Intentional binding – Is it just causal binding? A replication study of Suzuki et al. (2019)

Michael Wiesing, Eckart Zimmermann

Consciousness and Cognition 119 (2024) 103665

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Full Length Article

Intentional binding – Is it just causal binding? A replication study of Suzuki et al. (2019)

Michael Wiesing [*], Eckart Zimmermann

Institute for Experimental Psychology, Heinrich Heine University Duesseldorf, Germany

A R T I C L E I N F O

Keywords:
Virtual Reality
Intentional Binding
Sense of Agency
Open Data

A B S T R A C T

Intentional actions produce a temporal compression between the action and its outcome, known as intentional binding. However, Suzuki et al. (2019) recently showed that temporal compression can be observed without intentional actions. However, their results show a clear regression to the mean, which might have confounded the estimates of temporal intervals. To control these effects, we presented temporal intervals block-wise. Indeed, we found systematically greater compression for active than passive trials, in contrast to Suzuki et al. (2019). In our second experiment, our goal was to conceptually replicate the previous study. However, we were unable to reproduce their results and instead found more pronounced temporal compression in active trials compared to passive ones. In a subsequent attempt at a direct replication, we did not observe the same findings as the original study. Our findings reinforce the theory that intentions rather than causality cause temporal binding.

During the preparation of this work, the authors used ChatGPT in order to improve the readability of the paper. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

1. Introduction

The human brain faces the constant need to distinguish sensations generated by its own actions from sensations that were produced by external stimulation. This discrimination has remarkable effects on perception. Sensory consequences of self-generated actions typically appear attenuated (Knoetsch & Zimmermann, 2021; Wolpe et al., 2013). Additionally, intentional actions produce a temporal compression between the action and its outcome, a phenomenon termed *intentional binding* (Haggard et al., 2002).

However, some research has cast doubt on the assumed link between *intentional binding* and intentional actions. Instead, studies provided evidence suggesting that temporal binding results from the causal relation linking actions with their consequences (e.g., Buehner, 2012; Buehner & Humphreys, 2009; Kirsch et al., 2019).

In line with this assumption, a recent study by Suzuki et al. (2019) has cast doubt on the role of intentions in the phenomenon of *intentional* binding. They explored whether this binding occurs even in the absence of intentional actions and if these effects have the same magnitude as classical intentional effects. Utilizing a virtual reality setup, they created an environment in which participants received the same visual and tactile feedback for both, intentional actions and mere observations This allowed them to distinctly evaluate the role of intentionality. In their study, participants either: (a) actively pressed a button, (b) watched a virtual hand press the button, or (c) saw the button move on its own. Suzuki et al. (2019) observed the typical temporal compression effects when contrasting

* Corresponding author.
E-mail address: mi.wiesing@gmail.com (M. Wiesing).

https://doi.org/10.1016/j.concog.2024.103665
Received 16 November 2023; Received in revised form 23 January 2024; Accepted 6 February 2024
Available online 14 February 2024

the active and no hand conditions, with the active condition showing reduced interval estimates. Interestingly, no significant differences in temporal binding were observed when comparing the active and fake hand conditions. Moreover, the fake hand conditions showed significantly more temporal compression than the no hand condition. In a follow-up experiment, the authors examined whether differences in agency could explain the observed effects. Participants, however, reported higher agency in the active trials than in fake hand trials, suggesting that the observed effects are not the result of differences in agency across conditions. Suzuki et al. concluded that causal binding, rather than intentional binding, might be the predominant influence, a conclusion that, if replicated, would challenge existing the common understanding of intentional binding. Since only the intention of the action was varied between conditions but the magnitude of the temporal compression effect remained the same, the authors concluded that causal, not intentional binding was driving the effect. If true, these conclusions would challenge the common understanding of intentional binding.

In reviewing the data presented by Suzuki et al. (2019), particularly those shown in their Fig. 3, we observed a couple of noteworthy patterns. Firstly, the temporal compression seemed most pronounced in the longer temporal interval of 800 ms, while virtually absent from the short temporal intervals. Secondly, the temporal estimates appeared to show clear tendencies consistent with regression towards the mean. Perceptual judgments often must be reached when only impoverished sensory information is available. Under these circumstances, our sensory system takes additional sources of information into account, one of them being the recent perceptual history (Hollingworth, 1910; Jazayeri & Shadlen, 2010). The reliance on previous perception results in a regression towards the average sensory stimulation. The central tendency effect is particularly pronounced in time perception, in which judgments can gravitate towards an intermediate interval (Vierordt's Law). One of us has previously shown that temporal estimates in intentional binding can be influenced by central tendency effects, which might sometimes overshadow temporal compression. (Zimmermann & Cicchini, 2020).

The present study aims to determine whether differences in perceived duration between fake hand and active trials are influenced by the degree of the regression to the mean. We performed a conceptual replication of Suzuki et al.'s (2019) Experiment 1, with one key alteration: instead of using a random sequence for presenting temporal intervals, we employed a block design to potentially reduce the effect of regression toward the mean. In a follow-up experiment, we adhered to the original methodology of Suzuki et al. (2019), using a randomized and counterbalanced presentation of intervals. We hypothesized that the block-wise presentation in Experiment 1 would lead to a diminished regression to the mean, thereby revealing more distinct temporal compression differences between active and fake hand trials.

1.1. Experiment 1

Experiment 1 was a conceptual replication of Suzuki et al.'s (2019) Experiment 1, with one key alteration: instead of using a random sequence for presenting temporal intervals, we employed a block design to potentially reduce the effect of regression toward the mean.

2. Material and methods

2.1. Participants

26 participants (five males, mean age: 25.77 years, age range: 19 –54 years, two left-handed) took part in Experiment 1. All participants had a normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. Before the start of the experiment, written informed consent was obtained following the Declaration of Helsinki. Participants subsequently received either monetary compensation or course credits. The study was approved by the ethics committee of the Faculty of Mathematics and Natural Sciences of the Heinrich-Heine-University Düsseldorf, Germany. The sample size was determined based on a Bayesian stopping rule. We started to collect data from 20 participants before we initially examined the evidence for or against each hypothesis. Data collection was stopped once substantial evidence was acquired supporting or against each experimental hypothesis across the pertinent comparisons.

