Testing Temporal Binding and Memory Accounts: A Comparison of Interval and Line Length Judgments

Pernille Hemmer[1], Fatemeh Mahdinia[2], Sean Duffy[2], Robrecht van der Wel[2]
[1]Rutgers University, New Brunswick
[2]Rutgers University, Camden

Abstract

Temporal binding (TB), the perceived compression in time between an action and its effect versus an observed action and its effect, has been viewed as an implicit measure of sense of agency (SoA). SoA refers to the feeling of control over our actions and their outcomes. Recent work has raised methodological and theoretical concerns about the TB paradigm. Most problematically, the TB effect may be explained by a regression-to-the mean pattern commonly observed as a memory bias, rather than as an effect related to SoA. We tested this possibility using two structurally equivalent tasks. Participants completed a standard TB task and a well-established memory task (line length estimation) that we equated on all aspects to allow for direct comparison. We did not replicate the TB effect; instead, both the time interval and line length conditions showed classic regression-to-the-mean memory effects, consistent with predictions from the Category Adjustment Model (Huttenlocher, 2000).

Keywords: sense of agency; temporal binding; memory; explicit measures; metacognition

Introduction

The sense of agency (SoA), broadly defined as the feeling of control over one's voluntary actions and their associated outcomes, is a fundamental aspect of human experience. When individuals act intentionally, they typically experience themselves as the authors of those actions and their effects. This experience plays a critical role in motor control, social interaction, development, and clinical conditions such as schizophrenia.

Over the past two decades, cognitive science research has converged on temporal binding (TB) as the primary implicit measure of SoA. TB refers to the subjective compression of the temporal interval between a voluntary action (e.g., pressing a button) and its outcome (e.g., hearing a tone). Actions are perceived as occurring later than they objectively do, and outcomes are perceived as occurring earlier, resulting in a compressed interval. In contrast, when outcomes occur without voluntary action, intervals are often perceived as longer. The difference in time perception during an active action versus passive condition is referred to as the TB effect.

TB effects have been widely replicated and are broadly considered to be an implicit reflection of SoA (Hughes et al., 2013; Moore & Obhi, 2012). To date, accounts of TB have implicated sensorimotor mechanisms (Haggard et al., 2002), cognitive-level inferences about causality (Buehner & Humphreys, 2009; Hoerl et al., 2020), or a combination of these cues (Moore et al., 2009). Another interpretation has been that time estimates of actions and effects are based on a weighted average of all relevant sources of information, where the weighting is done by the estimated reliability of each information source (e.g., Suzuki et al., 2019). In this processing, top-down perceptual predictions and bottom-up sensory prediction errors together feed into SoA. While these accounts differ in their specifics, each of them focuses on the role of different information sources as they contribute to SoA.

However, an important and increasingly scrutinized feature of TB paradigms is that they are also fundamentally memory and time estimation tasks. For example, Gutzeit et al. (2023) controlled for a procedural confound between active and passive conditions and found that this eliminated the TB effect. They theorized that TB is a more basic temporal grouping effect. In addition, TB tasks are inherently memory tasks: participants must *recall* and reproduce temporal intervals after experiencing them. A growing body of work demonstrates that the behavioral signatures observed in TB tasks closely resemble regression-to-the-mean patterns well documented in episodic memory research. Saad et al. (2022) showed that when the temporal binding data replotted to show the error (binding) for each observed time interval the classic memory pattern emerged. According to the category adjustment model (CAM; Huttenlocher, 2000), estimates of stimulus magnitude are biased toward the mean of the experienced stimulus distribution, producing overestimation of short intervals and underestimation of long intervals. Critically, these patterns can emerge independently of the objective level of control people have in a task (e.g., Kirsch, Kunde, & Herbort, 2019). Thus, a core question for understanding the value of TB as a measure of particular cognitive processes, it is important to delineate to what extent TB reflects episodic memory, processes involved in the planning and execution of intentional actions (from which the SoA may then be derived), or both.

