Sense of Agency: Towards Empirically Driven Measures and Understanding

Fatemeh Mahdinia, Amir Lindor, Pernille Hemmer, & Robrecht van der Wel
Rutgers University

Abstract

Sense of Agency (SoA) is a core concept related to our experience as intentional agents in our environment. Explicit and implicit measures have been used to study SoA. Recent findings suggest that the most common implicit measure, namely Temporal Binding (TB), may reflect memory processes rather than SoA. Here, we implemented two TB measures and an explicit measure in a novel goal-directed extended action task to better understand SoA measures. Participants either watched or produced dot movements to a target of choice and then estimated the duration between two tones that played either upon movement completion (TB1, akin to traditional TB studies) or based on the start and end of movements (TB2). Participants reported stronger explicit SoA during active than passive movements. Results from neither TB version aligned with prediction based on TB-accounts as a reflection of SoA. We discuss memory-based and scaling accounts as alternative interpretations for our data.

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

Introduction

The sense of agency (SoA), i.e., the feeling of control over our actions and their associated outcomes, is one of the most fundamental aspects of human experience. The actions we perform allow us to experience a sense of ownership and control over the changes our actions bring about in the environment. For example, if I reach for a glass of water, I feel a sense of ownership over my actions as I grasp the glass and drink from it. View this in contrast to accidentally knocking a glass of water on the floor–I caused the outcome but might not have controlled it. Our ability to veridically experience this link between our actions and their outcomes is a core aspect of child development and has been linked to several clinical disorders, e.g., schizophrenia. SoA is also thought to play a role in voluntary movement. Notably, a better understanding of SoA has important implications for neurorehabilitation and socially cognizant robotics.

While there is a wealth of research on SoA, it has recently come to the fore that the most common implicit measure of SoA (temporal binding (TB), see below), may not reflect SoA. To address this issue, we introduce and present data on a novel approach to understanding common measures of SoA (including TB). We do so by introducing a goal-directed extended action task paradigm. Before we describe this task, we will first provide background of the most common theories of SoA.

Metacognition of Action

What do we know about SoA as a metacognitive phenomenon? SoA consists of several components (Gallagher, 2012). It includes both judgments of agency (JoA) as well as feelings of agency (FoA). The term *judgment of agency* refers to the notion of being the one who initiated an action, whereas the term *feeling of agency* refers to the notion of experiencing control over the action (see Haggard & Tsakiris, 2009; Pacherie, 2008; Synofzik et al., 2008). These concepts are linked, as people report lower FoA when they judge that they did not initiate an action as compared to when they did. For a person to judge that they initiated an action does not imply that they feel in control over it, however. For example, one could imagine walking into a classroom and flipping a light switch to turn on the lights. In this case, one would probably feel in control over the lights turning on, unless someone else happened to simultaneously flip another light switch while they entered the classroom through another door (Silver et al., 2021). Thus, the link between initiating an action and experiencing SoA over its effect is sometimes ambiguous. Becoming aware of these ambiguities may influence SoA after action completion (i.e., postdictively).

While SoA is a core psychological phenomenon, it is not straightforward to understand or measure. Predictive and postdictive accounts have been developed for how people experience SoA. The *predictive account* (e.g., Blakemore et al., 2002; Haggard, 2005; Tsakiris et al., 2006) postulates that SoA arises from the match between the predicted and actual sensory consequences of an action. SoA then increases as this match gets stronger.

The *postdictive account* (e.g., Hoerl et al, 2020) is most clearly represented by the theory of apparent mental causation. Whereas the predictive account establishes SoA dynamically during the action, the postdictive account claims that SoA is established *after* the action is completed by evaluating the extent to which three criteria are met – namely priority, consistency, and exclusivity. The light switch example above forms a straightforward way of illustrating these criteria. *Priority* dictates that a thought needs to precede an action (i.e., the thought of turning on the lights). *Consistency* dictates that the thought needs to be consistent with the action outcome (i.e., the lights turning on). *Exclusivity* dictates that no alternative causes for the action are perceived or known (i.e., realizing whether someone else flipped a light switch or not). Any reduction of these criteria then lowers SoA.

