78$, 78 %-CI = [−1.92, 0.36], $PP_{b>0} = 20$ %, $BF_{10} = 0.25$) with the Bayes factor indicating substantial evidence for the null hypothesis. Under visual memory load, judgment errors do not increase more strongly with increasing number of items in the contingent compared to the baseline condition. The main effect of number of items is non-significant as well, indicating no influence of visual memory load on judgment errors in both conditions.

3.2.3. Outcome binding

3.2.3.1. Under spatial memory load. The random effects for the model for judgment errors for tones under spatial memory load reduced to the intercept and the factor Condition (see Table 4). Outcome binding occurred under spatial memory load with a significant one-sided test of the factor condition ($B = -105.98$, 78 %-CI = [−116.34, −95.46], $PP_{b<0} = 100$ %, $BF_{10} \gg 100$). On average, the tone is reported as occurring 106 ms earlier in the contingent compared to the baseline condition (Fig. 8a, grey lines). The interaction of condition and number of items is not significant ($B = 0.36$, 78 %-CI = [−1.02, 1.75], $PP_{b<0} = 38$ %, $BF_{10} = 0.61$) with the Bayes factor indicating inconclusive evidence. This is contrary to Hypothesis 1. Judgment errors for the tone are not increased more strongly (more negative) with increasing number of items in the contingent compared to the baseline condition. Neither does the number of items influence judgment errors across both conditions indicated by the non-significant main effect.

3.2.3.2. Under visual memory load. The model for outcome binding embedded in the visual memory task had convergence issues indicated by $\hat{R} > 1.00$ for some predictors. This problem persisted with 4000 iterations and 1000 warm-up runs. The results in Table 5 are based on 5000 iterations and 1500 warm-up runs which led to satisfactory convergence. The remaining random effects were the intercept and the factor Condition (see Table 5). As with spatial memory load, outcome binding did occur under visual memory load (Fig. 8b, grey lines) with a one-sided significant condition effect ($B = -115.21$, 78 %-CI = [−125.73, −104.62], $PP_{b<0} = 100$ %, $BF_{10} \gg 100$). The interaction of condition and number of items was not significant ($B = 1.51$, 78 %-CI = [0.12, 2.89], $PP_{b<0} = 9$ %, $BF_{10} =$

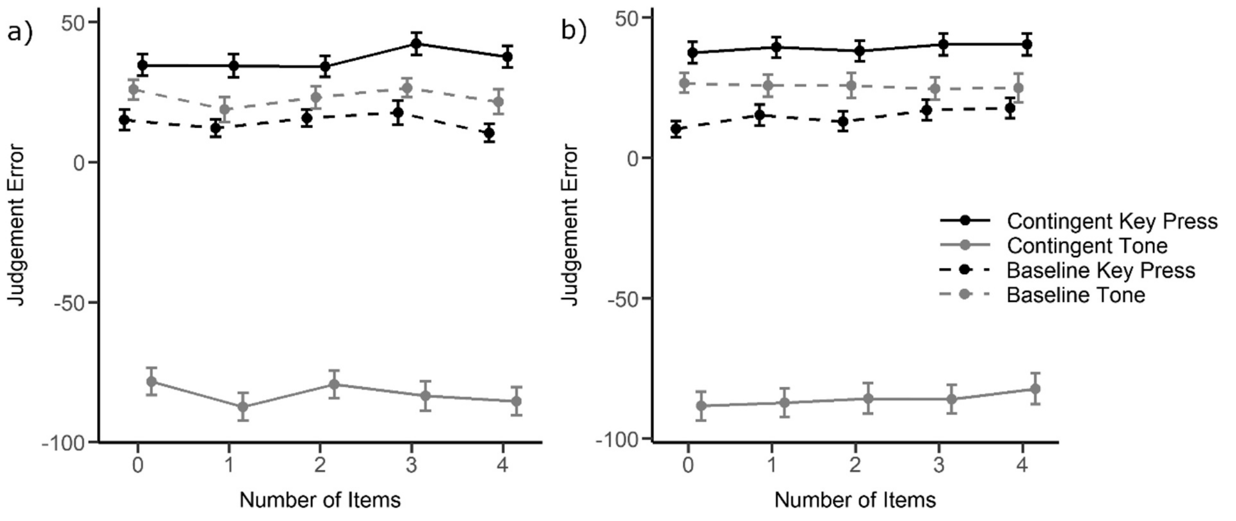

Fig. 8. Judgment errors in the Libet clock task embedded in the spatial memory task (a) and the visual memory task (b). Temporal binding is visible as the difference between the contingent (solid lines) and baseline conditions (dashed lines) for actions (black) and outcomes (grey). Error bars show ± 1 *SE*.

