The Role of Endogenous Input in Self-Generated Action: A Multi-Measurement Study Comparing Markers of Volition

Authors: Guillaume P. Pech, Eva Nicolay, Paulius Rimkevičius, Uri Maoz, Axel Cleeremans

Introduction

What makes us feel that we are who we are, if not the sense that our actions are willed? Although it might seem easy to recognize a voluntary action from a first-person perspective, providing an objective scientific characterization of such actions from a third-person perspective remains challenging (Haggard, 2019; Jeannerod, 2009). One fundamental characteristic of voluntary action is that it is self-generated. Therefore, understanding what makes an action self-generated could help characterize what a voluntary action is. Researchers have thus attempted to define the criteria and mechanisms that characterize self-generated action (Nachev & Husain, 2010; Obhi, 2012; Passingham et al., 2010b, 2010a; Schüür & Haggard, 2011, 2012). Schuur and Haggard (2011) proposed the following definition of self-generated action: "the motor consequence of integrating a range of different types of input". Input refers to the information that the brain integrates when we are deciding to initiate an action. So according to this definition, unlike automatic or reflexive actions (which are triggered by one type of input), self-generated actions are based on a combination of different types of input. Although insightful, the definition proposed by Schuur and Haggard (2013) does not fully consider a critical aspect: the nature of the input that contributes to self-generated actions, in particular, whether it is endogenous or exogenous determined. By "determined by exogenous input", we mean information about which action to perform, while, by contrast, "determined by endogenous input" refers to information that needs processing by the cognitive system in order to help determine which action to perform (i.e., it requires memory, reasoning, valuation). The objective of this study is thus to investigate how different markers of volition are influenced by the quantity of different types of input, as well as by their endogenous or exogenous nature determination.

The preliminary distinction between self-generated and other types of actions emerges from early distinctions between classical and operant conditioning. On one side we find actions driven by external stimuli (e.g., Pavlovian reflexes: Fanselow & Wassum, 2016; Pavlov (1927), 2010), on the other side those whose origin lies in the organism (i.e., reinforcement learning: Silvetti & Verguts, 2012; Skinner, 1984), which also constitutes the distinction between exogenous and endogenous actions (Schüür & Haggard, 2011). Another distinction has been made between operant behavior and underdetermined actions (i.e., actions not fully determined by external input), in which the determining factor arises from the agent's will (i.e., self-generated action; Passingham et al., 2010b; Schüür & Haggard, 2011). Neuroimaging and lesion studies have notably indicated that self-generated actions recruit distinct brain areas compared to operant actions, including the Supplementary Motor Area (SMA) and the pre-SMA (Passingham et al., 2010b; Seghezzi et al., 2019). For example, Thaler and colleagues (1995) showed that lesions to the SMA and pre-SMA in monkeys perturbs non-cue-generated action (where the monkeys had to make an arbitrary movement to deliver peanuts). However, the same monkeys preserved their performance for cued-generated action triggered by a tone to realize the same movement (i.e., operant conditioning). But, defining self-generation via non-cue-generated action (i.e., through underdetermination) involves a risk of circularity and has been criticized as unscientific (see Nachev et al., 2010; Passingham et al., 2010b). Self-generated action has to be determined by something, if the only defining feature of self-generated action is underdetermination, we are not given very specific directions on what we should be investigating.

Schuur and Haggard proposed a new definition of self-generated action which distinguishes it from action driven by operant conditioning, while avoiding the underdetermination argument. The authors define self-generation as the complex integration of different types of input. Input is defined as being “information from different times, locations, modalities, and information about different objects, and attributes” (Schuur & Haggard, 2013: page 6). This definition allows for a continuous rather than discrete categorization of self-generated action. The more types of input there are, the more self-generated the action is. Adding to this proposed definition, in this paper, we argue that this model omits the dimension of endogenicity, as inherited from operant conditioning theory and proposed by Passingham and colleagues (2010b) in response to Nachev’s criticism (2010). According to this literature, all actions would be endogenous in an operant conditioning situation (both on Passingham’s "underdetermination account" and on Schuur and Haggard's multi-input account). What we would like to stress here is that, even under those conditions, an action might be judged less voluntary if the input is largely exogenous. So, according to our account, whether an action is self-generated depends on whether there was endogenous input (like the underdetermination account says), but also, whether an action is self-generated can depend on the number of inputs (as the multiple-input account says), and it can be a matter of degree. We acknowledge that delineating the boundaries of what makes an input exogenous or endogenous is not easy. For our purposes here, as noted above, we define endogenous input as requiring processing by the cognitive system. However, we leave the question open whether the cognitive system includes only the brain (Gallagher, 2017, page 28; Goldman & de Vignemont, 2009), or also other parts of the body (Gallagher, 2017, page 38; Shapiro, 2019), or even parts of the environment (Gallagher, 2017, page 40; Thompson & Varela, 2001). Therefore, for example, perhaps one could characterize information coming from the stomach as both exogenous and endogenous. Where exactly to locate the boundary of the cognitive system is beyond the scope of this paper.

To measure volition, one can rely on self-reports (e.g., Charles & Haggard, 2020), but explicit ratings such as those are vulnerable to biases such as social-desirability bias (Fisher, 1993) and self-attribution bias (i.e., it has been shown that people tend to over-estimate their contributions to positive outcomes and under-attribute them to negative ones; Choi & Pak, 2005; Tsakiris & Haggard, 2005). Alternatively, two implicit markers of volition have been used in the literature: the Readiness Potential (RP) and Temporal Binding (TB). The RP—a slow ramping activity preceding movement—has been linked to volition and seems to originate from the (pre-)SMA regions (Shibasaki & Hallett, 2006), which are also regions described above as being linked to self-generated action (Passingham et al., 2010b; Seghezzi et al., 2019). It should be noted that the right interpretation of the RP is debated, with alternative views proposed. For instance, the RP may reflect noisy accumulation (Schurger et al., 2012), or alignment with slow cortical potentials (Schmidt et al., 2016).

Most RP studies have compared self-generated actions with actions triggered by an external stimulus, reporting weaker RPs for the latter (Di Russo et al., 2017; Jahanshahi et al., 1995; Jenkins et al., 2000; Khalighinejad et al., 2018). However, self-generated actions and actions triggered by an external stimulus differ notably in terms of urgency regarding the preparation of the movement. There is no imperative to act quickly in the former, while there is an imperative to move quickly in the latter. Since the RP has been described as a marker of motor preparation, controlling for urgency between the conditions is crucial. One possibility is to contrast the situation where the participant has to choose which action to perform (i.e., free), with a situation where the action is determined by a cue, but without an imperative to move quickly (i.e., instructed). Moreover, drawing on the multiple-input account, we can increase the number of action alternatives in order to see whether that contributes to the action being felt as more self-generated. To the best of our knowledge, no study has contrasted the RP in self-generated actions where there are alternatives (i.e., free), with stimulus-driven actions where there is only one alternative (i.e., instructed). Furthermore, we also did not find RP studies where the number of action alternatives was manipulated.

Interestingly, Filevich and colleagues investigated another EEG signal preceding voluntary movement, the Contingent Negative Variation (CNV; Brunia et al., 2011; Filevich et al., 2013). CNV is classically produced between the presentation of a preparatory cue, and an imperative cue signaling the requirement to move (Di Russo et al., 2017; Schurger et al., 2021). In their study, participants were asked to produce either a rapid reaction to a cue, or to respond after a short self-chosen delay. In the instructed condition, participants had to act as quickly as possible, or after a chosen delay depending on the specific instructions. In the free condition, participants could choose to act as quickly as possible, or after the chosen delay. Although the authors did not directly compare the instructed and free conditions, they found that in the free condition, the CNV amplitude was lower in the delayed responses compared to rapid responses. In contrast, in the instructed condition, the CNV was higher in the delayed responses compared to rapid responses. In our study, we aim to investigate whether the RP is similarly affected by comparing instructed and more free self-generated actions.