To adjudicate between agency-based and memory-based interpretations of the TB effect, the present work directly compares TB with a structurally matched line-length estimation task. Both tasks involve an initiating action, a target magnitude (time or space), and a continuous reproduction response. The line-length task is canonically interpreted as reflecting memory bias rather than agency. If TB uniquely reflects SoA, then agency-specific effects, i.e., stronger compression for active than passive trials, should emerge selectively in the tone task. Conversely, if both tasks

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

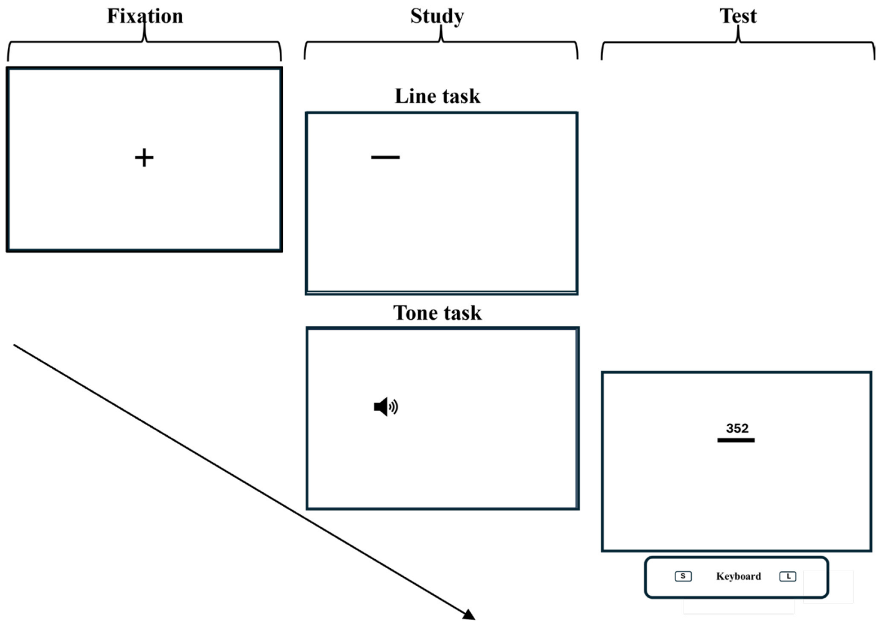

Figure 1. Task illustration. At the fixation mark either participants pressed a button (active) or the trail stared automatically (passive). At study in the line task a line appeared on the screen and in the tone task there was a blank screen and a tone sounded after an interval. A test a line appeared, and participants manipulated the length using the ‘s’ and ‘l’ keys on the keyboard.

show similar bias patterns consistent with regression-to-the-mean, this would favor a memory-based account and it would be unclear if and why there would be a reason to postulate that the TB effect also reflects SoA.

In addition, we assessed explicit ratings of perceived control to examine whether the subjective sense of agency tracks objective control manipulations and whether it aligns with implicit bias measures. Finally, in Experiment 2, we evaluated the temporal stability of explicit and implicit measures across a one-week interval. If TB reflects a core aspect of SoA, it should exhibit trait-like stability over time.

Experimental Setup

Experiment 1

Participants and design

Thirty-eight participants at Rutgers-Camden consented to participate and received psychology course credit for their participation. The study was approved by the Rutgers University Institutional Review Board.

The study employed a within-subjects design with two factors: Task (Line Length Estimation vs. Tone Interval Estimation) and Condition (Active vs. Passive). In both tasks, participants either triggered the stimulus themselves (Active) or observed the stimulus when it was generated automatically by the computer (Passive). Each task included three target magnitudes; short, medium, and long (=300, 500, and 700 units; pixels for lines and milliseconds for tones). The order of tasks was counterbalanced across participants.