Table 3
Bayesian mixed model explaining the judgment error for key presses under visual memory load with the factor Condition (reference category is the baseline condition) and the continuous predictor Number of Items and their interaction.

Variable	*B*	*SE*	89 %-CI LL	89 %-CI UL	$\hat{R}$	ESS Bulk	ESS Tail
Fixed Effects							
Intercept	9.98	5.15	1.79	18.28	1.00	2403	5246
Condition	27.35	5.33	18.82	35.81	1.00	4348	8538
n Items	0.89	0.65	−0.15	1.93	1.00	33,134	15,710
Condition: *n* Items	−0.78	0.93	−2.26	0.70	1.00	31,293	16,738
Random Effects							
SD(Intercept)	48.07	3.68	42.48	54.27	1.00	4648	8014
SD(Condition)	47.20	3.76	41.59	53.53	1.00	7439	11,964

Note. *n* Items = number of items, LL = lower limit, UL = upper limit.

Table 4
Bayesian mixed model explaining the judgment error for tones under spatial memory load with the factor Condition (reference category is the baseline condition) and the continuous predictor Number of Items and their interaction.

Variable	*B*	*SE*	89 %-CI LL	89 %-CI UL	$\hat{R}$	ESS Bulk	ESS Tail
Fixed Effects							
Intercept	23.34	5.00	15.45	31.43	1.00	3454	6862
Condition	−105.98	8.57	−119.59	−92.34	1.00	2184	5090
n Items	−1.13	0.79	−2.38	0.14	1.00	30,561	15,998
Condition: *n* Items	0.36	1.13	−1.46	2.16	1.00	30,668	14,869
Random Effects							
SD(Intercept)	44.37	3.49	39.18	50.29	1.00	5023	9534
SD(Condition)	79.11	6.03	70.23	89.33	1.00	4156	7404

Note. *n* Items = number of items, LL = lower limit, UL = upper limit.

Table 5
Bayesian mixed model explaining the judgment error for tones under visual memory load with the factor Condition (reference category is the baseline condition) and the continuous predictor Number of Items and their interaction.

Variable	*B*	*SE*	89 %-CI LL	89 %-CI UL	$\hat{R}$	ESS Bulk	ESS Tail
Fixed Effects							
Intercept	25.31	5.85	16.15	34.77	1.00	3492	7455
Condition	−115.21	8.64	−128.85	−101.40	1.00	3525	6605
n Items	−0.05	0.79	−1.32	1.22	1.00	42,703	22,321
Condition: *n* Items	1.51	1.13	−0.31	3.28	1.00	41,031	21,654
Random Effects							
SD(Intercept)	53.35	4.21	47.05	60.48	1.00	7179	13,315
SD(Condition)	80.14	6.10	70.92	90.46	1.00	5717	11,438

Note. *n* Items = number of items, LL = lower limit, UL = upper limit.

0.10), confirming Hypothesis 2 with moderate evidence. In the contingent condition, judgment errors do not become more negative more strongly with increasing number of items compared to the baseline condition. The main effect of number of items is also not significant, showing no influence of memory load on judgment errors across both conditions.

3.2.4. Analysis of three-way-interaction including memory type as a factor

As mentioned in the section on the statistical analysis, the preregistration only included separate two-way models including LC-condition and number of items as predictors. For further insight, we additionally report three-way models including the memory task (spatial vs. visual) as an additional factor. These results must be seen as exploratory. Going beyond the two-way models, the three-way model allows to directly test whether the number of items influences binding differently under spatial and visual memory load. To be more precise, looking at the three-way-interaction we test whether the number of items increases the judgment error more strongly in the contingent condition compared to the baseline condition and whether this differs under the two types of memory load. For action binding, our coding of predictor categories requires a positive regression coefficient of the three-way interaction to indicate a stronger increase of binding with an increasing number of items for the spatial compared to the visual memory task. For outcome binding, this is indicated by a negative regression coefficient.