The second implicit marker of volition, TB (Haggard et al., 2002), refers to the perceived compression of time between an action and its effect when the action is voluntary versus fully externally triggered (e.g., via TMS). Whether the intentional part of the voluntary action is necessary to produce TB, or if TB is due to an increase of perceived causality, is still debated (Haggard, 2017; Hoerl et al., 2020; Kong et al., 2024; Moore & Obhi, 2012; Wen & Imamizu, 2022). Nonetheless, the distinction between instructed and freer self-generated actions is more extensively documented in the literature on TB than on RP. Notably, it has been shown that those actions that are chosen more freely produce reliably stronger TB (i.e., shorter interval estimates) than instructed ones (Barlas et al., 2017; Barlas, 2019; Barlas & Obhi, 2013; Borhani et al., 2017; Caspar, Christensen, et al., 2016; Caspar et al., 2018, 2020; Pan et al., 2024; Pech & Caspar, 2023). Barlas and colleagues (2013, 2017), for instance, manipulated the number of action alternatives (i.e., one to five) and found significant TB differences between the instructed condition and the one with the greatest number of alternatives. Only specific contexts, such as military settings, appear to reduce or abolish the difference in TB between self-generated and instructed actions (Caspar et al., 2020).

In this pre-registered multi-measurement study, we investigated input integration, by varying actions based on a single input (i.e., instructed) to actions based on 4, 8 or 9 input, to test Schuur and Haggard's account of self-generation. Moreover, we contrasted decisions that are determined more by exogenous input or more by endogenous input (see below). We recruited 39 participants who, in a Random Dot Kinematic (RDK) task (Charles & Haggard, 2020), judged the motion direction under four conditions varying in input quantity and nature: Instructed (one direction option; only exogenous input), Semi-Instructed (four direction options; endogenous input, choosing from four options), Recommended (eight direction options; endogenous input, choosing from eight options; plus exogenous—a recommendation of one of the options), and Free (eight direction options; entirely endogenous input). Drawing on the literature cited above, we measured three main markers of volition: self-report, the RP, and TB. We formulated a first hypothesis (H1), drawing on the account of Schuur & Haggard (2011), which stated that integrating more input would produce actions that are felt as more self-generated. More precisely, we predicted the following direction in H1: *Recommended (9 input) > Free (8 input) > Semi-Instructed (4 input) > Instructed (1 input)*. To evaluate our suggestion that not only the quantity of input matters, but also the nature of the input, Hypothesis 2 contrasted decisions based predominantly on endogenous input (i.e., Free condition) with decisions integrating more exogenous input (i.e., Recommended condition).

According to Hypothesis 2, free action should be felt as more self-generated than recommended action. More precisely, the following direction was predicted in H2: *Free* > *Recommended* > *Semi-Instructed* > *Instructed* (Fig. 1).

Overall, our results showed that self-reports supported Hypothesis 2, whereas implicit and neural measures (TB and the RP) favored an alternative account: these markers discriminated mainly between exogenously and endogenously generated actions, without tracking input quantity or tracking endogenicity in a graded fashion.

Methods

Participants

We recruited 39 participants through the credit system of the Université Libre de Bruxelles. Participants were instructed to arrive with hair conditions compatible with EEG recording (dry hair and no products that could interfere with the EEG). All participants were first-year students in the Faculty of Psychological and Educational Sciences and received study credits as compensation for their participation.

We aimed for a minimum sample size of 16 participants to meet the criteria to ensure sufficient statistical power (Brysbaert, 2019; Brysbaert & Stevens, 2018). A properly powered experiment should have at least 1,600 observations per condition, which translates to 16 participants with 100 stimuli per individual in the four conditions in our experiment (Brysbaert, 2019; Brysbaert & Stevens, 2018). We oversampled to 20 to account for potential data loss due to noise in EEG imaging, and failure to perform the interval estimation task. We pre-registered this sample size of 20 with the condition of having enough certainty in our posterior distribution using Bayesian models. We observed that the models were still inconclusive for the null and alternative hypothesis with 20 participants for the implicit markers (i.e., the RP and TB). We therefore collected data until the end of the academic semester. This resulted in a final sample of 39 participants (age: M = 20.4, SD = 5.0; 33 female). All participants provided written informed consent before participating. This study was approved by the Ethical Review Board of the Faculty of Psychology (Reference: 1544/2023).

Apparatus and Procedure

The experiment consisted of a computerized task where participants judged the general direction of a moving dot cloud. The dot cloud could move in one of eight randomly assigned directions and was visible for 2 seconds (see Figure 1). The motion of the dots was partially random and partially consistent with one of the eight directions, making the discrimination task ambiguous. Task difficulty was calibrated for each participant to enhance comparability and ensure ambiguity, following Lau et al. (2015), to minimize the influence of a predictable outcome on perceived voluntariness. After the dot cloud disappeared, arrows representing the directional options appeared on-screen, and participants selected the direction they believed was correct using a numeric keypad. There was no time limit for selecting. After a variable delay (200ms, 500ms, or 800ms; (Pech et al., 2025; Pech & Caspar, 2023), the selected arrow turned green, and a tone was played. Subsequently, participants performed an additional task depending on the trial type: (1) in estimation trials, they judged the time interval between their button press and the tone on a scale from 0 to 1000ms; (2) in question trials, they rated their agreement with statements related to the sense of volition on a scale from 0 (not at all) to 100 (completely). The questionnaire comprised ten questions (see supplementary S1 for the list of questions), repeated three times per condition. The questions were adapted from the questionnaires of Tapal and colleagues (2017). Feedback was given at the end of each trial to maintain motivation, showing participants their score based on accuracy (e.g., 2 points for correct direction, 1 for adjacent directions, 0 otherwise)

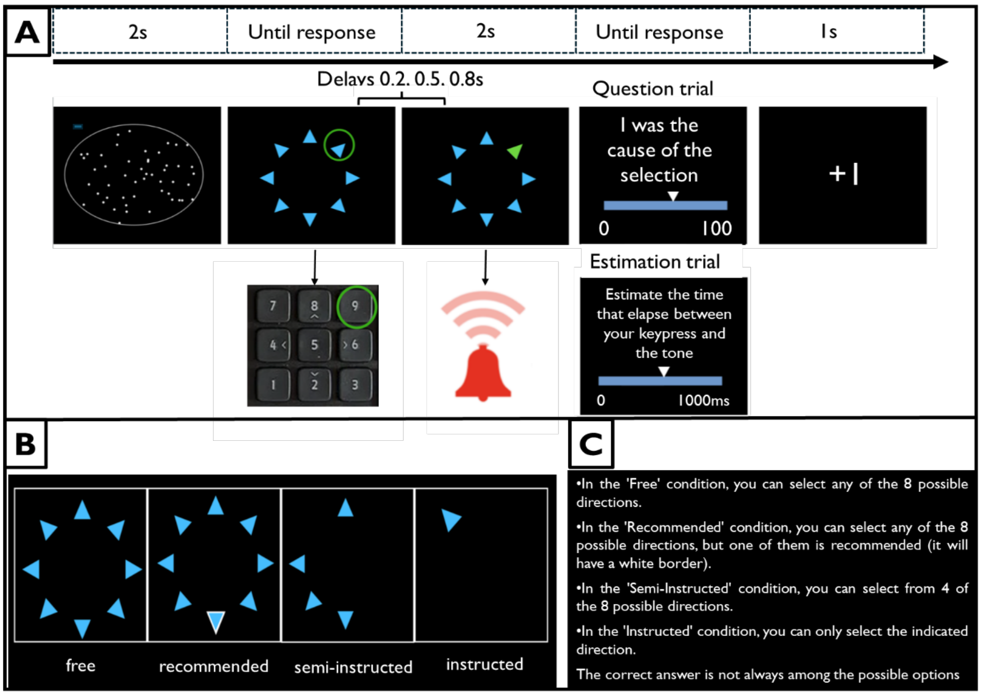

Figure 1 Experimental design: A) Schematic representation of the task presentation on the screen. B) Representation of the directional arrow displays available depending on the condition. C) Exact instructions provided to participants regarding the different conditions.