The participants completed 3 blocks of 9 for a total of 27 trials in each condition (amounting to a total of 108 trials). We based this number and the number of participants on previous line length estimation and temporal binding studies. Participants experienced all 4 conditions of the experiment (Line Passive, Line Active, Tone Passive, Tone Active). The order of the conditions varied by subject based on their subject ID number.

Materials and Procedure

The experiment was implemented in E-Prime (Psychology Software Tools, 2025). At the beginning of all trials a fixation point appeared at the center of the screen. In the active conditions, the fixation cross disappeared when participants pressed the spacebar. In the line task, they then saw a reference line. Reference line stimuli consisted of horizontal black lines (height = 10 pixels) presented on a gray background at three lengths (300, 500, and 700 pixels). The reference line disappeared after 1.5 seconds and was followed by an adjustable line. To avoid possible anchoring effects, the initial length of the adjustable line was drawn from a uniform distribution between 0 and 1000 pixels. Participants needed to adjust this line to match the length of the reference line, using the S and L keys to make it shorter or longer. Once they were satisfied with their reproduction, they pressed the spacebar to continue.

In the tone condition, participants matched the duration between the disappearance of the fixation cross and a tone. Tone stimuli consisted of pure 800 Hz tones that played at

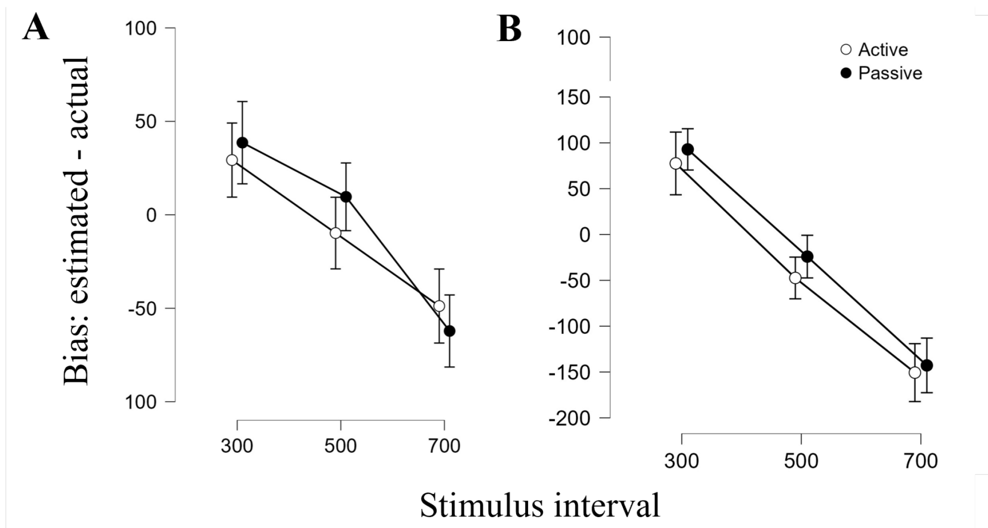

Figure 2. Bias for line (panel A) and tone (panel B) tasks as a function of active/passive and stimulus interval.

300, 500, and 700 ms (a silent interval) after the disappearance of a fixation cross. Just like the line task, the same randomly generated horizontal adjustable line then appeared on the screen after the tone played. A number appeared above the line indicating the “duration estimate”, for example 352 ms. Participants adjusted the line’s length to represent the length of the delay between the pressing action and the tone.

The passive conditions for both tasks were identical, except that the fixation cross disappeared by itself after two seconds (thus not requiring an action to start the task).

After each block of 9 trials, participants answered an explicit sense-of-agency question: “*How much control did you feel over the past block*?” on a Likert scale from 1 - 9 with 1 indicating no control and 9 indicating complete control.

Experiment 2

The goal of E2 was to 1) replicate the findings from Experiment 1, and 2) test for the stability across time in subjective bias. If TB is in fact an implicit measure of SoA this would be expected to be very stable (i.e., an individual trait), such that measurements of TB within an individual should positively correlate over time. In addition, if our line task reflects memory processes (Huttenlocher et al. 2000) while our tone task reflects SoA (or at least something other than memory, such as Vierordt’s law (Lejeune & Waerden, 2009), then we should not find strong correlations between tasks (either within or across sessions). Positive correlations would suggest that similar mechanisms underly these two tasks.