3.2.4.1. Action binding. The model for judgment errors of the key press containing memory task as an additional factor retained the random effects of the intercept, condition, number of items, memory task and the interaction of condition and memory task (see Table 6). As with the analyses separated by memory task, the main effect of condition is significant (one-sided, $B = 27.17$, 78 %-CI = [20.76, 33.61], $PP_{b>0} = 100$ %, $BF_{10} \gg 100$). The two-way interaction of condition and number of items is not significant ($B = -0.71$, 78 %-CI = [−1.86, 0.42], $PP_{b>0} = 29$ %, $BF_{10} = 0.22$). This means that across both memory tasks, action binding does not increase with the number of items. The three-way interaction is also not significant ($B = 0.87$, 78 %-CI = [−0.75, 2.49], $PP_{b>0} = 74$ %, $BF_{10} = 2.89$). This means that the influence of memory load on action binding does not differ significantly between memory tasks. The Bayes factor favors the alternative hypothesis that there is a three-way interaction but stays just under the preregistered threshold of 3. Descriptively, this result aligns with our hypotheses. However, the Bayes factor indicates inconclusive evidence, and the 78 %-CI includes 0.

3.2.4.2. Outcome binding. The model for outcome binding including memory type as a factor had problems with convergence with $\widehat{R} > 1.00$ for some predictors. Increasing iterations to 4000 and warm-up runs to 1000 did not alleviate this, so the results in Table 7 were calculated with 5000 iterations and 1500 warm-up runs which yielded satisfactory convergence. Identically to the three-way model for action binding, the three-way model for judgment errors of the tone retained the random effects of the intercept, condition, number of items, memory task, and the interaction of condition and memory task (see Table 7). Again, the main effect of condition is significant (one-sided, $B = -114.46$, 78 %-CI = [−125.11, −103.73], $PP_{b<0} = 100$ %, $BF_{10} \gg 100$). The two-way interaction of condition and number of items is not significantly below 0 (one-sided, $B = 1.53$, 78 %-CI = [0.16, 2.9], $PP_{b<0} = 9$ %, $BF_{10} = 0.09$). According to the 78 %-CI, outcome binding decreases with an increasing number of items across both memory tasks. The three-way interaction is also not significantly below 0 (one-sided, $B = -1.06$, 78 %-CI = [−3.01, 0.86], $PP_{b<0} = 75$ %, $BF_{10} = 2.99$) meaning that the influence of memory load on outcome binding does not differ significantly between memory tasks. As with action binding, the Bayes factor favors the alternative hypothesis that there is a three-way interaction staying barely under the preregistered threshold of 3.

3.3. Preliminary discussion

The goal of Experiment 2 was to test the theory that temporal binding in the Libet clock task stems from processes in spatial memory. To test this, we compared binding under increasing levels of spatial memory load with binding under increasing levels of visual memory load. For both action and outcome binding, our preregistered analyses separated by memory type fail to confirm Hypothesis 1. Results suggest that binding does not increase with spatial memory load. However, the Bayes factors remain inconclusive. The preregistered analyses confirm Hypotheses 2. Bayes factors indicate substantial to strong evidence against an effect of visual memory load on binding for actions and outcomes.

Our additional exploratory analysis added memory type as a factor into the models. Descriptively, the effect of memory load on binding does differ between spatial and visual memory load in the expected direction. Binding is increased more strongly with spatial than visual memory load for both, action and outcome binding. However, for both types of binding the results are non-significant. Although favoring the hypothesis of the presence of this effect, Bayes factors stay just below the threshold of what we preregistered as substantial evidence.

A limitation of the memory tasks is that they might not bind purely spatial or visual resources respectively. First, we did not use an articulatory suppression task to prevent subvocal verbalization as a memory strategy. This was in part due to testing multiple participants at once. Spatial and visual content might have been remembered partly using phonological resources. Second, we cannot exclude that participants used other strategies to remember the stimuli. For example, they could have imagined shapes between the target squares in the spatial task (e.g., a triangle when the number of targets was three) or they might have relied on local features of

Table 6
Bayesian mixed model explaining the judgment error for tones under visual memory load with the factor Condition (reference is the baseline condition), the continuous predictor Number of Items, the factor Memory Task (reference category is the visual memory task) and their interactions.