Calibration

Task difficulty was calibrated using a staircase procedure to tailor the level of ambiguity for each participant. Initially, dot motion coherence was set to 100%, meaning all dots moved in the same direction. Through three adaptive phases, coherence was adjusted based on participant performance to achieve an accuracy level slightly above chance. The individualized coherence level was then applied to the experimental task.

Training

Following calibration, participants completed a time estimation training phase. This involved trials identical to those in the experimental task, except that a tone was played 200ms, 500ms, or 800ms after the directional choice. Participants estimated the time interval between their button press and the tone on a scale from 0 to 1000ms. Training continued until participants reported verbally to the experimenter that they were confident in their ability to differentiate between the delays.

Experimental Task

The experiment compared four conditions varying in the number of directional options and the degree of constraint. The experiment was divided into three parts, each containing four blocks representing the conditions, presented in a fixed but counterbalanced order. Each block included 13 estimation trials and 10 question trials (see Figure 1A), resulting in 276 trials per participant (156 estimation and 120 question trials). At the beginning of each block, participants were informed of the upcoming condition and trial type. After completing a block, participants received feedback on their overall score and accuracy rate to maintain motivation.

The experimental task included four distinct conditions designed to vary the number of options and constraints (See Figure 1B and 1C). In the Free trials, participants were able to select freely from all eight possible directions. In the Instructed trials, participants were required to select the only available option, which was randomly determined by the computer and not necessarily the correct direction. The Semi-Instructed trials limited participants to choosing from four randomly selected directions, which did not always include the correct one. Finally, the Recommended trials allowed participants to choose from all eight directions, but one direction was highlighted as a recommendation. These recommendations were adapted, ensuring that participants did not adhere too frequently (e.g., incorrect recommendations were provided if adherence exceeded 70%) or too rarely (e.g., correct recommendations were made if adherence was below 30%).

EEG recordings

EEG data were acquired at a sampling rate of 2,048 Hz from 64 channels, placed according to the international 10-20 system, using the Biosemi equipment (see http://www.biosemi.com for hardware details). All data were recorded using Actiview software. Due to error of recording while using Activew, 2 participants did not have triggers record (the files are accessible on osf). We therefore had 37 participants remaining.

EEG processing

Data were processed using MNE-Python (Gramfort et al., 2013). We down-sampled the data to 512 Hz using the default Fast Fourier Transform method of the resample() function in MNE-python. We then applied a bandpass filter between 0.1 and 40 Hz, with the Finite Impulse Response method, a zero-phase delay and the Hamming window (in line with other literature that measured the RP; Maoz et al., 2019; Nann et al., 2019; Pech et al., 2025; Travers et al., 2021; Verbaarschot et al., 2019). We detected and interpolated bad channels automatically using the find_all_bads() function of the of the pyprep library(mean = 3.2, SD = 2.6; Bigdely-Shamlo et al., 2015). This function uses a combination of criteria: "extreme amplitudes (deviation criterion), lack of correlation with any other channel (correlation criterion), lack of predictability by other channels (predictability criterion), and unusual high-frequency noise (noisiness criterion)" in addition to the random sample consensus method (Fischler & Bolles, 1987). After interpolating, a copy of the data was created in order to apply an independent component analysis (ICA), to detect eye movements, with a high-pass filter of 1 Hz, with a number of components calculated in order to represent 99.99% of the data. The high-pass filter of 1Hz allows an improvement of the performance of the ICA (Winkler et al., 2015). ICAs were computed with the number of components defined in order to represent 99.99% of the data. To detect eye movements (i.e., blink and saccades), we visually selected the components (M = 2.1, SD = 0.3) to remove from the original data, and the copy of the data was not further used. During the visualization of components, we also looked if participants had more than 32 components (out of 64 electrodes) as a quality check. ICA is a source separation technique, if fewer components were founded, we took this as a sign of long segment of noise remaining. For these participants (4/37), we annotated manually bad segments based on visual inspection, and reperformed the ICAs. If the ICAs decomposed into fewer than 32 components, we did not keep those participants for further analysis, but this was not the case.

We then re-referenced the channels using the reference-electrode standardization technique with a point at infinity with a head model and the forward method (Gramfort et al., 2013; Yao, 2001; Yao et al., 2019). The data were epoched on electrode Cz in a window from -3s to 1s around the hand press. We used a baseline from -5ms to 5ms around the action onset to avoid making any assumptions about the onset of the RP (Khalighinejad et al., 2018; Pech et al., 2025). We also extracted and analyzed the slope of the RP, which is not influenced by the selection of baseline. Moreover, measuring the mean and the slope helps in getting more confidence in the results if they do converge. We considered the time from –1s to 0s relative to movement onset as the timing of the RP, guided by the literature and based on a visual inspection of the grand averages across all conditions and participants (Gavenas et al., 2025; Khalighinejad et al., 2018, 2019; Parés-Pujolràs et al., 2019; Pech et al., 2025; Travers et al., 2021; Travers & Haggard, 2021). Epochs containing artifacts were rejected based on the value of the mean, the peak-to-peak magnitude, and the slope during this period within each participant (see “Removed data” section below).

Statistical analysis

Our main analysis method relied on Bayesian linear mixed models, using the ‘brms’ R package (Bürkner, 2017). For each parameter we report the estimated medians (Med) and the 89% Highest Density Interval (HDI89%). Furthermore, for each comparison we report the estimated medians difference (Meddiff), the 89% Highest Density Interval (HDI89%; Kruschke, 2018; Makowski et al., 2019; McElreath, 2016), the Probability Direction (PD; Kruschke, 2018; Makowski et al., 2019), the Bayes Factor in favor of H1 (BF10 > 3 in favor of H1, BF10 < 1/3 in favor of H0; Dienes, 2014; Kruschke, 2018; Makowski et al., 2019), and the Robustness Range that leads to the same Bayes Factor conclusion (e.g. RRH0 if BF10 < 1/3 or RRIN if 1/3<BF10<3; Dienes, 2014, 2019). See Supplementary S2 Statistical Analysis for further information. Note that the PD is roughly equivalent to a p-value (Makowski et al., 2019; Shi & and Yin, 2021). Hence, a p-value of 0.05 would correspond to a PD of 97.5% for a two-sided test, and to a PD of 95% for a one-sided test.

For each parameter we report the estimated medians (Med) and the 89% Highest Density Interval (HDI89%). Furthermore, for each comparison we report the estimated medians difference (Meddiff), the 89% Highest Density Interval (HDI89%), the Probability Direction (PD), the Bayes Factor in favor of H1 (BF10), and the Robustness Range that leads to the same Bayes Factor conclusion (e.g. RRH0 if BF10 < 1/3 or RRIN if 1/3<BF10<3).

Removed data

Because we wanted to ensure that participants had paid attention to their selection, trials in which response times (RTs) were faster than 350ms were discarded (Cohen & Donner, 2013; Pech et al., 2024; Pech & Caspar, 2022; Semmelmann & Weigelt, 2017). Overall, this resulted in the rejection of 0.3% of trials (SD = 1.1%) for all our dependent variables (i.e., questions, interval estimations and RP).

We also removed trials that were outliers within each participant, due to behavioural or EEG data. We used the Inter-Quartile Range (IQR) method with a threshold of 2 IQR below the 25th percentile, and above the 75th percentile to demarcate outliers (Hoaglin et al., 1986; Jones, 2019; Leys et al., 2013; Osborne & Overbay, 2019). This method rejected 4.5% of trials (SD = 2.1%) for the RTs, 13.3% (SD = 7.5%) for the RP, and 1.3% (SD = 2.0%) for the TB.