All materials and procedures were identical to Experiment 1 (tone × line × active × passive), but participants completed two sessions one week apart. Experiment 2 included thirty-two participants (25 females).

Results

We separately present the results of Experiments 1 and 2 about bias, slopes, and the explicit ratings first and then include an integrated section to report on correlations across tasks (for both experiments) and sessions (for Experiment 2).

Experiment 1

Implicit Bias Measures

For both the line and tone tasks, reproduction error was calculated as the difference between estimated and actual magnitudes (error = estimated – actual). For the tone task (the presumed agency task) this is referred to as Temporal Binding (TB). For the line task (the presumed memory task) this is referred to as bias. Positive values indicate overestimation (or repulsion), while negative values reflect underestimation (or compression consistent with TB). The temporal binding **effect** results from differences in compression across active and passive conditions.

We first analyzed TB biases as implicit indicators of agency. If TB was present, the tone task should show a clear pattern where Active bias (i.e., compression) is greater than Passive bias. In other words, the intervals when participants acted should be perceived as shorter. Figure 2 shows the results. In contrast to the TB prediction, the results failed to show a main effect of active vs passive ($p > 0.10$), or an interaction between task (line/tone) and condition (active/passive), ($p > 0.10$).

Instead, both line and tone tasks exhibited strong regression-to-the-mean patterns ($F(2, 80) = 166.22$, $p < 0.001$, $\eta^2 = 0.412$): short magnitudes were overestimated, and long magnitudes were underestimated. Note that this is problematic for the tone task, because overestimation is repulsion, and is not the predicted effect for SoA.

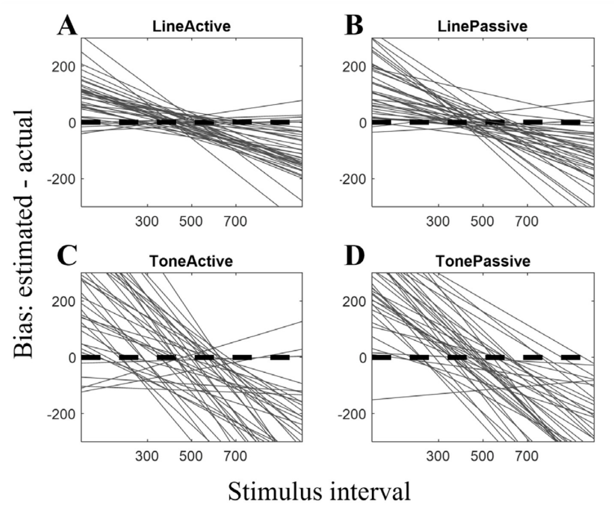

Figure 3. Slopes for each participant in the line task, active (panel A) and passive (panel B) and tone task, active (panel C) and passive (panel D).

Slope Measures

For both the line and tone tasks, to assess regression-to-the-mean, a linear regression model was first fitted to each participants responses across the stimulus intervals 300, 500, and 700 (seconds or pixels) to obtain a slope measure. This is an often-used measure in the memory literature, used to evaluate category bias. A negative slope different from zero would indicate a regression-to-the-mean effect, where values below the mean are overestimated and values above the mean are underestimated.

We found that slopes were negative across both tasks and conditions, consistent with a regression-to-the-mean effect (all $p< 0.001$). We then analyzed the difference in slopes between conditions (see Figure 3) to determine if there was a difference in regression between the active and passive conditions and between the line and tone tasks. There was a strong main effect of task ($F(1,37) = 59.093$, $p < 0.001$, $\eta^2 = 0.491$), such that the slopes were steeper for the tone task – meaning there was a greater influence of the category mean. There was no main effect of active/passive, meaning we did not observe a difference in the influence of the category means whether the participant pressed the button or passively observed the trial. Lastly, there was no interaction. These results are consistent with memory-based category adjustment.