There is broad consensus that the predictive and postdictive accounts are complementary rather than conflicting in nature. In fact, the *Cue Integration Theory* integrates predictive and postdictive aspects of SoA based on Bayesian mechanisms (Legaspi & Toyoizumi, 2019; Moore & Fletcher, 2012; Synofzik et al., 2013). According to this theory, the derivation of SoA arises based on a weighting of multiple cues. These cues could be sensorimotor cues or other predictive and postdictive cues.

Measuring SoA Implicitly: Temporal Binding

Interestingly, SoA may not just be reactive to the integration of predicted and actual signals but may also modulate this integration. Evidence for this assertion stems from the phenomenon of temporal binding (or action-effect binding). Temporal binding (TB) is the most considered *implicit* measure of SoA. TB refers to the observation that one's own actions and action effects are bound together in time, i.e., a subjective compression of the time interval between the action (e.g., pressing a button) and its effect (e.g., hearing a tone; Haggard et al., 2002). To be precise, actions people perform are perceived later in time than they objectively occur, and action effects are perceived earlier in time than they objectively occur. In contrast, passively produced action effects as well as action effects produced by others are temporally separated from the actions that produced them. To emphasize the link to intentionality, TB is sometimes referred to as intentional binding (Haggard, 2017).

TB effects have been widely replicated and are frequently conceptualized as 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). TB may also be a specific example of a more general process of causal binding across time and space rather than being specifically indicative of intentionality. When seen as a form of Bayesian predictive processing, the notion is 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. Considering SoA to arise from this process is different from the traditional TB account, as it does not require postulations about intentionality. Thus, it is important to carefully examine how TB relates to SoA.

This last point is of critical importance because recent findings cast doubt on whether TB forms an implicit reflection of SoA, or if so, to what extent it does. There are at least two shortcomings in the work on TB as a measure of SoA. One is that *TB tasks typically do not involve a clear goal or an extended action*, such as reaching for and grabbing a coffee cup (but see Kumar & Srinivasan, 2014, for an alternative task). Instead, they center around the somewhat arbitrary production of tones through keypresses. A second and more serious concern is that it has recently been shown that *TB effects may be accounted for through a regression-to-the-mean pattern* commonly observed as an effect of memory, rather than SoA (e.g., Saad, Musolino, & Hemmer, 2022). The TB task is inherently a memory task (recall of time intervals after producing them), and the behavioral patterns are indistinguishable from performance (specifically regression to the mean) in episodic memory. Thus, it is possible that TB effects do not reflect sensorimotor mechanisms or SoA. This possibility is also consistent with Vierordt's law (Vierordt, 1868).

Measuring SoA Explicitly: Rating Scales

While the most common implicit measure for SoA is TB, *explicit* measures of SoA involve conscious reflection and self-report. In most studies that use an explicit measure of SoA, participants are asked to rate the extent to which they felt in control over a preceding action on some continuous scale. While explicit rating scales come with inherent challenges (such as response biases, etc.), they have been shown to be sensitive to objective changes in control across different conditions in a way that one would expect. For example, van der Wel (2015) showed in a joint action task that explicit SoA ratings varied systematically with the actor's role in the action. If explicit ratings and implicit measures both measure SoA, then they should provide a converging picture of how SoA works. Unfortunately, however, it should be noted that explicit ratings and temporal binding effect have repeatedly been shown to be dissociable (e.g., Barlas & Obhi, 2014; Dewey & Knoblich, 2014; Pfister et al., 2021; Saito et al., 2015, but also see Makwana & Srinivasan, 2017). This suggests that TB and explicit ratings are tracking different cognitive processes, raising the question of which measure reflects SoA more clearly.