In addition to removing outlier trials within participants, we also removed outlier participants. For the TB measure, we wanted to ensure that participants correctly discriminated the three durations of the delays (i.e., 200ms, 500ms, 800ms). We set up a contrast with –1, 0, 1 for the 200 ms, 500ms, and 800ms delays, respectively. We then performed a linear regression analysis using the 'lm()' function in R with interval estimate as an outcome, and the three delays as a predictor, for each participant separately. We only kept participants with a significant positive linear trend (i.e., $p < .05$) for the contrast across the delays (similarly as: Caspar et al., 2016, 2018; Pech et al., 2025; Pech & Caspar, 2022). This resulted in the removal of 6/39 (15%) participants. We also removed participants who were outliers using the 2 IQR method for each measurement (Pech et al., 2025; Pech & Caspar, 2025). Based on this method, we removed 0/39 participants for the reported sense of volition (SoV), 1/37 participants for the slope and the mean of the RP (2 were already selected to be removed due to a recording problem), 4/33 participants for TB (6 were already selected to be removed due to the absence of significant linear trend).

Sense of Volition (SoV) questions

We calculated Cronbach's alpha over the 10 questions to ensure inter-item reliability. To do so, we used the 'cronbach.alpha()' function of the ltm library in R (Rizopoulos, 2007), with 1,000 bootstraps and 89% confidence interval. A value above 0.7 is usually taken as evidence for reliability (Tavakol & Dennick, 2011). We found evidence for reliability between the 10 questions ($\alpha = 0.906$, $CI_{89\%} = [0.886\ 0.921]$). The questions were thus considered as similarly measuring the reported sense of volition.

Results

This section displays the results for the reported SoV, the TB, and the RP. For each of these measurements, different models were computed, which are accessible in Supplementary

Information S3, with Figure 2 for self-reports, Figure 3 for TB, and Figures 4-5-6 for the RP. We used a model with a Gaussian family distribution, an identity link for both mu and sigma for all the models (see Supplementary Information S3 for more information). We included the condition of self-generation (i.e., Instructed, Semi-Instructed, Recommended, and Free conditions) as main effect. Participants were included as random intercepts. The main effect was included as random slopes per participant. Moreover, we observed differences across our conditions in reaction times to select the direction, and in accuracy of the reported direction (see Supplementary S3). A previous study showed an effect of reaction times on both the TB and the RP (Pech et al., 2025); reaction times was therefore added (scaled and centered) as a fixed effect to the models. Finally, we observed differences in accuracy across conditions (see Supplementary S3). As previous studies have shown that SoV, TB, and RP can be affected by outcome valence (Kaiser et al., 2021; Pech et al., 2025; Tsakiris & Haggard, 2005; Yoshie & Haggard, 2013), we included the correctness of the chosen direction as a fixed effect in the models. This was the minimum structure of all the models, except if stated, it was the model use. All models are described in full details in the supplementary S3 section.

Self-reported Sense of Volition

We analysed the reported SoV (from 0 = "totally disagree" to 100 = "totally agree").

For the Instructed condition (Med = 35, HDI89% = [31 39]), there was evidence for lower reported SoV compared to the Semi-Instructed condition (Med = 56, HDI89% = [53 59]; Meddiff = –21, HDI89% [-23 -18], PD = 100%, BF10 = 5.50e+15, RRH1 = [<5 1526]), the Recommended condition (Med = 63, HDI89% = [60 66]; Meddiff = –28, HDI89% [–32 –23], PD = 100%, BF10 = 1.35e+16, RRH1 = [<5 2018]), and the Free condition (Med = 70, HDI89% = [66 73]; Meddiff = -35, HDI89% [–39 –30], PD = 100%, BF10 = 1.33e+16, RRH1 = [<5 2489]). For the Semi-instructed condition (Med = 56, HDI89% = [53 59]), there was evidence for lower reported SoV compared to the Recommended condition (Med = 63, HDI89% = [60 66]; Meddiff = –7, HDI89% [–11 –4], PD = 99.9%, BF10 = 146.4, RRH1 = [<5 124]), and the Free condition (Med = 70, HDI89% = [66 73]; Meddiff = –14, HDI89% [–18 –10], PD = 100%, BF10 = 7.01e+15, RRH1 = [<5 815]). Finally, for the Recommended condition (Med = 63, HDI89% = [60 66]), there was evidence for lower reported SoV compared to the Free condition (Med = 70, HDI89% = [66 73]; Meddiff = –7, HDI89% [–9 –5], PD = 100%, BF10 = 375607.3, RRH1 = [<5 329]).

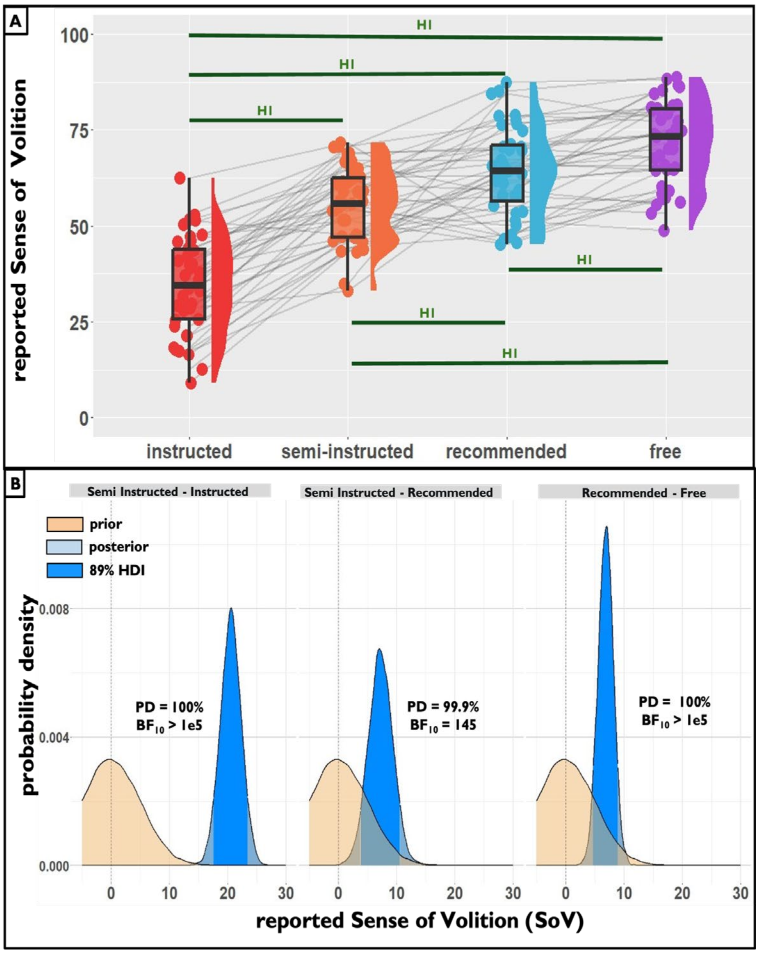

Figure 2. Reported SoV. (A) In each boxplot, the thick horizontal black line represents the median. The top and bottom parts of the box represent the 75% and 25% percentiles of the data. The vertical lines represent the 1.5 interquartile range starting from the 25% percentile at the bottom and the 75% percentile on top. Each line connecting the boxplots and dots represents the reported SoV of a single participant. H1 represents evidence for a difference (see Statistical analysis for more details). (B) Posterior analysis of reported SoV. Each column represents the posterior distribution of the difference between two conditions (e.g. "Semi Instructed minus Instructed," is the resulting posterior distribution of the Semi-Instructed condition minus the Instructed condition). The distribution of the prior used to run the model and to calculate the BF10 is in orange. It has a normal distribution centered on 0 with a SD of 5. The blue distribution is the posterior distribution, with the 89%HDI (in darker blue). The dashed vertical line is at 0 (no difference).

Resume results SoV

The reported SoV increased progressively across the conditions. The Instructed condition showed the lowest reported SoV. The Semi-instructed condition displayed a moderate increase in reported SoV. The Recommended condition further showed reported SoV. The Free condition showed the highest reported SoV.