2.2. Apparatus and stimuli

Participants were seated in a chair wearing a head-mounted display (HMD) and had motion-controllers attached to each hand. Stimuli were delivered by an Intel i7-based PC (Intel, Santa Clara, CA) with an NVIDIA GTX 3070 connected to an HTC Vive Pro Eye (HTC Corporation, Taoyuan, Taiwan). The HMD presents stimuli on two low-persistence organic light emitting diode (OLED) displays with a resolution of 1,440 $\times$ 1,600 pixels per eye and a refresh rate of 90 Hz. Before each session, the SRanipal eye tracking calibration software was employed to ensure optimal HMD fit and to adjust the inter-pupillary distance.

Participants used Valve Index motion controllers, designed with straps for secure attachment to the hands. This design allows participants to fully open their hands without the risk of dropping the controllers. Furthermore, the controllers contain capacitive sensors, which can be used to approximate the postures of the fingers. This was used to track the hand and finger movements. However, in the current study, we decided to not show naked hands, but instead, the standard SteamVR gloves were presented. Past research indicates that discrepancies between a participant's actual hands and their virtual representation can influence factors like the sense of presence (Jung et al., 2018; Ogawa et al., 2019) and perceptions of size and body ownership (Jung et al., 2018). By presenting gloves we were able to avoid discrepancies like differences such as skin color or gender. The controllers were also used to provide vibrotactile feedback during the experiment.

Head and hand movements were tracked via the HMD and controllers using the standard SteamVR tracking system. According to previous research (Niehorster et al., 2017), this tracking system provides a robust tracking of head and hand motion with a 360° coverage provided tracking loss is prevented. Tests by (Verdelet et al., 2019) demonstrated a submillimeter precision (0.237 mm) and an accuracy of 8.7 mm for static and 8.5 mm for dynamic objects. Although the system can update the user's pose (position and orientation) at a higher rate (up to 1,000 Hz for the HMD and 250 Hz for the controllers), in this study the sampling rate for both HMD and controllers was limited by the HMD's refresh rate of 90 Hz.

The virtual environment was rendered in Unreal Engine 4.26 (UE4). We adapted a version of UE4′s "Archviz Interior" sample project (EpicGames, 2019) to suit our needs (Fig. 1). While the original sample project utilized real-time raytracing—a rendering technique that's currently too resource-intensive for VR—we made several optimizations to ensure a consistent 90 frames per second (FPS). This involved disabling raytracing as well as disabling post-processing and opting for baked lighting and reflections, Moreover, the polygon count and texture resolutions of most objects were reduced. In the virtual environment, participants found themselves seated before a virtual table (160 × 60 × 70 cm) that aligned with a real-world table. A red buzzer was centrally positioned for participants that could be pressed with either their palm or index finger of the right hand. It required being pressed down to a depth of 1 cm to register an action. Once fully pressed, the controller vibrated for 200 ms, to provide tactile feedback.

To the right of the buzzer was a small mat (17 × 21 cm), designating the resting position for the participants' right hand between trials. A tangible rubber mat was placed at the same location in the real world. The center of the mat was placed about 50 cm in front and 24 cm to the right of the participants.

To the buzzer's left, a numeric keypad was available for participants to input their interval estimates using their right index finger. Similar to the buzzer, It required being pressed down to a depth of 0.4 cm to register an action. Upon a full press, the controller emitted vibrotactile feedback. After entering the estimate, they had to press "accept" to confirm their choice. If participants wanted to correct their estimate they could press "delete" and enter the corrected estimate before confirming their choice.

2.3. Procedure

The experiment started with 18 training trials of the interval estimation task. Participants heard two tones (both 880 HZ for 50 ms each) and had to estimate the duration of the interval that was marked by the tones. The duration was randomly chosen for 3 levels (200 ms, 400 ms, and 800 ms). The same interval durations were also used in the main experiment. During the training session, the duration was further randomly jittered between −100 and 100 ms to prevent participants from learning the specific intervals of the main experiment. After both tones were played, participants were instructed via the display on the table, to enter their duration estimate in milliseconds. After the participants hit accept, a prompt on the display instructed the participants to return to the starting position with their right hand, to start the next trial. Furthermore, during the training trials, participants received feedback about their estimates on the virtual display. The feedback consisted of the difference between the estimate and the true interval duration. Hence, overestimations were indicated as positive numbers, and underestimations as negative numbers.

Afterward, participants performed 10 practice trials. Participants were instructed to press the red button on the table with the right hand and to return to the starting position afterward. After randomly selected interval duration (200 ms, 400 ms, 800 ms), a tone (880 HZ) followed, and participants had to estimate the duration between the button press and the tone.

Fig. 1. Virtual environment created in Unreal Engine 4. Participants viewed virtual gloves that mirrored their hand and finger motions. Between trials, the hand rested on the small mat to the right of the red buzzer. A physical rubber mat was positioned at the corresponding location to offer haptic feedback, signifying the starting point. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In Suzuki et al. (2019), practice trials served to capture participants' hand movements, which were subsequently replayed during the fake hand condition. In contrast, our approach utilized a consistent set of ten pre-recorded hand animations for all participants. This decision was made to eliminate the possibility of inadvertently recording and later replaying any tracking anomalies or user errors. By using a standardized set of animations, we aimed to ensure a consistent experience for all participants, thereby minimizing potential variables that could influence the outcome of the experiment.

In Suzuki et al. (2019), the practice trials served to capture participants' hand movements, which were subsequently replayed in the fake hand condition. In contrast, our approach utilized the same set of ten pre-recorded hand animations for all participants. This decision was made to eliminate the possibility of inadvertently recording and later replaying any tracking anomalies or user errors occurring during recording. By using a standardized set of animations, we aimed to ensure a consistent experience for all participants, thereby minimizing potential variables that could influence the outcome of the experiment.