A Novel Task

Here, we aimed to address concerns raised above about measures of SoA by implementing TB and explicit measures in a novel task. First, we replaced the single button press common in TB studies with a goal-directed extended action task. The logic here is that more complex actions should be more sensitive to differences in SoA. Second, we modified the TB task to include two different versions; the first version (TB1 in Figure 1) is most like what happens in regular TB tasks, except that rather than a single button press we asked participants to choose between two possible targets and then move a dot displayed on a screen to that target through a sequence of button presses (with movement speed varying across conditions). Arrival at the target then caused a first tone (like the single button press in standard TB tasks), which is followed by a second tone at varying time intervals. Participants then completed an interval estimation task (i.e., they estimated the duration between the tones on a slider scale). We included both an active and passive condition. In the active condition, participants moved the dot. In the passive condition, participants watched the dot move to the target. We included these two conditions because they are

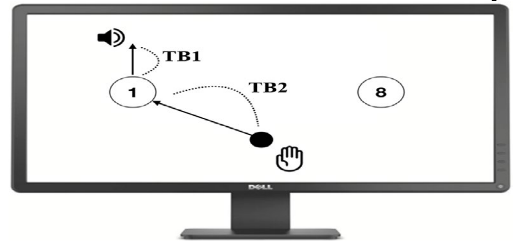

Figure 1. Overview of our experimental task. Participants first entered the number shown in the target of their choice. TB1 and TB2 are our implicit measures, see text.

common in standard TB paradigms. Participants also provided an explicit SoA rating after each block of 10 trials.

The second version of the TB task (TB2 in Figure 1) used the same task (including experimenter-controlled movement speed variations), except that starting the dot movement caused the first tone, and arrival at the target caused the second tone. Participants then again estimated the interval between the two tones. In other words, the time between tones was the same as the time it took to move the dot to the target. Participants also provided explicit ratings after each block of 10 trials.

This new version (TB2) is arguably much more interesting than standard TB tasks, as the interval duration is now under the participants' control. The task itself is an extended goal-directed action more akin to reaching for a glass of water, instead of a single button press resulting in an arbitrary tone. While the interval estimation task is still a recall task that happens after action completion, the action effect of relevance here is directly linked to the specifics of the actions (e.g., trajectory length, movement speed, and duration) and should be more revealing in terms of SoA.

Experiment 1

We provided a broad overview of our task and conditions above but provide further experimental detail here. Again, our novel experimental paradigm targeted the link between TB1, TB2, explicit ratings, and SoA. We tested 21 participants. Completion of the experiment took 75 minutes. Participants first completed a consent form and provided basic information such as handedness and neurological status. They were then seated behind a standard desktop computer (Dell Optiplex 7010 with a 22-inch monitor). We ran the practice and experiment through custom-written Matlab scripts using PsychToolbox (Brainard, 1997).

Practice Trials

Participants first completed 12 practice trials (3 trials for each of our four conditions; TB1/TB2 crossed with active/passive control of the dot). Each block of three trials within the practice portion first showed a screen that provided instructions for the specific version of the upcoming task (i.e., TB1/TB2 Active/Passive).

Each trial started with a presentation of two targets that varied in their location (targets labeled 1 and 8 in Figure 1 are examples). There were four possible target locations, one in each quadrant of the screen (i.e., top-left/right and bottom-left/right). We randomized which target combination was presented across trials. We also implemented small variations in coordinates along the x and y axes to ensure that participants had to adjust their steering slightly across different trials with the same target quadrant. The targets showed a random number between 1 and 9 in them. Once participants saw the two targets for a trial, they first indicated which target they would move to by entering the number corresponding to the target of their choice on a keyboard. After a 1000 ms delay, the trial then started and participants either moved (Active condition) or watched (Passive condition) the dot move to the target (details below). Once the dot reached the target and the two tones had played, participants then saw the next screen with a slider scale. Participants estimated the duration of the time interval between the tones by moving the index on the slider scale. Their estimate of milliseconds duration was displayed above the slider scale. The scale ranged from 0 to 1000 ms. Once they reached the value corresponding to their estimate, they submitted the estimate by pressing the spacebar. We ensured that the initial position of the index was randomized and needed to be changed before participants could submit their interval estimate. Participants received feedback on the accuracy of their duration estimate and the actual duration was shown on the screen after estimate submission (feedback was not provided during the experimental trials). After the 12 practice trials, participants were shown an example screen of the explicit rating scale they would complete after each of 10 experimental trials (details below). They then continued with the experimental trials.