			89 %-CI			ESS	
Variable	*B*	*SE*	LL	UL	$\widehat{R}$	Bulk	Tail
Fixed Effects							
Intercept	10.27	5.02	2.17	18.10	1.00	1536	2925
Condition	27.17	5.28	18.74	35.45	1.00	2486	4824
n Items	0.82	0.66	−0.24	1.88	1.00	14,062	15,342
Memory Task	0.31	3.23	−4.85	5.46	1.00	7291	12,064
Condition: *n* Items	−0.71	0.93	−2.21	0.76	1.00	13,603	14,386
Condition: Memory Task	−1.89	4.30	−8.79	4.99	1.00	9121	13,080
n Items: Memory Task	−0.65	0.92	−2.13	0.82	1.00	12,790	15,170
Condition: *n* Items: Memory Task	0.87	1.33	−1.23	2.99	1.00	12,247	14,640
Random Effects							
SD(Intercept)	47.12	3.57	41.73	53.02	1.00	2889	5950
SD(Condition)	45.53	3.61	40.19	51.64	1.00	5320	9337
SD(*n* Items)	1.16	0.63	0.15	2.16	1.00	3750	5357
SD(Memory Task)	22.18	2.15	18.93	25.76	1.00	7893	12,573
SD(Condition: Memory Task)	26.55	2.84	22.28	31.23	1.00	7734	12,527

Note. *n* Items = number of items, LL = lower limit, UL = upper limit.

Table 7
Bayesian mixed model explaining the judgment error for tones under visual memory load with the factor Condition (reference is the baseline condition), the continuous predictor Number of Items, the factor Memory Task (reference category is the visual memory task) and their interactions.

			89 %-CI			ESS	
Variable	*B*	*SE*	LL	UL	$\hat{R}$	Bulk	Tail
Fixed Effects							
Intercept	25.07	5.67	16.16	34.16	1.00	3255	6087
Condition	−114.46	8.75	−128.45	−100.38	1.00	2124	5572
n Items	−0.05	0.80	−1.35	1.23	1.00	20,647	21,309
Memory Task	−1.30	3.85	−7.46	4.83	1.00	13,278	18,779
Condition: *n* Items	1.53	1.12	−0.23	3.32	1.00	20,314	22,055
Condition: Memory Task	8.04	5.65	−1.01	17.06	1.00	11,861	17,108
n Items: Memory Task	−1.14	1.11	−2.92	0.63	1.00	19,635	21,896
Condition: *n* Items: Memory Task	−1.06	1.58	−3.58	1.44	1.00	19,020	21,412
Random Effects							
SD(Intercept)	51.03	3.99	45.01	57.69	1.00	6059	11,102
SD(Condition)	79.31	6.08	70.25	89.50	1.00	4689	7774
SD(*n* Items)	1.74	0.80	0.35	2.95	1.00	5488	8331
SD(Memory Task)	26.20	2.64	22.18	30.62	100	12,975	17,716
SD(Condition: Memory Task)	40.57	3.87	34.72	47.08	1.00	12,615	17,914

Note. *n* Items = number of items, LL = lower limit, UL = upper limit.

the Chinese characters such as the orientation of certain lines that stood out to them. Some participants reported using parts of the characters that resembled familiar objects, such as a house or a Latin letter. These factors might contribute to the overlap in the resources that are bound as indicated by the relatively high correlation between the two tasks.

4. General discussion

The aim of this study was to test the theory that binding in the Libet clock task results from interference in spatial working memory. In favor of this theory, we found reduced outcome binding in the Libet clock task with increasing spatial working memory performance in the 2-back task in Experiment 1. We did not find this relationship with action binding. As action and outcome binding seem uncorrelated (Siebertz & Jansen, 2022; Tonn et al., 2021), it is no surprise to find such a relationship with one but not the other binding type. At the same time, the evidence was not strong enough to completely rule out a similar relationship with visual working memory

Table 8
Overview of the hypotheses and respective level of evidence.