After the training, participants completed 2 blocks of each condition: active, no hand, and fake hand. The order of the conditions was counterbalanced across participants. The active condition was identical to the practice trials. In the fake hand condition, participants observed the prerecorded hand movements pressing the button. While the fake hand was visible, the virtual hand controlled by the participant remained invisible to induce the impression that the fake hand belonged to the participant.

For the no-hand condition, the same prerecorded hand movements from the fake hand condition were used, but the virtual hand remained invisible. Furthermore, the virtual hand controlled by the participants remained visible on the table during no hand trials. Hence, in both fake hand and no hand trials, the button press happened without intentional action.

In each condition, participants received vibrotactile feedback through the controller as soon as the button was pressed. Each block consisted of 51 trials consisting of 3 sub-blocks, one for each interval duration. The experiment lasted about 1 h including instructions an debriefing and a short break.

2.4. Analysis

The free statistical software R (R Core Team, 2021) RStudio (RStudio Team, 2020), and JASP (JASP Team, 2023) were used to analyze the behavioral data. Plots were generated using ggplot2 (Wickham, 2016).

For the analysis, we followed the procedure described by Suzuki et al. (2019), and a Bayesian stopping rule was applied to determine the final sample size. We first collected data from 20 participants and examined if the analysis provided substantial Bayesian evidence for or against each comparison. A Bayes factor (BF) greater than 3 indicates substantial evidence in favor of the experimental hypothesis, while a BF below 0.3 would provide substantial evidence in favor of the null hypothesis (Dienes, 2014). The sample size was incrementally increased until we obtained substantial evidence either in support of or against each comparison.

Following Suzuki et al. (2019), we conducted a 3 × 3 repeated measure analysis of variance (ANOVA) on the duration with the factors *Condition* (active, fake hand, and no hand) and *Interval* (200 ms, 400 ms, and 800 ms). In the second step, we calculated the average interval estimates across the three intervals, for each participant and condition, and conducted paired-sample *t-tests* and Bayesian paired-sample t-tests.

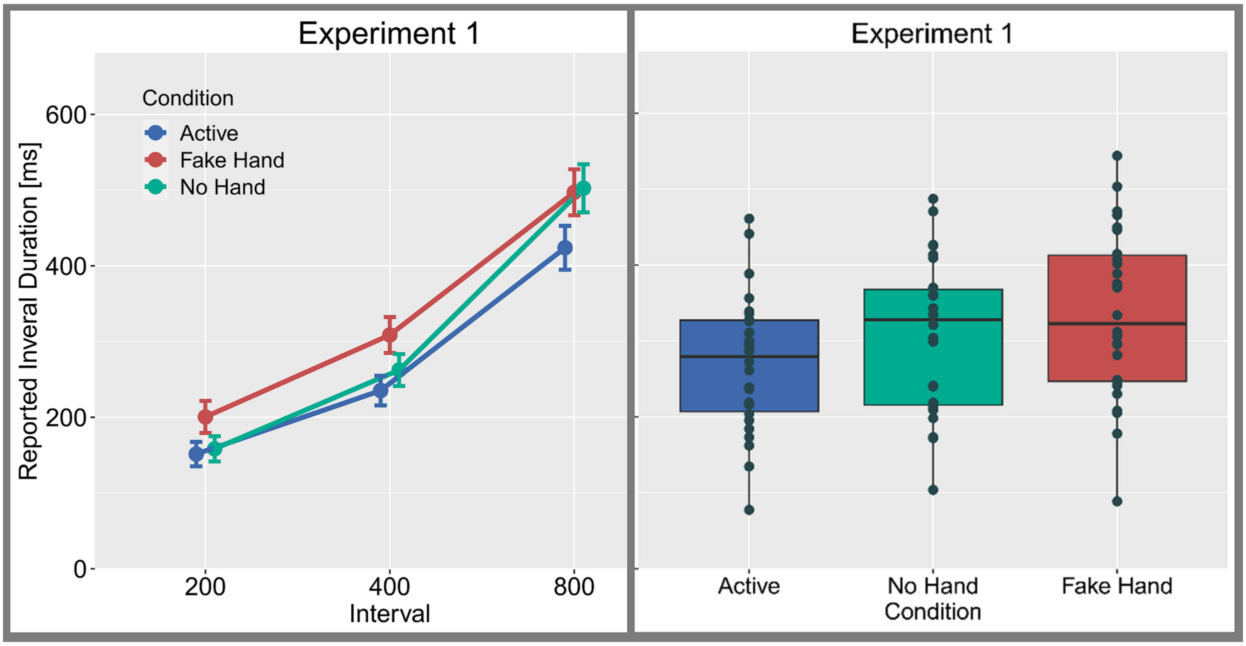

Fig. 2. Left: Average reported duration for each actual duration and condition in Experiment 1. Right: Average reported interval duration reported across the three conditions in Experiment 1. Error bars indicate the standard error of the mean.

For the Bayesian analyses, we utilized JASP and applied its standard prior settings. Specifically, JASP's default prior for t-tests is a Cauchy distribution with a width of 0.707.

2.5. Results

The ANOVA revealed a significant main effect for *Condition* ($F(2,50) = 17.85$, $p < 0.001$, $\eta_p^2 = 0.42$) as well as for the factor *Interval* ($F(1.12,28.06) = 119.08$, $p < 0.001$, $\eta_p^2 = 0.87$, Greenhouse-Geisser corrected). The interaction was also significant ($F(4,100) = 4.87$, $p = 0.001$, $\eta_p^2 = 0.15$) (Fig. 2).