Experimental Trials

Participants completed four blocks of 90 experimental trials per condition (360 trials in total). We counterbalanced the order of blocks across participants.

TB1 Active/Passive

During TB1 blocks, participants either moved (active) or watched movements of (passive) a dot from the start location in the center of the screen to their chosen target. In the active condition, participants moved the dot by using the direction keys on the keyboard. Movements varied with three levels of speed, such that there was a slow, medium, and fast condition within active and passive blocks (30 trials per speed). For active trials, speeds corresponded to 12, 17, and 22 pixels per button press. For passive trials, speeds corresponded to 22, 28, and 34 pixels per unit time. These speeds were based on pilot testing to ensure a feeling of control for each speed in the active condition and based on approximating movement intervals between ~300 and 700 ms in the passive condition.

Once the dot reached the target, this caused the first tone to play immediately (i.e., no temporal delay relative to the end of the action). We added this tone to the sensorimotor information participants got from ending their action (in active trials) and the visual effect of reaching the target on the screen to make the active and passive condition as similar as possible, except for the difference in actively acting. The first tone was followed by one of three intervals (300, 500, or 700ms, 30 trials each, randomized within a condition block) and then another tone. The participant then provided their duration estimate for the interval between the two tones using the slider scale described above.

TB2 Active/Passive

Trials in TB2 blocks were the same as TB1 blocks, except that the start of the dot movement caused the first tone (again, for the same reason as provided above) and the arrival of the dot at the target caused the second tone. This implies, in the active condition, that participants controlled the interval durations themselves.

As a preview of the results, it should be noted that while duration estimates in the passive condition were similar to TB1 (as the interval was under experimental control), it was possible for participants to generate interval durations (i.e., for movements to take longer) beyond 1000 ms in the active condition. This implies that the slider scale range was technically not sufficiently wide for these trials, and that it was not a priori possible to design the scale to cover the correct width. We planned to address this issue in our analyses, by considering the provided responses for these trials relative to their absolute duration values, as well as relative to the movement speed condition (i.e., slow, medium, or fast dot movements).

Explicit Ratings of SoA
In each block, participants also provided an explicit SoA rating after each 10 trials. They did so by entering on the keyboard a numerical response between 1 and 9 to the prompt: “To what extent did you feel in control during the past 10 trials?” A scale from 1 to 9 was displayed below this prompt, with 1 labelled as (“Not at all”) and 9 as (“Completely”).

Results

Explicit Ratings
We first analyzed the explicit ratings with a 2 (TB1/TB2) x 2 (Active/Passive) repeated-measures ANOVA. Figure 2 shows the results. The results indicated a main effect for Active (M = 6.30, SE = 0.29) vs Passive (M = 5.51, SE = 0.33), $F(1,19) = 7.28$, $p < .05$. The results did not show a main effect of TB1/TB2, $p > .10$. While the difference between the active and passive condition was numerically greater for TB2 versus TB1 trials, the interaction between TB1/TB2 and Active/Passive did not reach significance, $p > .05$. A Bayesian analysis confirmed that the model with Active/Passive as a factor was 4.64 times more likely than the null and 2.76 times more likely than a model with TB version * Active/Passive. In sum, participants reported a stronger SoA when they controlled the dot movements than when they watched the dot move. This finding suggests that our Active/Passive manipulation meaningfully tapped into changes in SoA.