	Hypothesis	Level of Evidence
Experiment 1	H1: Action and outcome binding occur in the Libet-clock task.	BF ≫ 100, overwhelming in favor of H1 for action and outcome binding.
	H2: Action and outcome binding occur in a spatially random task.	Action binding: BF = 0.04, strong against H2. Outcome binding: BF ≫ 100, overwhelming in favor of H2.
	H3: Action binding in the Libet-clock task correlates positively with action binding in the new, spatially random task.	BF = 0.18, moderate against H3.
	H4: Outcome binding in the Libet-clock task correlates positively with outcome binding in a spatially random task.	BF = 0.16, moderate against H4.
	H5: Binding in both tasks correlates negatively with spatial working memory performance.	LC Action binding: BF = 0.24, moderate against H5. LC Outcome binding: BF = 6.76, moderate in favor of H5. SR Action binding: BF = 0.29, moderate against H5. SR Outcome binding: BF = 0.60, inconclusive.
	H6: Binding in both tasks does not correlate with visual working memory performance.	LC Action binding: BF = 0.12, substantial in favor of H6. LC Outcome binding: BF = 0.57, inconclusive. SR Action binding: BF = 0.66, inconclusive. SR Outcome binding: BF = 3.01, moderate against H6.
Experiment 2	H1: In the Libet clock task embedded in the spatial memory task, temporal binding increases the more items have to be remembered.	Action binding: BF = 1.51, inconclusive. Outcome binding: BF = 0.61, inconclusive.
	H2: In the Libet clock task embedded in the visual object-releated memory task, the number of items does not increase temporal binding.	Action binding: BF = 0.25, substantial in favor of H2. Outcome binding: BF = 0.10, substantial in favor of H2.

Note. BF: Bayes Factor, LC: Libet Clock Task, SR: Spatially Random Task.

performance. The link to spatial working memory might still depend on more general attentional resources that are shared between the two working memory systems. Table 8 shows an overview of the hypotheses and the respective level of evidence in favor or against each one. The correlation between binding and visual working memory performance could also be attenuated due to the relatively low reliability of the visual 2-back task. Although somewhat higher, the reliability of the spatial 2-back task is not optimal either. This shows the importance of investigating the reliability of cognitive tasks for each adaptation of the task and for every new sample (Parsons et al., 2019).

Weighing against our theory, we could not find related binding when we removed the rotating clock from the Libet clock task in Experiment 1. This could mean that the binding effects in the Libet clock task require a continuously moving event timing reference or that they depend on reduced visual precision, especially between action and outcome. We found lower reliability for the spatially random task than in the Libet clock task. Nonetheless, it seems unlikely that we did not find a relationship between binding in the two tasks solely because of reliability-based correlation attenuation since the correlation coefficient is very small.

In the preregistered analyses for Experiment 2 we found evidence against an effect of visual memory load on binding but inconclusive evidence regarding the effect of spatial memory load. Formally, this fails to confirm the hypothesis of spatial working memory processes being responsible for binding in the Libet clock task. Our exploratory analysis looking at the three-way interaction between condition, memory load and memory type yielded a similar result. We descriptively found a stronger increase in action and outcome binding with increasing spatial working memory load compared to increasing visual working memory load in Experiment 2. However, this result stays below the significance threshold, and the analysis was not preregistered. Future research might replicate our experimental approach in Experiment 2 with sufficient power to investigate the three-way interaction between Libet clock task condition, working memory load, and working memory type. Taken together, our study yields only weak evidence for the theory that binding in the Libet clock task is based in spatial working memory processes. The few findings supporting the theory could still be explained by more general attentional processes (Schwarz & Weller, 2022). This would require the spatial and visual working memory tasks to differ in attentional demands. We designed the two tasks to be identical in terms of number of trials and items, duration, response modality and visibility. We counterbalanced the order of the two tasks across participants to account for effects of fatigue and practice. Difficulty does also not seem to differ much between the two tasks. One reviewer pointed out that suitable memory strategies might make the tasks demand fewer attentional resources. As discussed before, both tasks provide opportunities to abstract the stimulus material into more familiar or summarized memory content. Taken together, the two tasks do not differ in attentional demands in an obvious way. Future research could use instructions against strategy use or ask participants to report their memory strategies to better judge the impact they may have.