For the comparison between active (Mean = 270.71 ms, SD = 91.33 ms) and no hand (Mean = 307.77 ms, SD = 101.67 ms), we expected longer interval durations for the no hand condition, which was supported by the analysis ($t(25) = 3.41$, $p = 0.001$ (one-tailed), Cohen's d = 0.67, BF = 34.31) (Figs. 2 and 1C). Hence, a difference between both conditions is about 34 times more likely than the H0 hypothesis, which provides sensitive evidence for the H1 hypothesis. Next, we compared the fake hand (Mean = 334.94 ms, SD = 112.36 ms) and no hand condition, which was also significant ($t(25) = -2.63$, $p < 0.05$ (two-tailed), Cohen's d = −0.47, BF = 3.44). Finally, we compared the active and fake hand conditions. We expected to find longer duration estimates for fake hand conditions, which turned out to be true ($t(25) = 5.38$, $p < 0.001$ (one-tailed), Cohen's d = 1.06, BF = 3193.63), indicating that the alternative hypothesis is about 3194 times more likely than the null hypothesis.

2.6. Experiment 2

In Experiment 1, we conceptually replicated Experiment 1 of Suzuki et al. (2019). However, in contrast to the original experiment, we presented temporal intervals in blocks, to reduce uncertainty about interval durations and thereby reducing the regression to the mean, we observed in the results of the original study. In line with our hypothesis, the results of Experiment 1 did not replicate the original results reported in Suzuki et al. (2019). Instead, the results provided evidence for stronger binding in Active trials than in both N Hand trials as well as Fake Hand trials, while evidence for no difference between the latter two conditions was observed. In other words, in line with what would be expected according to the assumption that intentional actions are what drives temporal binding, we observed stronger temporal binding for the active condition as compared to both passive conditions.

To further support the idea that the discrepancies between Suzuki et al. (2019) and our study stem from the use of blocked interval durations, which likely reduced regression to the mean, we conducted a second experiment. Here we adhered to the original methodology of Suzuki et al. (2019), using a randomized and counterbalanced presentation of intervals. We hypothesized that the randomized presentation in Experiment 2 would lead to an increased regression to the mean, thereby diminishing the differences between Active and Fake Hand trials.

3. Material and methods

26 participants (5 men, mean age: 25.89 years, age range: 19–30, two left-handed) took part in Experiment 2. All participants had a normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. Before the start of the experiment, written informed consent was obtained following the Declaration of Helsinki. Participants subsequently received either monetary compensation or course credits. The study was approved by the ethics committee of the Faculty of Mathematics and Natural Sciences of the Heinrich-Heine-University Düsseldorf, Germany. Like in the previous experiment, the sample size was determined based on a Bayesian stopping rule. We started to collect data from 20 participants before we initially examined the evidence for or against each hypothesis. Data collection was stopped once substantial evidence was acquired supporting or against each experimental hypothesis across the pertinent comparisons.

3.1. Apparatus and stimuli

The experimental setup was identical to Experiment 1.

3.2. Procedure

The procedure was essentially identical to Experiment 1. However, in contrast to Experiment 1, the interval durations were presented in randomized order.

3.3. Results

The analysis revealed a significant main effect for *Condition* ($F(2,50) = 6.43$, $p < 0.01$, $\eta_p^2 = 0.20$) as well as for the factor *Interval* ($F(1.076,26.906) = 87.57$, $p < 0.001$, $\eta_p^2 = 0.78$, Greenhouse-Geisser corrected). The interaction was also significant ($F(2.919,72.98) = 4.03$, $p < 0.05$, $\eta_p^2 = 0.14$, Greenhouse-Geisser corrected) (Fig. 3).

When comparing the active condition (Mean = 315.13 ms, SD = 131.34 ms) and no hand condition (Mean = 346.01 ms, SD = 151.47 ms), we expected longer interval durations for the no hand condition, which was supported by the analysis ($t(25) = 2.53$, $p <$

0.01 (one-tailed), Cohen's d = 0.50, BF = 5.65). Hence, the alternative hypothesis is 5.65 times more likely than the H0 hypothesis. This finding is consistent with the results of Suzuki et al. (2019).

Next, we compared the fake hand (Mean = 350.74 ms, SD = 143.46 ms) and no hand conditions. This comparison turned out to be not significant (t(25) = −057, p = 0.57 (two tailed), Cohen's d = −0.11, BF = 0.24. Instead, the results indicate that the null hypothesis is 4.154 times more likely than the alternative hypothesis, suggesting no differences in interval estimation between fake hand and no hand trials.

Finally, we compared the active and fake hand conditions. Following the results of Suzuki et al. (2019), we expected to observe no difference between active and fake hand trials. Contrary to this hypothesis the analysis revealed significantly higher duration estimates for the fake hand condition (t(25) = 3.130, p < 0.01 (two-tailed), Cohen's d = 0.61, BF = 9.48), indicating that the alternative hypothesis is 9.478 times for likely than the null hypothesis.

Furthermore, to examine if the block interval durations in Experiment 1 successfully reduced the regression to the mean, a linear regression analysis was performed between the interval duration estimate and the actual interval duration for each participant and interval duration. The extracted slopes of each participant were then compared between Experiment 1 and Experiment 2 using a *t*-test for independent samples. The independent samples *t*-test revealed no significant difference between Experiment 1 (Mean = 0.492, SD = 0.230) and Experiment 2 (Mean = 0.57, SD = 0.225), (t(49) = −1.198, p = 0.237, Cohen's d = −0.34, BF = 0.51), compare which provides equivocal evidence, suggesting neither the null nor the alternative hypothesis is strongly supported by the data. In other words, the results suggest that the regression to the mean was not successfully reduced in Experimern1.

3.4. Experiment 3

In contrast to Suzuki et al. (2019), Experiment 1 and Experiment 2 both provided substantial evidence that estimates of temporal intervals are lower in the active than in the fake hand condition. However, in both experiments, we utilized the same set of pre-recorded hand animations for the fake hand condition for all participants. Conversely, Suzuki et al.(2019) employed hand animations that were individually recorded for each participant. Hence, this difference in experimental design was considered to be the root of the failed attempt to replicate Suzuki et al and to explain the different results. For instance, observing pre-canned animations in the fake hand trials, instead of observing animations of one's own hand movement, might have reduced the precision in predicting the timing of the button press (Lush et al., 2019). To address this, we conducted another experiment mirroring the methodology of Experiment 2. However, instead of using a standard set of pre-recorded animations, we recorded individual hand animations for each participant and used them in the fake hand condition. Consistent with the first two experiments of the present study, the results showed more pronounced temporal compression in the active condition compared to the fake hand trials, differing from the findings of Suzuki et al. (2019).