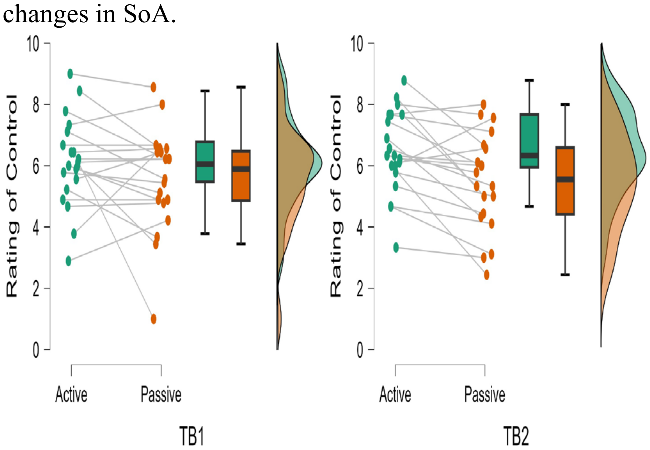

Figure 2. Explicit SoA ratings as a function of condition.

Implicit TB Measures
One of our main objectives was to better understand TB measures in the context of goal-directed actions that were more extended in time than a single button press. We incorporated TB1 as a version that is closer to traditional TB studies, as arrival at the target (like a single button press) started the interval and a second tone (after 300/500/700 ms) ended the interval to be estimated. In this version, interval duration was under full experimental control for active (TB1 Active) and passive (TB1 Passive) trials (as the tones did not depend on the produced movement). We also included our novel TB2 task, in which tones corresponded to the start and end of the goal-directed movements. In the passive condition (TB2 Passive), interval duration was also under experimental control (as we simulated movement trajectories that then controlled the dot movement). We controlled movement duration by feeding these trajectories with different gains (resulting in fast/medium/slow dot movements) to create different interval durations. In the active condition (TB2 Active), we also varied the gain to create fast/medium/slow dot movement conditions, but it is important to note that interval duration was not under experimental control. Instead, it depended on the participants’ movements in that case. For this reason, we report our results for each TB version separately below.

TB1: Binding for Tones after Movement Completion
Our first analysis focused on Bias for TB1 as a function of Active/Passive and Interval Duration (300, 500, or 700 ms). We started with this analysis because TB1 was most like standard TB studies that use button presses.

To calculate TB, we first calculated a measure of bias for each trial. We did so by taking the interval duration estimate a participant provided in a trial and subtracting the actual duration between the tones from this value. Positive resulting values thus indicated overestimation (or repulsion in TB terms) of the duration between tones, and negative values indicated underestimation (or binding). For each of the following analyses, we removed outlier values if Bias fell outside of the mean ± 2 STDs for a given participant within a given condition (so relative to the 90 trials per condition). We applied a Greenhouse-Geisser correction to the degrees of freedom for violations of the sphericity assumption.

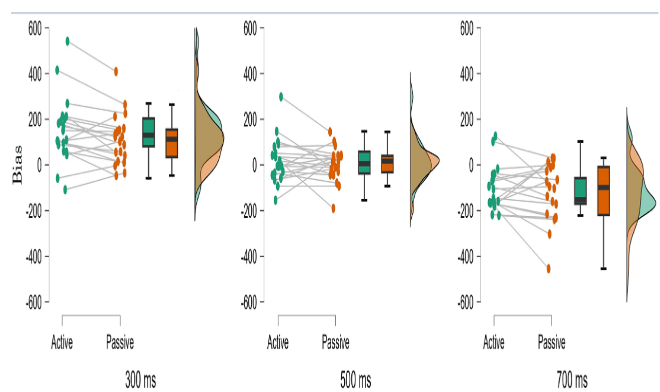

Figure 3. Bias as a function of Active/Passive and Interval Duration for TB1 trials.

The 2 (Active/Passive) x 3 (Interval Duration) repeated-measures ANOVA indicated a main effect for Interval Duration ($F(1.04, 19.74) = 45.22$, $p < .001$), but not for Active/Passive, $p > .10$. The interaction also did not reach significance, $p > .10$. As can be seen in Figure 3, mean Bias was positive for the 300 ms interval (M = 131.75, SE = 27.66), close to zero for the 500 ms interval (M = 10.79, SE = 16.04), and negative for the 700 ms interval (M = -114.83, SE = 21.42). A Bayesian analysis indicated that a model with just interval duration was 1.56 times more likely than a model with interval duration and Active/Passive, and at least five times more likely than any other model.