While our study does not support the theory, a few alternative explanations for the lack of evidence should be considered. Spatial working memory is a group of interlocking processes (e.g. Baddeley, 2000; Oberauer & Lin, 2017; Ren et al., 2023). Depending on the demands of a specific task, these processes contribute in varying degrees to solving it. The 2-back task for example does not require a large working memory capacity but constant updating of the relevant information contrary to other tasks such as the digit span backwards (Haatveit et al., 2010; Hilbert et al., 2015). We chose the 2-back task in Experiment 1 because it heavily relies on maintaining and updating which we hypothesized to be relevant processes leading to binding in the Libet clock task. More specifically, we hypothesized that maintenance is required to precisely report the action-related position and updating for the outcome-related position. Further research could use a task that, e.g., isolates updating (Ecker et al., 2014) and could look for a relationship to outcome binding. Another consideration is that working memory processes can be subdivided further. Updating for example consists of retrieval, transformation and substitution (Ecker et al., 2010). Transformation or substitution might be particularly important mechanisms for binding in the Libet clock task and using tasks that do not isolate these subprocesses would attenuate any observable correlation. This is especially challenging for a dual task design as we used in Experiment 2. What we imposed simultaneously to the Libet clock task where additional storage demands. It would be interesting to find a dual task that imposes additional processing demands instead.

Finally, it is worth mentioning that despite the importance of understanding the underlying processes of binding it does not necessarily answer the question of the role of intention and agency. If for example, spatial attention is the cause of binding (Schwarz & Weller, 2022) intentional action might still have a strong influence on the focus of attention. This will have to be investigated separately.

4.1. Conclusion

This study investigated whether spatial working memory processes are the cause of binding in the Libet clock task. Our results suggest that this is likely not the case. Especially, isolating the spatial aspect from the continuous movement did not reproduce binding as observed in the Libet clock task. We only found a weak relationship between outcome binding and spatial working memory measured which might warrant further inquiry. However, general attentional ability could not be excluded as an explanation for this finding.

CRediT authorship contribution statement

Markus Siebertz: Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Petra Jansen:** Writing – review & editing, Validation, Supervision, Resources, Funding acquisition, Conceptualization.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data, analysis code and PsychoPy code are available and linked in the article text.

References

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

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

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software, 80*(1). https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal, 10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Cavazzana, A., Begliomini, C., & Bisiacchi, P. S. (2014). Intentional binding effect in children: Insights from a new paradigm. *Frontiers in Human Neuroscience, 8*, 651. https://doi.org/10.3389/fnhum.2014.00651

Danböck, S. K., Franke, L. K., Miedl, S. F., Liedlgruber, M., Bürkner, P.-C., & Wilhelm, F. H. (2023). Experimental induction of peritraumatic dissociation: The role of negative affect and pain and their psychophysiological and neural correlates. *Behaviour Research and Therapy, 164*, Article 104289. https://doi.org/10.1016/j.brat.2023.104289

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working memory. *Annual Review of Psychology, 66*(1), 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Di Plinio, S., Arnò, S., Perrucci, M. G., & Ebisch, S. J. H. (2019). Environmental control and psychosis-relevant traits modulate the prospective sense of agency in non-clinical individuals. *Consciousness and Cognition, 73*, Article 102776. https://doi.org/10.1016/j.concog.2019.102776

Dienes, Z. (2019). How do I know what my theory predicts? *Advances in Methods and Practices in Psychological Science, 2*(4), 364–377. https://doi.org/10.1177/2515245919876960

Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language, 74*, 77–90. https://doi.org/10.1016/j.jml.2013.09.003

Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 170–189. https://doi.org/10.1037/a0017891

Engbert, K., Wohlschläger, 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., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian Workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society, 182*(2), 389–402. https://doi.org/10.1111/rssa.12378

Graham-Schmidt, K. T., Martin-Iverson, M. T., Holmes, N. P., & Waters, F. A. V. (2016). When one's sense of agency goes wrong: Absent modulation of time perception by voluntary actions and reduction of perceived length of intervals in passivity symptoms in schizophrenia. *Consciousness and Cognition, 45*, 9–23. https://doi.org/10.1016/j.concog.2016.08.006

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A. (2010). The validity of *d* prime as a working memory index: Results from the "Bergen *n* -back task. *Journal of Clinical and Experimental Neuropsychology, 32*(8), 871–880. https://doi.org/10.1080/13803391003596421

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

Hartley, A. A., Speer, N. K., Jonides, J., Reuter-Lorenz, P. A., & Smith, E. E. (2001). Is the dissociability of working memory systems for name identity, visual-object identity, and spatial location maintained in old age? *Neuropsychology, 15*(1), 3–17. https://doi.org/10.1037/0894-4105.15.1.3

Hilbert, S., Nakagawa, T. T., Puci, P., Zech, A., & Bühner, M. (2015). The digit span backwards task: Verbal and visual cognitive strategies in working memory assessment. *European Journal of Psychological Assessment, 31*(3), 174–180. https://doi.org/10.1027/1015-5759/a000223

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental Brain Research, 203*(2), 465–470. https://doi.org/10.1007/s00221-010-2199-1

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

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General, 133*(3), 355–381. https://doi.org/10.1037/0096-3445.133.3.355

Libet, B., Gleason, C. A., & Wright, E. W. (1983). *Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential)*. 20.