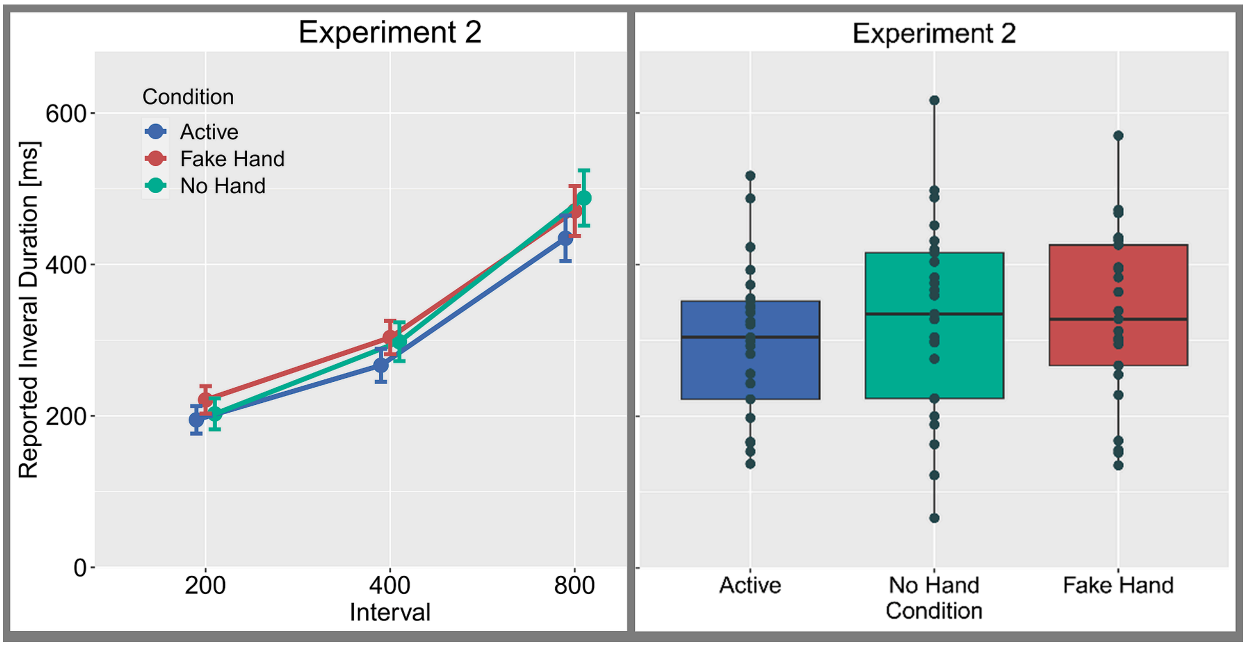

Fig. 3. Left: Average reported duration for each actual duration and condition in Experiment 2. Right: Average reported interval duration reported across the three conditions in Experiment 2. Error bars indicate the standard error of the mean.

4. Material and methods

26 participants (8 men, mean age: 26.8 years, age range: 18–30, one left-handed) took part in Experiment 3. All participants had a normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. Before the start of the experiment, written informed consent was obtained following the Declaration of Helsinki. Participants subsequently received either monetary compensation or course credits. The study was approved by the ethics committee of the Faculty of Mathematics and Natural Sciences of the Heinrich-Heine-University Düsseldorf, Germany.

Like in the previous experiments, the sample size was determined based on a Bayesian stopping rule. We started to collect data from 20 participants before we initially examined the evidence for or against each hypothesis. Data collection was stopped once substantial evidence was acquired supporting or against each experimental hypothesis across the pertinent comparisons.

4.1. Apparatus and stimuli

The experimental setup was identical to Experiment 1.

4.2. Procedure

The procedure was essentially identical to Experiment 2. The key difference was that, in Experiment 3, the ten practice trials were individually recorded for each participant and then randomly replayed during the fake hand trials. To minimize the chance of recording user errors, participants were trained on the hand movement required to press the button and return to the starting position before the experiment commenced. If a practice trial was compromised by a participant's or technical error, it was repeated to ensure accurate recordings.

5. Results

The ANOVA with the factors Condition (active, fake hand, no hand) and Interval (200 ms, 400 ms, 800 ms) revealed a significant main effect for the Condition ($F(2,50) = 4.83$, $p < 0.05$, $\eta_p^2 = 0.16$) and for the Interval factor ($F(1.12,28.07) = 121.30$, $p < 0.001$, $\eta_p^2 = 0.83$, Greenhouse-Geisser corrected). The interaction was not significant ($F(2.64,65.96) = 2.69$, $p = 0.06$, $\eta_p^2 = 0.10$, Greenhouse-Geisser corrected) (Fig. 4).

As anticipated, the no hand condition (Mean = 410.80 ms, SD = 104.82 ms) yielded longer interval estimations compared to the active condition (Mean = 377.77 ms, SD = 96.21 ms). This expectation was confirmed by the analysis ($t(25) = 2.51$, $p < 0.01$ (one-tailed), Cohen's $d = 0.49$, BF = 5.44), suggesting that the H1 hypothesis is 5.44 times more likely than the Null hypothesis.