This result is consistent with previous findings in the TB literature – namely a regression effect and both compression (at 700ms) and repulsion (at 300ms). However, this result is not consistent with predictions from *theories* of TB which would predict compression for all time intervals, and a difference in degree of compression between active and passive.

TB2: Binding for Tones during Movement Completion

For TB2, we calculated Bias in two different ways; First, we considered Bias for TB2 Active and Passive relative to the minimum possible movement duration based on the experimenter-controlled movement speed (i.e., gain). By this, we mean that we determined how long it would take to move the dot from the start location to the target if there was no deviation from the straightest possible path and the participants did not pause anywhere along the movement trajectory. For TB2 Passive, this was identical to the actual interval duration they judged (as the dot movements were simulated at different gains and the interval duration was therefore under experimental control).

Figure 4 shows the results relative to the minimum possible duration based on dot movement speed. We then analyzed Bias for TB2 by conducting a 2 (Active/Passive) x 3 (Movement Speed: Slow/Medium/Fast) repeated-measures ANOVA (which is identical to how we analyzed it for TB1, except that we used Movement Speed instead of Interval Duration). The results revealed a main effect of Movement Speed ($F(1.11, 21.12) = 33.12$, $p < .001$), such that the fast (M = 65.51, SE = 23.37) and medium (M = 52.92, SE = 19.26) movement speed resulted in positive bias, and the low movement speed resulted in negative bias (M = -42.64, SE = 18.46). The results also indicated a main effect ($F(1, 19) = 17.40$, $p < .001$) of Active (M = 77.32, SE = 26.01) versus Passive (M = -26.79, SE = 18.25). These effects were qualified by a significant interaction between Active/Passive and Movement Speed, $F(1.383, 26.28) = 12.06$, $p < .001$, such that there was more positive bias for active than passive trials at fast and medium movement speeds, whereas bias did not differ significantly at slow speeds. A Bayesian analysis confirmed that a model with interval duration, Active/Passive, and their interaction term was over seven-hundred times more likely than any other model.

Second, we considered TB2 Active trials on their own. The analysis for TB2 we just presented used minimum possible movement durations based on movements straight to the

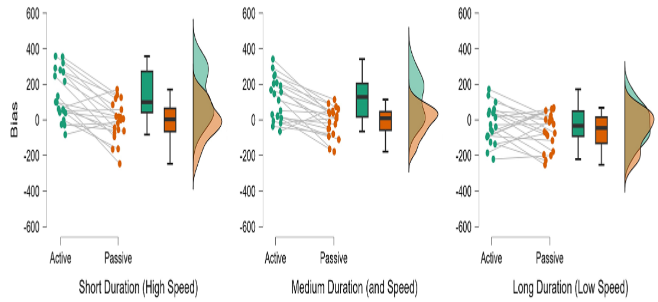

Figure 4. Bias as a function of Active/Passive and Movement Speed for TB2 trials.

target without any temporal delays. However, participants often deviated from this path and did not move continuously. As a result, the movement durations in TB2 Active could (and often did) exceed the maximum range on the slider scale.

To get a better sense of bias as a function of actual interval durations between the tones during TB2 Active, we calculated two measures. First, we calculated the proportion of cases for which participants generated movement (and thus tone interval) duration that exceeded the maximum of 1000 ms on the slider scale. This was the case for 44.81%, 45.56%, and 57.78% of trials for fast, medium, or slow movement speeds, respectively. Interestingly, however, we also calculated the proportion of cases for which the interval estimates participants provided was at the maximum of the range of the slider scale (i.e., 1000 ms). If participants considered the estimation task in an absolute way, then 1000 ms would be the most reasonable response when the actual interval duration exceeded this value. We observed, however, that participants responded with 1000 ms on 26.67%, 7.00%, and 3.67% of trials for fast, medium, and slow speeds, respectively. As these values are substantially lower than the proportion of actual durations exceeding 1000 ms, this suggests that participants judged the durations for TB2 Active in a relative way. They likely did so relative to the width of the slider scale rather than in actual movement durations.