Explicit Agency Ratings

Explicit ratings (Figure 4) were analyzed using repeated-measures ANOVAs. Participants reported significantly higher control in the active condition than in the passive condition, $F(1,73) = 5.818$, $p = 0.018$, $\eta^2 = 0.03$, indicating that explicit SoA tracked the objective manipulation of control. There was also a main effect of task, as participants reported higher control ratings for the line task than the tone task ($F(1,73) = 34.938$, $p < 0.001$, $\eta^2 = 0.149$). The interaction was not significant. This is consistent with a SoA explanation where people feel stronger SoA over active than passive actions.

Experiment 2

Implicit Bias Measures

Figure 5 shows the bias results obtained for Experiment 2. As in Experiment 1, the results failed to provide evidence for agency-specific temporal binding. Bias patterns in both tone and line tasks were dominated by regression-to-the-mean effects and were highly similar across sessions. Specifically, a 2 (session) x 2 (task: line/tone) x 2 (condition: active/passive) x 3 (stimulus interval: 300/500/700 ms/pixels) repeated-measures ANOVA indicated a main effect of stimulus interval ($F(2,74) = 168.507$, $p < 0.001$, $\eta^2 = .0.214$), such that short intervals were overestimated and long intervals were underestimated. There was also a main effect of session ($F(1,37) = 18.124$, $p < 0.001$, $\eta^2 = 0.024$), and an interaction between session and interval ($F(2,74) = 5.813$, $p < 0.01$, $\eta^2 = .002$). These results indicate that participants became more accurate (i.e., less bias) for each stimulus interval, except for the shortest one between sessions. The stability of the regression-to-the-mean bias patterns further supports a memory-based interpretation. Importantly, the results failed to show evidence for the temporal binding effect, as there was no significant main effect or interaction for the active versus passive condition, all $p > 0.10$.

Slope Measures

As in exp 1, for both the line and tone tasks, to assess regression-to-the-mean, a linear regression model was first fitted to each participants responses across the stimulus

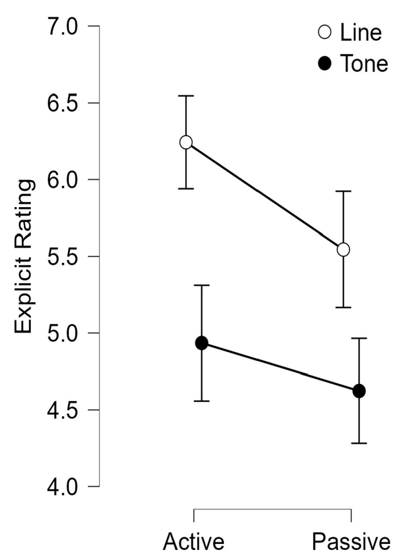

Figure 4. SoA ratings as a function of the active and passive conditions for the line and tone tasks in Experiment 1.