Furthermore, the data provided substantial evidence for no difference between the fake hand (Mean = 409.64 ms, SD = 100.10 ms)

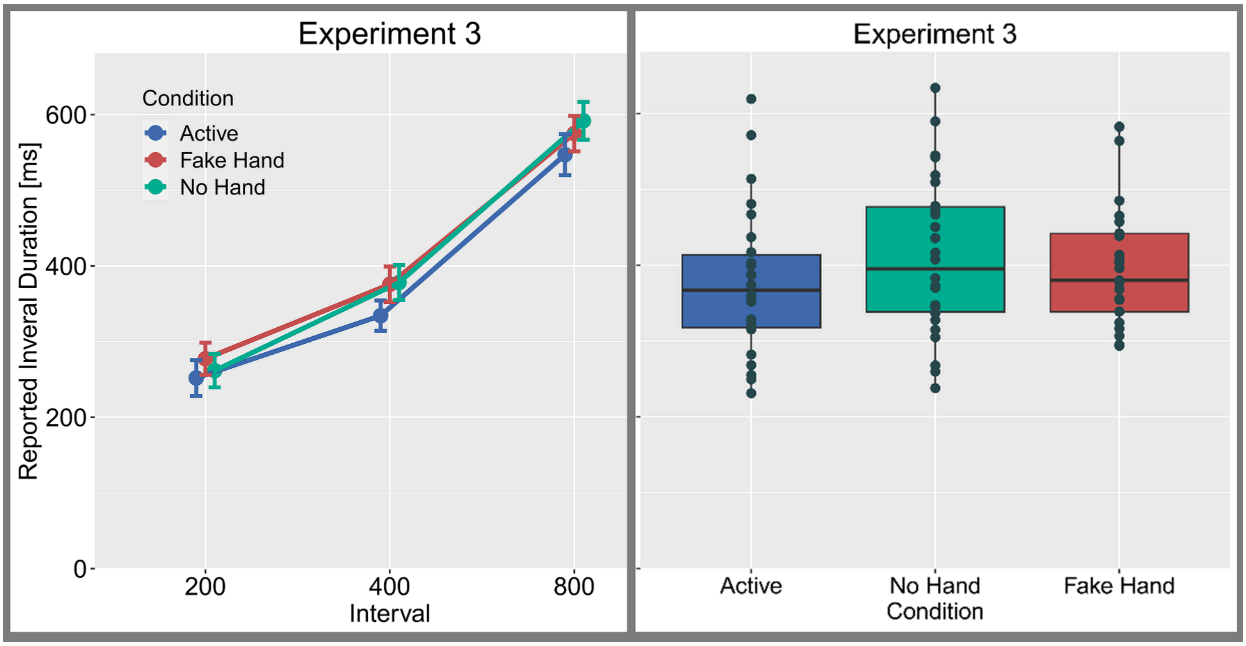

Fig. 4. Left: Average reported duration for each actual duration and conditions in Experiment 3. Right: Average reported interval duration reported across the three conditions in Experiment 3. Error bars indicate the standard error of the mean.

and no hand conditions ($t(25) = -0.11$, $p = 0.91$ (two-tailed), Cohen's $d = -0.02$, $BF = 0.21$).

Consistent with our previous experiments, the analysis revealed strong evidence for elevated interval estimates in the fake hand condition compared to the active condition ($t(25) = 2.72$, $p < 0.05$ (two-tailed), Cohen's $d = 0.53$, $BF = 4.11$). This indicates that the alternative hypothesis is 4.11 times more likely than the null hypothesis.

6. Discussion

The phenomenon of temporal binding, inducing a perceived temporal compression between an action and its outcome, has been a subject of interest in cognitive neuroscience. The traditional understanding posits that temporal binding is a direct result of the intentionality behind the action and has been taken as an implicit measure of the sense of agency (Moore & Obhi, 2012).

However, the idea that intention a critical for temporal binding has been challenged by different studies. The results of Buehner & Humphreys (2009) showed that intentional action that is not causal does not result in temporal binding, indicating that a causal relationship between an action and its effects is required. Similarly, unintentional mechanical causes have been found to result in temporal binding (Buehner, 2012). Other studies have also shed doubt on the link between intentionality and temporal binding (e.g., Buehner, 2015; Kirsch et al., 2019; Ruess et al., 2020), suggesting that instead of intentionality the perceived causal relationship between action and effects is what drives temporal binding effects.

In line with the above-discussed studies, a recent study by Suzuki et al. (2019) presented experimental evidence that temporal compression can be observed even in the absence of intentional actions. Leveraging the capabilities of modern virtual reality technology, they ensured consistent visual and tactile feedback while varying the intentional and motor activities associated with the button press. Beyond the conventional active–passive dichotomy, they introduced an additional passive scenario where participants witnessed pre-recorded button presses. Consequently, the distinction between active and fake hand trials lies in the intention to act and the actual motor performance. Visual and tactile feedback remained consistent across both conditions.

Suzuki et al. (2019) results provided substantial evidence against a difference between active hand movements and observed hand movements (fake hand), a result that contrasts with intentions being what drives temporal binding. While it's plausible to observe enhanced temporal binding in fake hand trials—given that participants viewed their own hand animations, potentially leading to an amplified sense of agency – one would expect a higher sense of agency in the active condition as compared to each of the passive conditions based on previous research on intentional binding (Hughes et al., 2013). Yet, Suzuki et al. (2019) present evidence supporting the absence of any disparity in interval estimates between active and fake hand trials. They conclude that intentional binding effects are the result of multisensory causal binding rather than being related to intentions or agency.

Upon examining the data from Suzuki et al. (2019), we noted that temporal compression was markedly pronounced for extended intervals and virtually non-existent for the 200 ms interval durations. Additionally, their findings displayed a distinct trend of regression to the mean. In prior research, we demonstrated that central tendency effects, such as regression to the mean, can mask temporal compression effects in intentional binding tasks (Zimmermann & Cicchini, 2020).

Hence, we speculated whether the regression to the mean might have led to an underestimation of the temporal binding effects reported in Suzuki et al. (2019).

To mitigate uncertainty regarding timing and subsequently minimize central tendency effects, we ordered temporal intervals in a block-wise manner. Consistent with prior studies on temporal binding, we detected a more pronounced temporal compression in active trials compared to no hand trials (citations needed). Contrary to the findings of Suzuki et al. (2019), our results consistently showed greater temporal compression for active trials than for fake hand trials.