Even though movement durations often exceeded the maximum value on the slider scale, we could calculate bias based on the actual movement durations participants accomplished. To do so, it is useful to consider median bias (median values better account for outliers and for the non-normal distributions of movement times here) as a function of dot movement speed (as movement durations varied with dot movement speed). The median bias for each movement speed was strongly negative (*Median* = -783.43 for fast speeds, *Median* = -683.86 for medium speeds, and *Median* = -749.64 for slow speeds). These values further indicate that participants rescaled their responses to the width of the slider scale.

In sum, for TB2 we found the same pattern of regression to the mean (with a negative slope) as we found for TB1. Although we did find a difference between active and passive this should be interpreted with some caution, as the active trial movements tended to exceed the response scale, and thus result in larger bias.

Discussion

Here, we introduced a novel paradigm aimed to address several concerns about standard SoA tasks; 1) we used more extended actions than a button press, 2) we introduced two versions of a TB task, with participants controlling the time interval between tones in one (TB2) but not the other (TB1), and 3) we obtained explicit ratings for the TB tasks and active/passive conditions.

First, several methodological points should be made about our novel task for studying SoA. Our experimental setup differed from most TB studies as we included a goal-directed action task that was extended in time. By this, we mean that most TB studies use a single button press that participants press when they feel the urge to do so in the active condition. This results in a tone that plays after some interval duration, and they then estimate the interval. In passive conditions, they typically observe the same task and hear a first tone at the time of the button press and another tone some interval later (e.g., Saad et al., 2022).

Our paradigm differs in several important ways. One is that participants chose one of two targets rather than pressing a single key to generate a tone. They chose the target in our passive conditions as well, such that they indicated the target and the dot then moved to that target 3 seconds later. Thus, participants had some level of control in our passive conditions, albeit much reduced and with a substantial delay compared to our active conditions. Second, our participants did not just hit a key once in the active condition, but completed a sequence of button presses in a goal-directed manner. As such, our task involved more extensive and continuous control than typical TB studies. In addition, participants controlled the first tone but not the second tone for the TB1 Active condition, and they controlled both tones and the interval between them for TB2 Active. In terms of task structure, then, one would expect TB2 Active to be more reflective of SoA than TB1 would be.

This is exactly what we found. Our explicit SoA ratings showed sensitivity to the objective amount of control over the actions, as active conditions resulted in higher SoA ratings than passive conditions. This was particularly the case for TB2 Active (although the interaction was not significant). These results suggest that, indeed, giving participants control over the timing of both tones (and the resulting interval duration between them) results in a stronger SoA.

When we considered bias (i.e., the difference between estimated and actual interval durations), our results showed two important patterns across TB1 and TB2. First, bias did not systematically change as a function of active or passive control in a trial. This finding is problematic for SoA accounts of TB, as they would predict more compression of interval durations for active versus passive trials. For example, Weller et al. (2020) hypothesized that TB would be weaker (less compression) for trials where participants did not perform an action (i.e., passive trials) relative to trials where they did perform an action (active trials).

Second, bias changed as a function of interval duration (TB1) and movement speed (TB2). While this pattern has repeatedly been observed in TB data (see Saad et al., 2022), it is not predicted by SoA account of TB. Importantly, overestimation (or repulsion between the action and outcome) should only occur for non-intentional actions - for example a finger twitch caused by transcranial magnetic stimulation (Haggard 2002). However, we found over-estimation (repulsion) for 300ms intervals for both active and passive trial for TB1 and overestimation for active trials for TB2. It is worth noting that the Weller et al. (2020) data also illustrates this pattern of repulsion at shorter intervals.