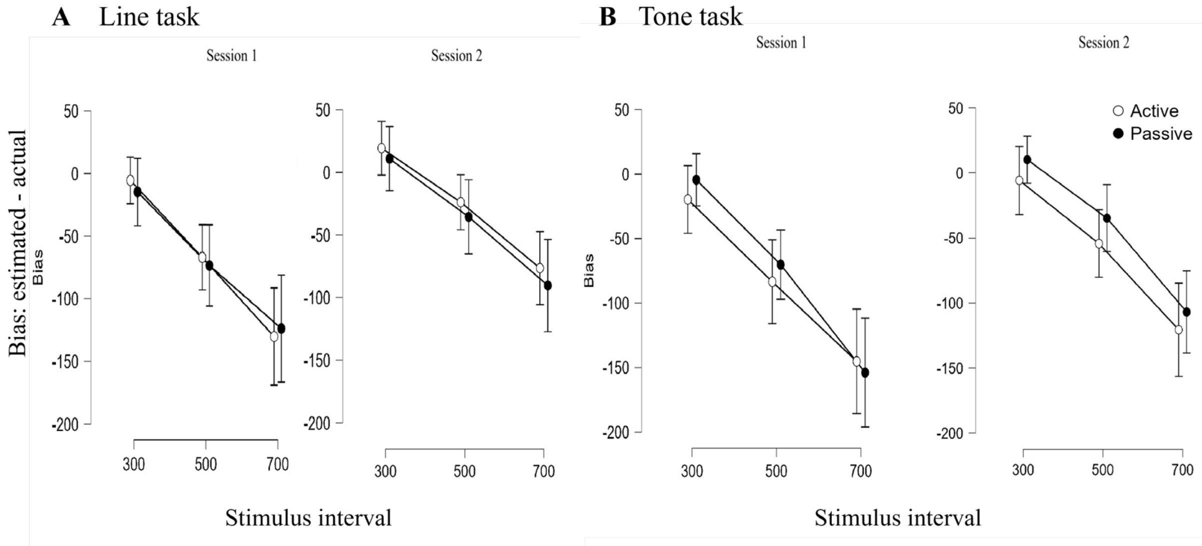

Figure 5. Bias for the line task (panels A) and tone task (panels B) as a function of active/passive, stimulus interval and session for experiment 2.

intervals 300, 500, and 700 (seconds or pixels) to obtain a slope measure. We found that slopes were negative across both tasks and conditions in both sessions 1 and 2, consistent with a regression-to-the-mean effect (all $p<0.001$). We then analyzed the difference in slopes between conditions and sessions to determine if there was a difference in regression between the active and passive conditions and between the line and tone tasks. For session 1 there was a strong main effect of task ($F(1,37) = 46.999$, $p < 0.001$, $\eta^2 = 0.416$), such that the slopes were steeper for the tone task – meaning there was a greater influence of the category mean. For session 2 There was a main effect of task ($F(1,37) = 15.405$, $p < 0.001$, $\eta^2 = 0.229$), such that the slopes were steeper for the tone task – meaning there was a greater influence of the category mean. There were no main effect of active/passive and no interaction.

Lastly, we compared slopes between sessions (see Figure 6). For the line active/passive condition for session 1 vs 2 there were no main effects or interaction. We confirmed the lack of main effects with a Bayesian repeated measures ANOVA, finding in favor of the null of no difference between sessions BF10 = 0.209 or condition BF10 = 0.287. For the tone active/passive condition for session 1 vs 2 there was a small main effect of session ($F(1,37) = 9.502$, $p = 0.004$, $\eta^2 = 0.103$), such that there was more regression in session 1 than session 2. This could be interpreted as participants learning the task and performing better (more accurate recall) in session 2. There were no main effect of condition (active/passive) and no interaction.

Explicit Agency Ratings

Like the bias results, we analyzed the explicit ratings with a 2 (session) x 2 (task: line/tone) x 2 (condition: active/passive) x 3 (stimulus interval: 300/500/700 ms/pixels) repeated-

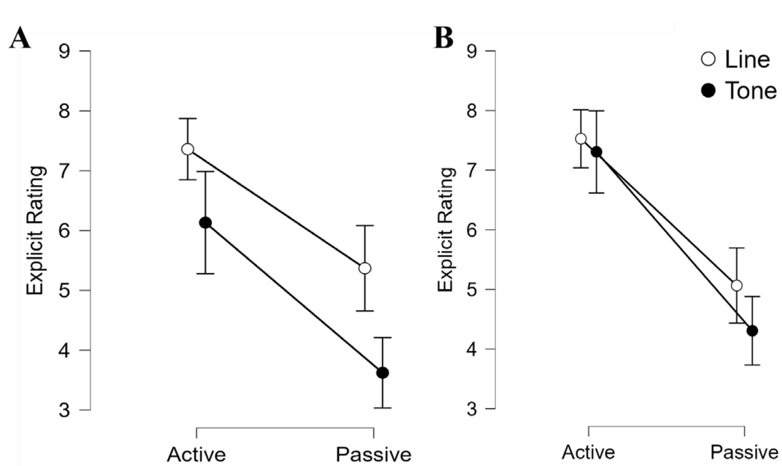

Figure 6. SoA ratings for the active and passive conditions in the line and tone tasks for sessions 1 (top) and 2 (bottom).