To examine, if the blocked temporal intervals in Experiment 1 explain the contrasting results to Suzuki et al. (2019), we conducted another experiment. In Experiment 2, we followed Suzuki et al. (2019) and we randomized the sequence of temporal intervals on a trial-by-trial basis. Once more, our findings underscored the distinction between active and no hand trials. However, again we found substantial evidence for a difference between active and fake hand trials, further contrasting with the outcomes presented by Suzuki et al. (2019).

However, the results of the present study do not confirm the hypothesis that the regression to the mean accounted for the missing effect between active and fake hand trials in Suzuki et al. (2019). A comparison between the first two experiments presented here did not reveal any significant difference in regression to the mean. This suggests that our intervention was not strong enough to mitigate central tendency effects. Therefore, future studies should explore more robust strategies to counteract a regression to the mean.

A potential distinction between the original study and Experiment 2 of our research, which could account for our inability to replicate the initial findings, may lie in different implementations of the fake hand condition. In the study by Suzuki et al. (2019), the hand animations were individually recorded by each participant before the experiment. However, recording individual hand animations can introduce errors, either from the participant or due to technical glitches, which could then be replayed consistently during the experiment. This is particularly true for individuals unfamiliar with VR hardware, who often face challenges when using new VR input devices (Bachmann et al., 2014). Therefore, to sidestep potential confounds arising from difficulties in recording hand animations, we opted to use a set of 10 pre-recorded animations for all participants.

However, this difference might have resulted in a reduced precision of the timing information to predict the action onset (i.e., the button press), as suggested by Bayesian cue combination accounts on temporal binding (Lush et al., 2019). Consequently, participants might more accurately predict a hand animation that mirrors their typical movement behavior/speed compared to an arbitrary animation from another individual, which might result in distorted binding effects. This time, we closely mirrored the methodology of Suzuki et al. (2019) by recording individual hand animations for each participant before the experiment. Hence, we recorded individual hand animations of each participant prior to the experiment, which were then replayed randomly during fake hand trials. Great care was taken that no errors or irregularities were recorded. During the recording sessions, a researcher observed the participant, and any recordings with observed irregularities were redone.

However, regardless of this modification, the analysis of Experiment 3 provided substantial evidence for a stronger compression in active as compared to no hand trials. of a more pronounced compression in active trials compared to no hand trials. Likewise, the data revealed significant differences between active and fake hand trials, confirming the findings of the initial two experiments.

In summary, the experiments reported in the current study consistently replicated the effect between active and no hand trials, as reported by Suzuki et al. (2019). This comparison represents the typical passive vs active contrast, tested in many studies about temporal binding, that has been interpreted as evidence that intentional actions underpin the observed temporal compression (Moore & Obhi, 2012).

However, while we hypothesized to observe differences in temporal binding in Experiment 1, it was unexpected to also find evidence of enhanced temporal compression in active trials compared to fake hand trials in both a conceptual replication in Experiment 2 and a direct replication in Experiment 3. In both these experiments, the data consistently showed a more pronounced temporal compression in active trials, contrasting with the findings presented by Suzuki et al. (2019). Moreover, across all three experiments in our study, we observed similar magnitudes of temporal binding across the three interval durations. This stands in contrast to Suzuki et al., whose results suggest that binding predominantly influenced the longer interval durations, with minimal impact on the shorter interval duration.

What could account for the disparate results if not the regression to the mean? In Experiment 3, we meticulously adhered to the experimental protocol of the initial experiment detailed by Suzuki et al. (2019). We used the temporal intervals, trial number, and procedure as in the original study Thus, the primary distinctions were lying in the design and implementation of the virtual environment and the VR hardware. For example, the difference in rendering performance between both studies could affect the accuracy and precision of temporal measurements (Wiesing et al., 2020). Furthermore, performance problems and the resulting lag and irregularities are known to trigger motion sickness in many users (Kourtesis et al., 2023), and in consequence, reduce the task performance of participants. While we cannot evaluate the performance metrics of Suzuki et al. (2019), in our study, the virtual environment was meticulously optimized to consistently operate at the HTC Vive's native 90 Hz refresh rate.

Another relevant difference between both studies is the VR hardware employed. Here we used the HTC Vive Pro HMD, which is mostly identical to the Oculus Rift CV 1 used by Suzuki et al. (2019) apart from having the higher resolution. However, other technical specifications, such as the display's refresh rate, are consistent across both HMDs, making it improbable as the source of the contrasting outcomes. Yet, a potentially pivotal difference can be found when comparing the hand tracking devices used in each study. Here, we used Valve Index Controllers, which are motion controllers, attached to the hand by a strap system. The hand location is tracked with these controllers using the SteamVR 2.0 tracking system, which can provide robust spatial tracking with 360° coverage with sub-millimeter precision (Niehorster et al., 2017; Verdelet et al., 2019).

In Suzuki et al. (2019), the Leap Motion Controller (LMC) was used for tracking hand movements, utilizing infrared cameras on the HMD for contact-free tracking. Unlike Index Controllers that use a combination of inertial and optical tracking for 360° coverage, the LMC is limited to a 150° field of view and may lose tracking if hands leave this area. It also faces issues with finger occlusion and tracking discrepancies, especially for new users. Studies (Ganguly et al., 2021; Niechwiej-Szwedo et al., 2018) have noted LMC's spatial measurement inconsistencies, particularly in dynamic tracking. While the Index Controllers provide reliable spatial tracking, LMC is known for latency issues and underestimating movement duration by up to 13 % (Niechwiej-Szwedo et al., 2018), possibly due to irregular sampling rates. These factors might explain differences in findings between Suzuki et al. (2019) and our study, suggesting that irregular sampling in LMC could affect the precision of recorded hand animations.