Temporal Binding

Different delays were used in the procedure between the action and the outcome. This different delays seems to produces variation at least on the effect on the TB (Caspar, Christensen, et al., 2016; Kong et al., 2024; Pech et al., 2025; Wiesing & Zimmermann, 2024). Therefore, we added the actual delays (i.e., 200ms, 500ms, and 800ms) as a fixed effect, as well as in interaction with our self-generation conditions in the model. The fixed effect and their interaction were also added as random slopes per participants. We analysed the Temporal Binding (TB) effect using interval estimation (in ms), where shorter interval estimations reflect stronger TB.

For the 200ms delay, there was evidence for stronger TB in the Instructed condition (Med = 224ms, HDI89% = [177 274]), compared to both the Semi-Instructed condition, though inconclusively (Med = 243ms, HDI89% = [192 290]; Meddiff = –18.5ms, HDI89% [–40.0 0.7], PD = 92.6%, BF10 = 2.29, RRIN = [11 106]), and the Recommended condition (Med = 253ms, HDI89% = [202 301]; Meddiff = –28.2ms, HDI89% [–50.7 –5.7], PD = 97.5%, BF10 = 5.96, RRH1 = [<15 30]). There was inconclusive evidence for similar TB between the Instructed condition and the Free condition (Med = 228ms, HDI89% = [179 277]; Meddiff = -3.4ms, HDI89% [–27.8 21.0], PD = 58.9%, BF10 = 1.02, RRIN = [5 49]). There was inconclusive evidence for similar TB in the Semi-Instructed condition compared to the Recommended condition (Meddiff = –9.7ms, HDI89% [–27.4 9.1], PD = 80.5%, BF10 = 1.18, RRIN = [5 53]), and inconclusive evidence for weaker TB in the Semi-Instructed condition compared to the Free condition (Meddiff = –14.9ms, HDI89% [–36.4 6.9], PD = 86.4%, BF10 = 1.58, RRIN = [7 75]). Finally, there was evidence for stronger TB in the Free condition compared to the Recommended condition (Meddiff = –24.5ms, HDI89% [–45.0 -5.0], PD = 97.5%, BF10 = 5.6, RRH1 = [<15 30]).

For the 500ms delay, there was evidence for weaker TB in the Instructed condition (Med = 423ms, HDI89% = [370 478]), compared to both the Recommended condition (Med = 402ms, HDI89% = [350 453]; Meddiff = -20.4ms, HDI89% [–40.8 -0.3], PD = 94.3%, BF10 = 2.99, RRIN = [14 140]) and the Free condition (Med = 399ms, HDI89% = [348 449]; Meddiff = -23.5ms, HDI89% [–46.2 0.7], PD = 94.4%, BF10 = 3.17, RRH1 = [<15 17]). There was inconclusive evidence for similar TB between the Instructed condition and the Semi-Instructed condition (Med = 412ms, HDI89% = [360 465]; Meddiff = –10.0ms, HDI89% [–30.0 8.3], PD = 80.2%, BF10 = 1.13, RRIN = [6 53]). There was inconclusive evidence for similar TB in the Semi-Instructed condition, compared to both the Recommended condition (Meddiff = –10.4ms, HDI89% [–25.7 5.6], PD = 85.5%, BF10 = 1.16, RRIN = [5 53]), and the Free condition (Meddiff = –13.5ms, HDI89% [–34.2 6.6], PD = 85.1%, BF10 = 1.4, RRIN = [7 65]). Finally, there was inconclusive evidence for similar TB between the Recommended and Free conditions (Meddiff = –3.2ms, HDI89% [–22.0 15.5], PD = 60.2%, BF10 = 0.82, RRIN = [4 37]).

For the 800ms delay, there was evidence for weaker TB in the Instructed condition (Med = 584ms, HDI89% = [524 648]) compared to the Semi-Instructed condition (Med = 554ms, HDI89% = [493 613]; Meddiff = -29.9ms, HDI89% [–54.4 –5.7], PD = 97.5%, BF10 = 6.99, RRH1 = [<15 37]), the Recommended condition (Med = 543ms, HDI89% = [482 605]; Meddiff = –41.1ms, HDI89% [-68.3 -13.9], PD = 98.9%, BF10 = 16.53, RRH1 = [<15 86]), and the Free condition (Med = 552ms, HDI89% = [486 613]; Meddiff = –31.3ms, HDI89% [–62.7 –

0.9], PD = 94.7%, BF10 = 5.13, RRH1 = [<15 26]). We observed inconclusive evidence for similar TB in the Semi-Instructed condition, and both the Recommended condition (Meddiff = –10.8ms, HDI89% [–27.6 7.0], PD = 83.1%, BF10 = 1.13, RRIN = [5 >53]), and the Free condition (Meddiff = –1.3ms, HDI89% [-23.6 21.7], PD = 53.5%, BF10 = 0.95, RRIN = [5 46]). Finally, we observed inconclusive evidence for similar TB in the Recommended condition compared to the Free condition (Meddiff = 9.4ms, HDI89% [–11.1 29.6], PD = 76.7%, BF10 = 1.08, RRIN = [5 53]).

Resume results TB

For the 200ms delay, we found evidence of stronger TB in the Free and Instructed condition compared to the Recommended, and inconclusive evidence compared to the Semi-Instructed condition.

For 500 and 800ms delays, we found evidence of weaker TB in the Instructed and Semi-Instructed conditions compared to the others.

The 500-800ms effects are in line with the results of previous studies (Barlas, year) showing stronger TB with more alternatives available. For the 200ms it might be an effect of residual stronger cognitive conflict. In the Instructed condition, participants know that they must select the displayed arrow. In the Free condition they can just follow what they feel is the right answer. In contrast, in both the Recommended and Semi-Instructed conditions, participants had to adjust what they felt was the good answer according to the external recommendations, or by the restricted alternative. The 200ms is the shortest delay with the reported answer that therefore is closer to the cognitive conflict.