measures ANOVA. Figure 7 shows the results. Explicit ratings again showed a robust effect of condition, $F(1,37) = 134.372$, $p < 0.001$, $\eta^2 = 0.482$), such that participants provided higher ratings in the active versus passive condition. The results also indicated main effects for session, $F(1,37) = 4.793$, $p < 0.05$, $\eta^2 = 0.115$), and task, $F(1,37) = 16.419$, $p < 0.001$, $\eta^2 = 0.307$). There was also a two-way interaction between session x task, $F(1,37) = 10.191$, $p < 0.01$, $\eta^2 = 0.216$), and between task x condition, $F(1,37) = 4.486$, $p < 0.05$, $\eta^2 = 0.108$). As can be seen in Figure 6, the effect of task was present for session 1 but not for session 2, and the difference between the active and passive conditions was more pronounced for the tone task than the line task.

Correlations across tasks and sessions

To better understand how performance in the active and passive conditions in the line and tone tasks related to each other as a stable feature of individuals, we tested for correlation within and between tasks. For each person we find the best fitting line across time intervals given active/passive and line/tone. This results in a slope measure for each participant for each condition in each task. We then correlate these slope measures.

For Experiment 1, we found that performance on the active and passive condition slopes correlated strongly within task (Line task active/passive: $r = 0.511$, $p = 0.008$, Tone task active/passive: $r = 0.669$, $p < 0.001$). However, the results failed to show a correlation of slopes between tasks for either the active condition (line/tone active: $r = 0.131$, $p = 0.524$) or the passive condition (line/tone passive: $r = 0.359$, $p = 0.072$).

We conducted the same analysis for sessions 1 and 2 for Experiment 2. For session 1, we found a similar pattern similar to Experiment 1; For the Line task there was a weak but non-significant correlation between the active and passive conditions ($r = 0.281$, $p = 0.088$) and a strong correlation between performance in the active/passive conditions in the Tone task ($r = 0.526$, $p < 0.001$), but no correlation across tasks (line/tone active: $r = 0.109$, $p = 0.517$, line/tone passive: $r = -0.016$, $p = 0.925$). For session 2, there was a strong correlation in slopes within tasks (line active/passive: $r = 0.592$, $p < 0.001$, tone active/passive: r = 0.718, $p < 0.001$), and no correlation of slopes across tasks (line/tone active: $r = 0.109$, $p = 0.514$, line/tone passive: $r = 0.162$, $p = 0.331$).

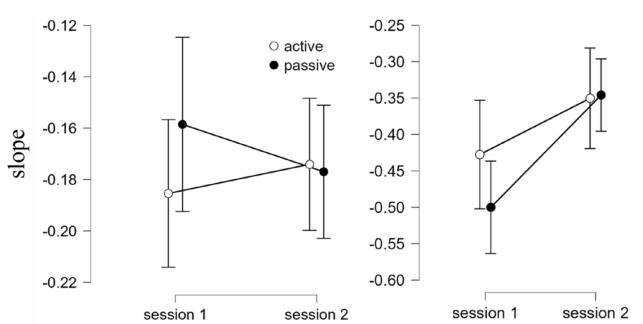

Figure 7. Slopes for session 1 and 2 for line task (left panel) active/passive and tone task (right panel) active/passive.