Lush, P., Roseboom, W., Cleeremans, A., Scott, R., Seth, A. K., & Dienes, Z. (2019). Intentional binding as Bayesian cue combination testing predictions with trait individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 45*(9), 1206–1217. https://doi.org/10.1037/xhp0000661

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software, 4*(40), 1541. https://doi.org/10.21105/joss.01541

Martin, J.-R. (2013). Experiences of activity and causality in schizophrenia: When predictive deficits lead to a retrospective over-binding. *Consciousness and Cognition, 22*(4), 1361–1374. https://doi.org/10.1016/j.concog.2013.09.003

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" Tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moore, J. W., Cambridge, V. C., Morgan, H., Giorlando, F., Adapa, R., & Fletcher, P. C. (2013). Time, action and psychosis: Using subjective time to investigate the effects of ketamine on sense of agency. *Neuropsychologia, 51*(2), 377–384. https://doi.org/10.1016/j.neuropsychologia.2012.07.005

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

Nolden, S., Haering, C., & Kiesel, A. (2012). Assessing intentional binding with the method of constant stimuli. *Consciousness and Cognition, 21*(3), 1176–1185. https://doi.org/10.1016/j.concog.2012.05.003

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review, 124*(1), 21–59. https://doi.org/10.1037/rev0000044

Parsons, S., Kruijt, A.-W., & Fox, E. (2019). Psychological Science needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in Methods and Practices in Psychological Science, 2*(4), 378–395. https://doi.org/10.1177/2515245919879695

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Ren, M., Xu, J., Li, Y., Wang, M., Georgiev, G., Shen, L., Zhao, J., Cao, Z., Zhang, S., Wang, W., Xu, S., Zhou, Z., Chen, S., Chen, X., Shi, X., Tang, X., & Shan, C. (2023). Neural signatures for the n-back task with different loads: An event-related potential study. *Biological Psychology, 177*, Article 108485. https://doi.org/10.1016/j.biopsycho.2023.108485

Ruess, M., Thomaschke, R., & Kiesel, A. (2017). The time course of intentional binding. *Attention, Perception, & Psychophysics, 79*(4), 1123–1131. https://doi.org/10.3758/s13414-017-1292-y

Schwarz, K. A., & Weller, L. (2022). Distracted to a fault: Attention, actions, and time perception. *Attention, Perception, & Psychophysics.*. https://doi.org/10.3758/s13414-022-02632-x

Siebertz, M., & Jansen, P. (2022). Diverging implicit measurement of sense of agency using interval estimation and Libet clock. *Consciousness and Cognition, 99*, Article 103287. https://doi.org/10.1016/j.concog.2022.103287

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schumacher, E. H., & Minoshima, S. (1995). Spatial versus Object Working memory: PET Investigations. *Journal of Cognitive Neuroscience, 7*(3), 337–356. https://doi.org/10.1162/jocn.1995.7.3.337

Tonn, S., Pfister, R., Klaffehn, A. L., Weller, L., & Schwarz, K. A. (2021). Two faces of temporal binding: Action- and effect-binding are not correlated. *Consciousness and Cognition, 96*, Article 103219. https://doi.org/10.1016/j.concog.2021.103219

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 1012–1028. https://doi.org/10.1037/a0015859

Voss, M., Moore, J. W., Hauser, M., Gallinat, J., Heinz, A., & Haggard, P. (2010). Altered awareness of action in schizophrenia: A specific deficit in predicting action consequences. *Brain, 133*(10), 3104–3112. https://doi.org/10.1093/brain/awq152

Wenke, D., & Haggard, P. (2009). How voluntary actions modulate time perception. *Experimental Brain Research, 196*(3), 311–318. https://doi.org/10.1007/s00221-009-1848-8