How then could one account for these results instead? One possibility is that the TB data from interval estimation paradigms show regression-to-the-mean. As we indicated in the introduction, such an account (e.g., Saad, Musolino, & Hemmer, 2023) can successfully simulate TB data, such as those by Weller et al. (2020). The core notion for this account is that the mean of the range of intervals used induces regression-to-the-mean for the short (300 ms) and long (700 ms) intervals. This then results exactly in the positive bias and negative bias observed at those intervals (regardless of whether it concerns active or passive trials). It is in our case remarkable to see that even our results for TB2 Active show some indications of a regression-to-the-mean pattern, despite the fact that the actual interval durations in many cases exceeded the range of the scale. Thus, participants rescaled their responses relative to interval durations across trials and relative to the width of the slider scale. This pattern would be expected based on a memory account or a scaling account, but not based on a SoA account of TB.

While we believe that our novel goal-directed task is a substantial improvement over standard TB button press tasks, our TB findings failed to show TB2 to be more reflective of SoA than TB1. In fact, our results failed to show any systematic relationship between the objective amount of control in our task and TB measures. This observation raises further doubt in terms of the usefulness of TB as an implicit measure of SoA.

References

Barlas, Z., Hockley, W. E., & Obhi, S. S. (2017). The effects of freedom of choice in action selection on perceived mental effort and the sense of agency. *Acta Psychologica*, *180*, 122–129.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242.

Buehner, M. J., & Humphreys, G. R. (2009). Causal Binding of Actions to Their Effects. *Psychological Science*, *20*(10), 1221–1228.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Dewey JA, Knoblich G (2014) Do Implicit and Explicit Measures of the Sense of Agency Measure the Same Thing? PLOS ONE 9(10): e110118.

Gallagher, S. (2012). Multiple aspects in the sense of agency. *New Ideas in Psychology*, *30*(1), 15–31.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, *9*(6), 290–295.
Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207.
Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385.
Haggard, P., & Tsakiris, M. (2009). The Experience of Agency: Feelings, Judgments, and Responsibility. *Current Directions in Psychological Science*, *18*(4), 242–246.
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*(5).
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), 133–151.
Kumar, D., & Srinivasan, N. (2014). Naturalizing sense of agency with a hierarchical event-control approach. *PLoS One*, *9*(3), e92431.
Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications*, *10*(1), 4250.
Makwama, M., & Srinivasan, N. (2017). Intended outcome expands in time. *Scientific Reports*, *7*(1), 6305.
Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59–68.
Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*(1), 546–561.
Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, *18*(4), 1056–1064.
Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107*(1), 179–217.
Pfister, R., Obhi, S. S., Rieger, M., & Wenke, D. (2014). Action and perception in social contexts: Intentional binding for social action effects. *Frontiers in Human Neuroscience*, *8*.
Saad, L., Musolino, J., & Hemmer, P. (2022). Bayesian rational memory model simulates temporal binding effect. *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*, 8.
Saito, N., Takahata, K., Murai, T., & Takahashi, H. (2015). Discrepancy between explicit judgement of agency and implicit feeling of agency: Implications for sense of agency and its disorders. *Consciousness and Cognition: An International Journal, 37,* 1–7.
Silver, C. A., Tatler, B. W., Chakravarthi, R., & Timmermans, B. (2021). Social Agency as a continuum. *Psychonomic Bulletin & Review*, *28*(2), 434–453.
Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional Binding Without Intentional Action. *Psychological Science*, *30*(6), 842–853.
Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219–239.
Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: An interplay between prediction and postdiction. *Frontiers in Psychology*, *4*.
Tsakiris, M., Prabhu, G., & Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, *15*(2), 423–432.
van der Wel, R. P. R. D. (2015). Me and we: Metacognition and performance evaluation of joint actions. *Cognition*, *140*, 49–59.
Vierordt, K. (1868). Der Zeitsinn nach Versuchen. Tübingen, Germany: Laupp.
Weller, L., Schwarz, K. A., Kunde, W., & Pfister, R. (2020). Something from nothing: Agency for deliberate nonactions. *Cognition*, *196*, 104136.