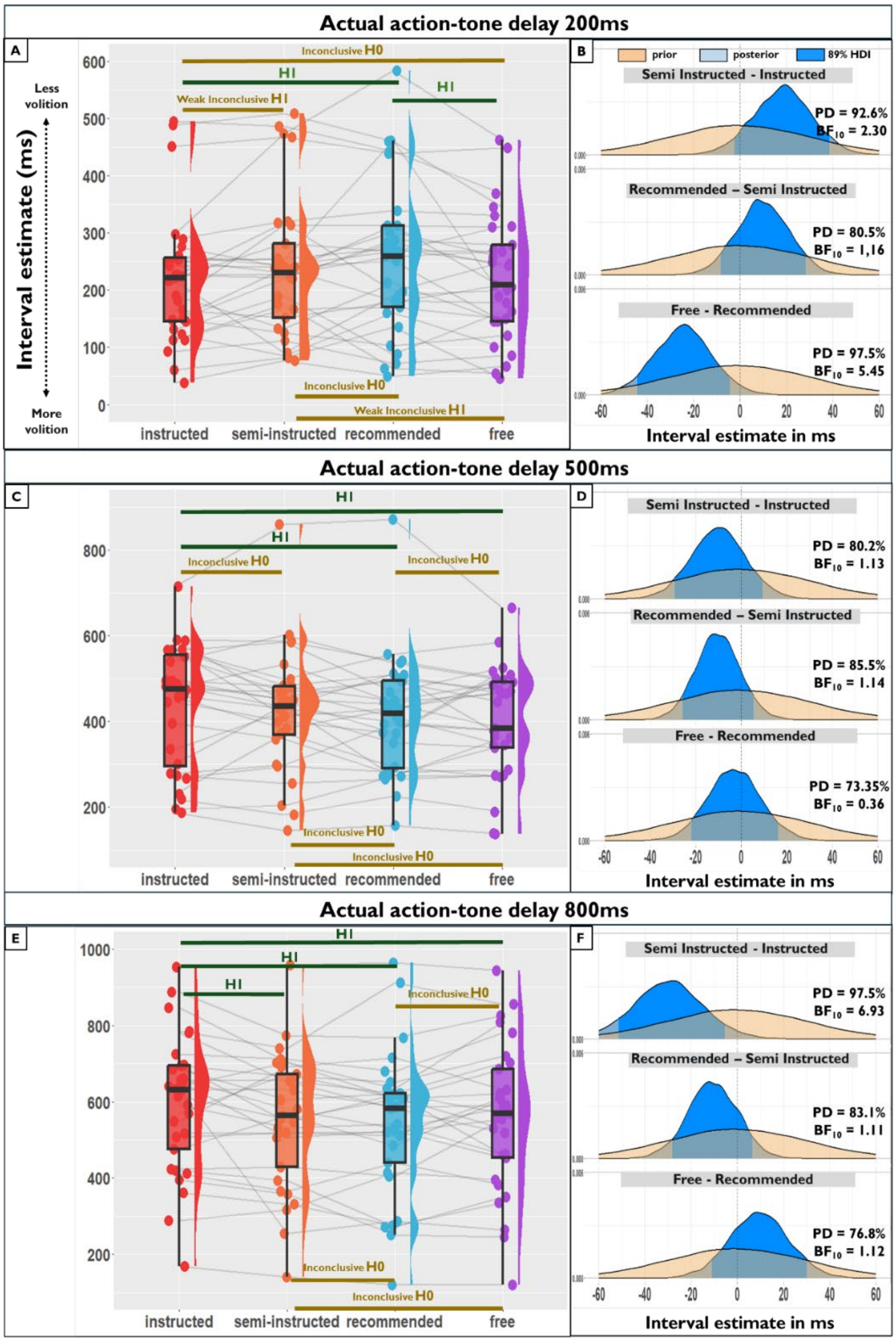

Figure 3. TB. (A-C-E) Similar as Fig. 2A across the three delays. H_0 represents evidence for similarity and H_1 for difference. (B-D-F) Posterior analysis of TB. Similar as Fig. 2B. Here, the SD of the prior distribution (in orange) is 15.

Readiness Potential

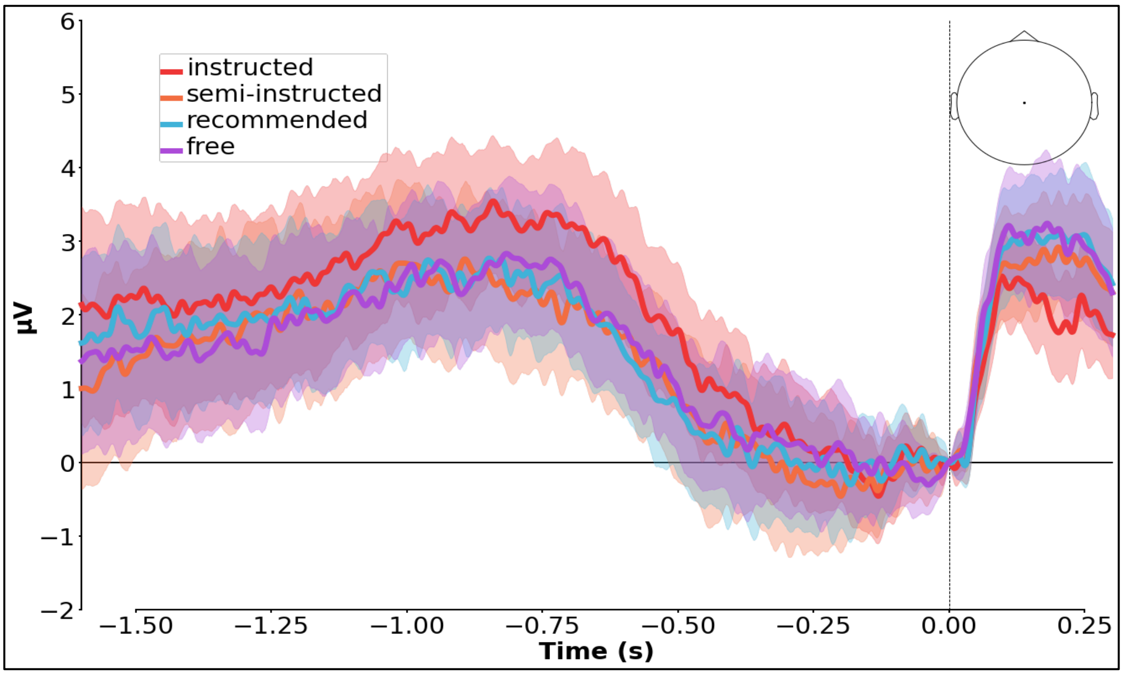

Figure 4. EEG plot of the RP. Each line represents a condition with its 95% confidence interval (i.e. the shaded area around each curve). The curves represent activity from the Cz electrode, time-locked on the keypress.

Mean amplitude RP

We analysed the mean of the RP (in µV) across the four conditions: Instructed, Semi-instructed, Recommended, and Free. As the baseline is cantered on –5ms to 5ms around the keypress, a higher value of the mean represents a stronger RP.

For the Instructed condition (Med = 1.590 µV, HDI89% = [0.862 2.310]), there was weak inconclusive evidence for a stronger RP compared to the Semi-Instructed condition (Med = 1.190 µV, HDI89% = [0.427 1.970]; Meddiff = -0.393 µV, HDI89% [–0.916 0.091], PD = 89.56%, BF10 = 1.38, RRIN = [0.232 2.164]), the Recommended condition (Med = 1.240 µV, HDI89% = [0.465 1.980]; Meddiff = –0.355 µV, HDI89% [–0.876 0.170], PD = 86.12%, BF10 = 1.18, RRIN = [0.188 1.882]), and the Free condition (Med = 1.160 µV, HDI89% = [0.386 1.940]; Meddiff = –0.436 µV, HDI89% [–0.938 0.052], PD = 91.74%, BF10 = 1.56, RRIN = [0.249 2.489]). For the Semi-Instructed condition (Med = 1.190 µV, HDI89% = [0.427 1.970]), there was inconclusive evidence for similar RP, with both the Recommended condition (Med

= 1.240 μV, HDI89% = [0.465 1.980]; Meddiff = 0.042 μV, HDI89% [–0.379 0.479], PD = 55.93%, BF10 = 0.55, RRIN = [0.087 0.874]), and the Free condition (Med = 1.160 μV, HDI89% = [0.386 1.940]; Meddiff = –0.041 μV, HDI89% [-0.504 0.422], PD = 55.36%, BF10 = 0.59, RRIN = [0.094 0.937]). We also observed inconclusive evidence for similar mean RP between the Recommended condition (Med = 1.240 μV, HDI89% = [0.465 1.980]) and the Free condition (Med = 1.160 μV, HDI89% = [0.386 1.940]; Meddiff = –0.078 μV, HDI89% [–0.500 0.336], PD = 62.33%, BF10 = 0.56, RRIN = [0.087 0.874]).