Lastly, we correlated performance within each task across sessions to test for stability. The results indicated that performance within tasks appeared to be stable, such that there was a correlation between session 1 and 2 slopes in both the Line task (active: $r = 0.659$, $p < 0.001$, passive: $r = 0.331$, $p = 0.043$) and the Tone task (active: $r = 0.365$, $p = 0.024$, passive: $r = 0.498$, $p = 0.002$).

Discussion

Here, we introduced a task design intended to directly compare the temporal binding effect – which is thought to be an implicit measure of Sense of Agency - with a structurally matched line-length estimation task to adjudicate between agency-based and memory-based interpretations of TB.

First, both tone and line tasks exhibited highly similar regression-to-the-mean patterns characteristic of episodic memory processes. When we considered bias (i.e., the difference between estimated and actual durations and line lengths), our results showed the exact pattern predicted by the category adjustment model, namely that short durations and lines were overestimated and long durations and lines were underestimated. These effects were robust, consistent across active and passive conditions, and stable over time within each task (but not across tasks). To our knowledge, we are the first to show stability in the slopes of regression lines over sessions for temporal binding data. Importantly, however, our results did not show the predicted differences in compression between the active and passive conditions for the tone (or the line) task. Binding itself (not the TB effect) was only present for the longest intervals. Together, these results strongly suggest that what is often labeled as temporal binding may reflect scaling mechanisms related to time perception or memory rather than an implicit sense of agency. With respect to memory, our results suggest that sensitivity of the priors people constructed for spatial (lines) and temporal (tones) intervals in our experiments do not clearly relate to one another.

Importantly, across the two experiments, explicit ratings consistently tracked objective control and were stable across time, supporting their validity as measures of SoA. In contrast, implicit bias measures derived from tone interval estimation - the standard paradigm for temporal binding - did not show the predicted agency-specific signatures, namely greater binding for active than for passive trials.

These findings have important implications for the interpretation of the TB effect in cognitive science and related applied domains. While TB paradigms remain valuable tools for studying time perception and memory, their use as implicit measures of the sense of agency should be reconsidered. Future work should focus on developing tasks and measures that more directly capture the online control and prediction processes central to agency, and on clarifying how different components of agency relate to distinct cognitive mechanisms.

References

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

Gutzeit, J., Weller, L., Kürten J., and Huestegge, L. (2023). Intentional binding: Merely a procedural confound? *Journal of Experimental Psychology. Human Perception and Performance*, doi.org/10.1037/xhp0001110

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

Hoerl, C., Lorimer, S., McCormack, T., Lagnado, D. A., Blakey, E., Tecwyn, E. C., & Buehner, M. J. (2020). Temporal Binding, Causation, and Agency: Developing a New Theoretical Framework. *Cognitive Science*, *44*. doi.org/10.1111/cogs.12843

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*, 133–151. doi.org/10.1037/a0028566

Huttenlocher, J., Hedges, L.V.,&Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, 129, 220-241. doi.org/10.1037/0096-3445.129.2.220

Kirsch, W., Kunde, W., & Herbort, O. (2019). Intentional binding is unrelated to action intention. *Journal of experimental psychology. Human perception and performance*, *45*, 378–385. doi.org/10.1037/xhp0000612

Lejeune, H., & Wearden, J. H. (2009). Vierordt's *The Experimental Study of the Time Sense* (1868) and its legacy. *European Journal of Cognitive Psychology*, *21*(6), 941–960. doi.org/10.1080/09541440802453006

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*, 546–561. doi10.1016/j.concog.2011.12.002

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, *18*, 1056–1064. doi.org/ 10.1016/j.concog.2009.05.004

Psychology Software Tools, Inc. (2025). E-Prime (Version 3.0). https://pstnet.com/products/e-prime/.

Saad, L., Musolino, J., & Hemmer, P. (2022). Bayesian rational memory model simulates temporal binding effect. *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society, (pp. 3549– 3556).

Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional Binding Without Intentional Action. *Psychological Science*, *30*, 842–853. doi.org/10.1177/ 0956797619842191