Contemporary theories suggest that uncertainty in an action can skew intentional binding towards its effect, like a sound (Klaffehn et al., 2021; Lush et al., 2019; Reddy, 2022), potentially reducing binding effects. Notably, in the no hand condition, where the hand wasn't rendered, temporal tracking irregularities would only impact active and fake hand conditions, possibly explaining why Suzuki et al. (2019) saw no difference between these conditions. It's important to recognize that these theories, particularly those of Klaffehn et al., Lush et al., and Reddy, are based on the Libet clock task, different from the interval estimation task used by us and Suzuki et al. Recent findings by Siebertz and Jansen (2022) indicate that intentional binding in these two paradigms is uncorrelated, hinting that they may measure different processes and factors influencing binding. Thus, one tentative explanation for the challenges in replicating Suzuki et al.'s (2019) findings might be the potentially reduced temporal and spatial precision of the LMC, which could, in turn, affect the precision of information used to predict action onset (i.e., the button press). Future research should consider exploring the potential impact of hand tracking devices on temporal binding effects.

In conclusion, while the study by Suzuki et al. (2019) provided valuable insights and opened up avenues for further exploration, our findings suggest that the traditional understanding of intentional binding remains robust, and is tied to intentionality, rather than causality. The intentionality behind an action plays a pivotal role in the observed temporal compression between the action and its outcome. Future studies in this domain could further explore the nuances of this phenomenon and the various factors that influence it.

CRediT authorship contribution statement

Michael Wiesing: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Eckart Zimmermann:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

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

The dataset supporting the conclusions of this research is available in the OSF repository, accessible at https://osf.io/6x9sd/.

Acknowledgments

This research was supported by the European Research Council (project moreSense grant agreement n. 757184).

References

Bachmann, D., Weichert, F., & Rinkenauer, G. (2014). Evaluation of the leap motion controller as a new contact-free pointing device. *Sensors, 15*(1), 214–233. https://doi.org/10.3390/s150100214
Buehner, M. J. (2012). Understanding the past, predicting the future: causation, not intentional action, is the root of temporal binding. *Psychological Science, 23*(12), 1490–1497. https://doi.org/10.1177/0956797612444612
Buehner, M. J. (2015). Awareness of voluntary and involuntary causal actions and their outcomes. *Psychology of Consciousness: Theory, Research, and Practice, 2*(3), 237.
Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*(10), 1221–1228.
Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781.
EpicGames (2019). *Archviz Interior*. https://www.unrealengine.com/en-US/blog/new-archviz-interior-rendering-sample-project-now-available.
Ganguly, A., Rashidi, G., & Mombaur, K. (2021). *Can Leap Motion Controller replace conventional marker-based motion capture systems?. International Society of Biomechanics*.
Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382–385.
Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods, 7*(17), 461–469.
Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin, 139*(1), 133.
JASP Team (2023). *JASP (Version 0.18.0)[Computer software]*. https://jasp-stats.org/.
Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13*(8), 1020–1026.
Jung, S., Bruder, G., Wisniewski, P. J., Sandor, C., & Hughes, C. E. (2018). Over my hand: Using a personalized hand in vr to improve object size estimation, body ownership, and presence. In *Proceedings of the 2018 ACM Symposium on Spatial User Interaction* (pp. 60–68).
Kirsch, W., Kunde, W., & Herbort, O. (2019). Intentional binding is unrelated to action intention. *Journal of Experimental Psychology: Human Perception and Performance, 45*(3), 378–385. https://doi.org/10.1037/xhp0000612
Klaffehn, A. L., Sellmann, F. B., Kirsch, W., Kunde, W., & Pfister, R. (2021). Temporal binding as multisensory integration: Manipulating perceptual certainty of actions and their effects. *Attention, Perception, & Psychophysics, 83*(8), 3135–3145.
Knoetsch, F., & Zimmermann, E. (2021). The spatial specificity of sensory attenuation for self-touch. *Consciousness and Cognition, 92*, Article 103135.
Kourtesis, P., Linnell, J., Amir, R., Argelaguet, F., & MacPherson, S. E. (2023). Cybersickness in virtual reality questionnaire (csq-vr): A validation and comparison against ssq and vrsq. *Virtual Worlds, 2*(1), 16–35.
Lush, P., Roseboom, W., Cleeremans, A., Scott, R. B., 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.
Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition, 21*(1), 546–561.
Niechwiej-Szwedo, E., Gonzalez, D., Nouredanesh, M., & Tung, J. (2018). Evaluation of the Leap Motion Controller during the performance of visually-guided upper limb movements. *PloS one, 13*(3), Article e0193639.
Niehorster, D. C., Li, L., & Lappe, M. (2017). The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research. *I-Perception, 8*(3), Article 2041669517708205.
Ogawa, N., Narumi, T., & Hirose, M. (2019). Virtual hand realism affects object size perception in body-based scaling. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)* (pp. 519–528).
R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.
RStudio Team (2020). *RStudio: Integrated development environment for R*. RStudio, PBC. http://www.rstudio.com/.
Reddy, N. N. (2022). The implicit sense of agency is not a perceptual effect but is a judgment effect. *Cognitive Processing, 23*(1), 1–13.
Ruess, M., Thomaschke, R., & Kiesel, A. (2020). Intentional Binding for Unintended Effects. *Timing & Time Perception, 8*(3–4), 341–349. https://doi.org/10.1163/22134468-bja10005
Siebertz, M., & Jansen, P. (2022). Diverging implicit measurement of sense of agency using interval estimation and Libet clock. *Consciousness and Cognition, 99*, 103287.
Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional binding without intentional action. *Psychological Science, 30*(6), 842–853.
Verdelet, G., Desoche, C., Volland, F., Farnè, A., Coudert, A., Hermann, R., Truy, E., Gaveau, V., Pavani, F., & Salemme, R. (2019). Assessing spatial and temporal reliability of the vive system as a tool for naturalistic behavioural research. In *2019 International Conference on 3D Immersion (IC3D)* (pp. 1–8).
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.
Wiesing, M., Fink, G. R., & Weidner, R. (2020). Accuracy and precision of stimulus timing and reaction times with Unreal Engine and SteamVR. *PloS One, 15*(4), e0231152.
Wolpe, N., Haggard, P., Siebner, H. R., & Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research, 229*, 467–474.
Zimmermann, E., & Cicchini, G. M. (2020). Temporal context affects interval timing at the perceptual level. *Scientific Reports, 10*(1), 1–10.