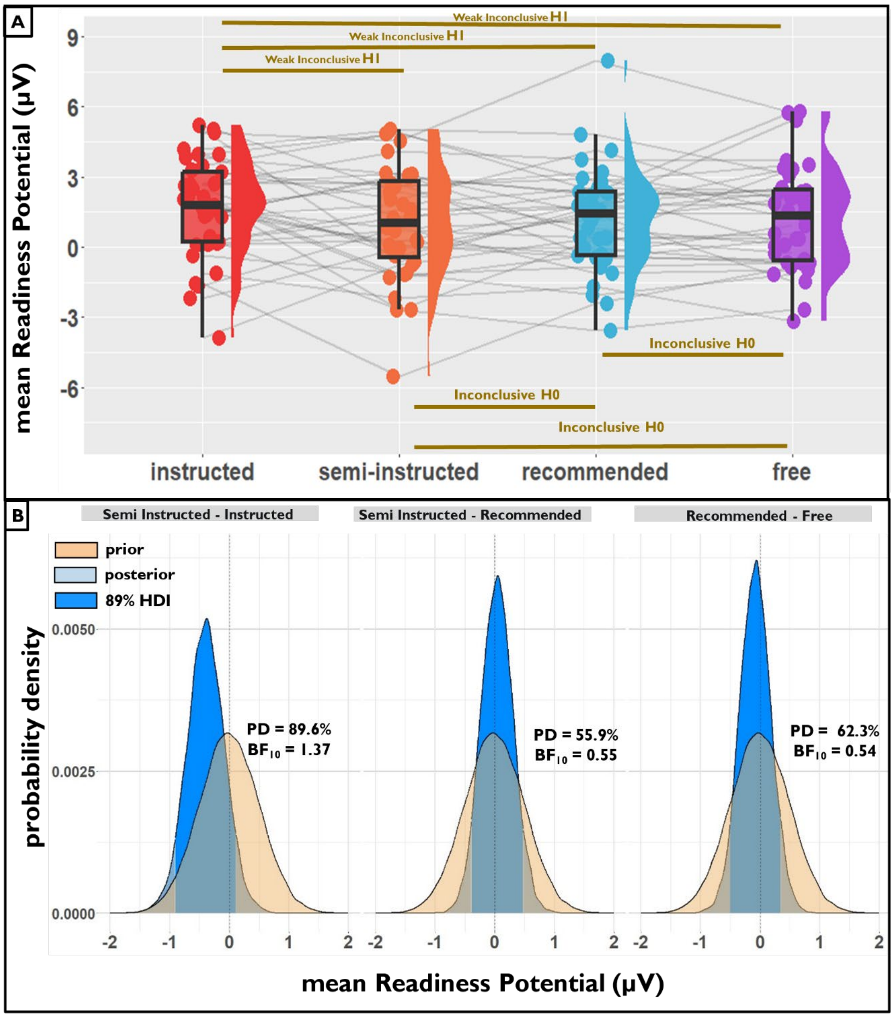

Figure 5 mean RP. (A) Similar as Fig. 2A. H_0 represents evidence for similarity and H_1 for difference. (B) Posterior analysis of the mean RP. Similar as Fig. 2B. Here, the SD of the prior distribution (in orange) is 1.

Slope RP

We analysed the slope of the RP (in µV) across the four conditions: Instructed, Semi-instructed, Recommended, and Free. The RP slope reflects the rate of change in neural activity preceding the keypress, with more negative values indicating steeper RP slopes (i.e., stronger RP). As the mean of the RP is highly influenced by the choice of the baseline, this gives us additional information that is not influenced by the baseline's choice.

For the Instructed condition (Med = –4.180 µV, HDI89% = [–5.420 –2.930]), there was weak inconclusive evidence for greater RP, compared to both the Semi-Instructed condition (Med = –3.740 µV, HDI89% = [–4.920 –2.510]; Meddiff = 0.451 µV, HDI89% [–0.333 1.203], PD = 82.07%, BF10 = 1.52, RRIN = [0.249 2.321]), and the Recommended condition (Med = –3.690 µV, HDI89% = [–4.890 –2.550]; Meddiff = 0.486 µV, HDI89% [–0.242 1.187], PD = 86.18%, BF10 = 1.69, RRIN = [0.267 2.668]). There was evidence for greater RP in the Instructed condition compared to the Free condition (Med = –3.430 µV, HDI89% = [-4.620 –2.160]; Meddiff = 0.751 µV, HDI89% [0.063 1.427], PD = 95.80%, BF10 = 3.61, RRH1 = [<0.5 0.616]). For the Semi-Instructed condition (Med = –3.740 µV, HDI89% = [–4.920 –2.510]), there was inconclusive evidence for similar RP slope compared to the Recommended condition (Med = –3.690 µV, HDI89% = [–4.890 –2.550]; Meddiff = 0.042 µV, HDI89% [–0.620 0.692], PD = 54.30%, BF10 = 0.81, RRIN = [0.133 1.239]), and to the Free condition (Med = –3.430 µV, HDI89% = [–4.620 –2.160]; Meddiff = 0.301 µV, HDI89% [–0.397 1.005], PD = 76.03%, BF10 = 1.13, RRIN = [0.176 1.756]). We also observed inconclusive evidence for similar RP slope between the Recommended condition (Med = –3.690 µV, HDI89% = [–4.890 –2.550]) and the Free condition (Med = –3.430 µV, HDI89% = [–4.620 –2.160]; Meddiff = 0.266 µV, HDI89% [–0.327 0.858], PD = 76.24%, BF10 = 0.97, RRIN = [0.164 1.527]).

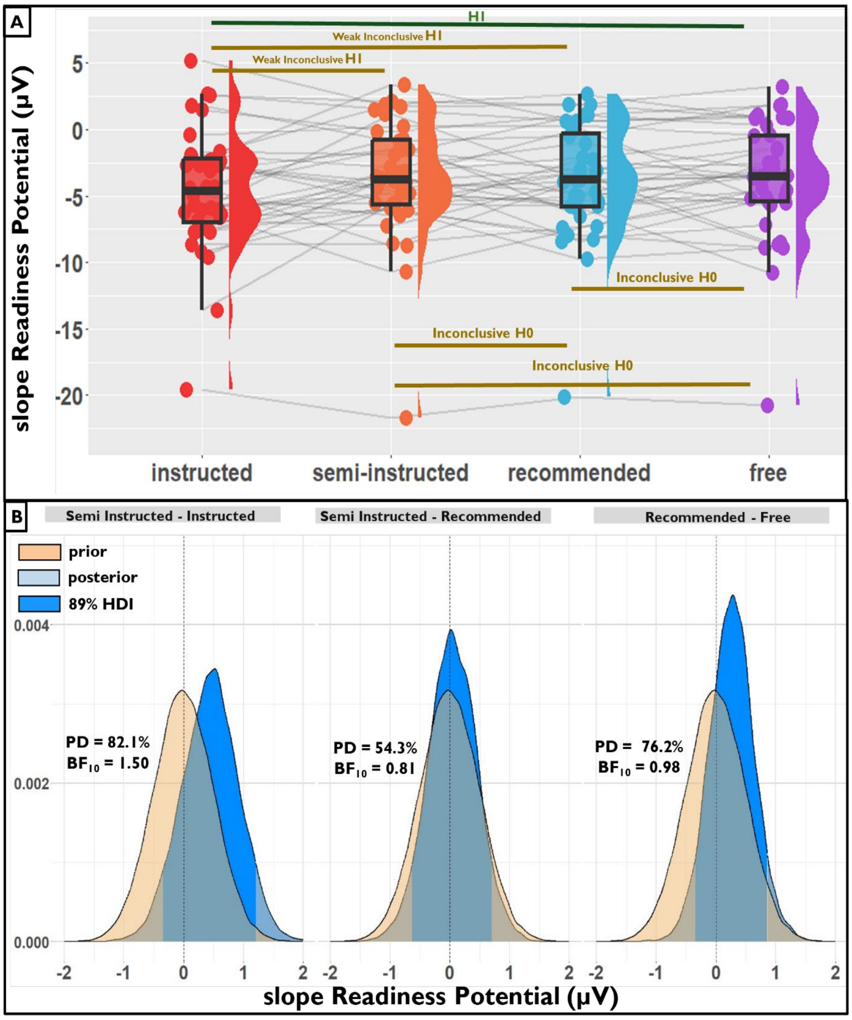

Figure 6 Figure 7 slope RP. (A) Similar as Fig. 2A. H_0 represents evidence for similarity and H_1 for difference. (B) Posterior analysis of the slope RP. Similar as Fig. 2B. Here, the SD of the prior distribution (in orange) is 1.

Resume RP

We found very little evidence for a greater RP in the Instructed condition compared to the others.

Discussion

Voluntary actions are frequently characterized as being self-generated. But what makes an action self-generated? Schuur and Haggard (2011) proposed that it depends on the integration of different types of input. Building on this account, we also tested the role of the nature of the input underlying the decision— whether the input is exogenous or endogenous. First, we examined how the quantity of different action alternatives (i.e., quantity or the input) affected markers of volition: a self-report measure (i.e., SoV), an implicit measure (i.e., TB), and a neural activity measure (i.e., RP). Second, we investigated how the quantity of endogenous versus exogenous input modulated the same markers. Thirty-nine participants completed a Random Dot Kinematic (RDK) task under four conditions. In the Instructed condition, they were required to indicate a given direction (most exogenous). In the Semi-Instructed condition, they chose one of four directions (more endogenous). In the Recommended condition, they chose among eight directions, with one recommended by the computer (mix of endogenous and exogenous). Finally, in the Free condition, they choose among eight directions without any recommendation (most endogenous).

The Recommended and Free conditions were important for testing the role of the nature of the input underlying the action. According to the initial proposal by Schuur and Haggard (2011), the Recommended condition—requiring the integration of more input—would be considered more self-generated. However, our alternative view suggests that the Recommended condition yields actions based on less endogenous input than the Free condition, resulting in actions considered less self-generated. Both views agree on the prediction that Instructed and Semi-Instructed conditions would produce actions considered less self-generated than both in the Recommended and Free conditions, due to the decrease in input. Following the same logic, these views also predicted that the Semi-Instructed condition would produce actions considered more self-generated than in the Instructed one.

Reported Sense of Volition

Our results showed that, subjectively, participants aligned more with our alternative view: decisions based on endogenous input were perceived as more self-generated. Interestingly, the RP and TB differed only when comparing the Instructed condition with the other conditions, indicating that these measures mainly reflect the contrast between exogenous and endogenous actions in a categorical rather than continuous manner. Incorporating the distinction between exogenous and endogenous input for decisions seems, therefore, to capture aspects at explicit, implicit, and neural levels.

Regarding the reported SoV, we hypothesized that the Instructed condition would yield the lowest score, followed by the Semi-Instructed condition, then the Recommended condition, and finally the Free condition (i.e., hypothesis 2). Our results supported this prediction: reliable differences emerged across these four conditions. According to Schuur and Haggard (2011), the quantity of input should render action more self-generated (i.e., hypothesis 1). Indeed, the Instructed condition (one input) produced lower reported SoV than the Semi-Instructed condition (four input), which in turn produced lower reported SoV than both the Recommended and Free conditions (nine and eight input, respectively). However, contrary to this multiple-input account, the Recommended condition elicited lower reported SoV than the Free condition, even though it involved more input. This pattern is readily explained by considering the notion of endogenous versus exogenous input: the recommendation introduces exogenous input, thereby reducing self-generation. Future studies should explore how the quantity of input and the balance between endogenous and exogenous input jointly shape the perception of self-generation, as their influence may not increase indefinitely, for instance. These findings also suggest that the account of self-generated action proposed by Passingham and colleagues (2011a, 2011b) and by Schuur and Haggard (2011, 2012) should be extended to include the endogenous–exogenous distinction of input to fully capture all the experience measured.

Temporal Binding

We hypothesized that integrating more input (i.e., hypothesis 1 & 2) and basing decisions on more endogenous input (i.e., hypothesis 2) would strengthen TB (i.e., shorten the estimated time interval). For the 200ms, we observed that the Recommended condition led to weaker TB than both the Free and Instructed conditions, with weak evidence for a similar effect in the Semi-Instructed condition. At the 500-ms delay, we found weak evidence for weaker TB in the Instructed condition compared to both the Free and Recommended conditions. Finally, we found weaker TB in the Instructed condition than in the other conditions at the 800-ms delay. The results for the 500ms and 800ms delays suggest that TB is sensitive to the contrast between decisions based on exogenous or endogenous input to act, rather than being in line with our two initial hypotheses. This pattern replicates previous findings contrasting Free (endogenous) and Instructed decisions. Notably, Barlas and colleagues (2013) showed that TB differences emerge only when contrasting decisions with the largest number of alternatives (i.e., 7 vs 1). Based on visual inspection of their plots, it seems that TB increased as alternatives increased (Barlas, 2013: pages 5). This was also the case in their follow-up study, they found a significant TB differences only when contrasting decision with the largest number of alternatives (i.e., 4 vs 1), but again, descriptively the TB seems to have increased as alternatives increased (Barlas et al., 2017, page 126). However, we cannot rule out that this is an artifact emerging from noise. Nevertheless, it may help explain the mixed results in our Semi-Instructed condition, where participants sometimes followed their chosen direction but at other times pressed a direction more randomly, thereby reducing the clarity of endogenous determination.

We might also interpret the results for the actual delay of 200ms considering studies showing that greater cognitive conflict weakened the TB (Barlas et al., 2017; Galang et al., 2025; Howard et al., 2016; but a study did not find such difference: Bussche et al., 2020). Semi-Instructed and Recommended trials are likely to produces a greater cognitive conflict than either Instructed (decision pre-specified) or Free trials (decision unconstrained). The effect emerged only at the shorter delay, possibly because cognitive conflict exerts its strongest influence near the time of action, whereas longer delays allow its effect to fade. Supporting this view, Barlas et al. (2017) found that subjective control increased with the number of alternatives, aligning with the "feeling of control" dimension of the SoV (Pacherie, 2008). Moreover, Galang and colleagues (2025) observed that following a rule led to stronger TB than breaking a rule, but this was the case only when looking at the shortest interval. Future work should test longer delays (e.g., 900, 1,100, or 1,300 ms), or temporally separate the choice from the execution phase, to test the robustness of the effect contrasting instructed and free decisions. Interestingly, in the literature, TB appears to vary depending on the delay: in some studies, it strengthens with longer delays (Caspar, Christensen, et al., 2016; Engbert et al., 2008; Kong et al., 2024; Ruess et al., 2017; Suzuki et al., 2019), while in others, it strengthens with shorter delays (Caspar, Desantis, et al., 2016; Galang et al., 2025; Haggard et al., 2002; Ruess et al., 2017). There is currently no clear explanation in the literature for why TB differs across delay manipulations. Future studies should investigate the underlying mechanisms responsible for these differences.

Readiness Potential

Regarding the RP, we expected larger RPs when more input was integrated or when decisions were more endogenous but found only weak evidence for a stronger RP in the Instructed condition compared to the other conditions. Thus, neither hypothesis was confirmed; in fact, the pattern was the opposite of the predicted one. Prior studies contrasting self-generated with external-stimulus-triggered actions usually report a higher RP in the former (Di Russo et al., 2017; Jahanshahi et al., 1995; Jenkins et al., 2000; Khalighinejad et al., 2018). These studies contrasted actions where the decision "when" to act was endogenously generated versus actions where it was exogenously generated. Critically, if the RP represents preparation to act, shortening the time of the preparation might impact it, regardless of the differences between exogenously and endogenously generated actions. Following this, stimulus-triggered reactions are confounded, as the preparation to act should be done as fast as possible in reaction to the cue, while in a self-generated condition, the preparation could arise slower. This could explain why we do not find the same effect as these earlier studies. We contrasted actions where the decision on “what” to do was either externally chosen (Instructed condition) or self-generated (Semi-Instructed, Recommended, Free conditions), but importantly, the decision as to “when” to act was always self-generated. One interpretation is that the size of the RP primarily reflects action preparation rather than decision preparation: when the decision is externally specified (Instructed condition), preparation is mainly about the motor plan, yielding a larger RP. Conversely, when the decision remains unresolved until late (Semi-Instructed, Recommended, Free), motor preparation fluctuates with the evolving choice (i.e., motor plan), reducing the RP. Whether this is more in line with the intentional-planning account or threshold-based accumulation accounts (e.g., Schurger, 2012) remains to be tested.

Limitations and Conclusion

Our findings might generalize only to the present type of recommendation (e.g., “press this key”). Different recommendations, such as avoiding an incorrect answer or making value-based choices, might yield different patterns. Future studies should also probe participants’ strategies explicitly and manipulate the proportion of correct recommendations to match accuracy across conditions. Moreover, because we did not save the exact arrow locations, we could not analyse spatial error or adapt difficulty to individual performance. In addition to this, Schuur and Haggard (2011) acknowledged that the notion of “input” was underspecified. Here, we treated as input the perceptual evidence (dot motion), the set of action alternatives, and the recommendation. Some may argue that the number of alternatives is not a distinct “type” of input; experiments that vary qualitatively different information, such as moral attributes, would clarify this issue. Finally, our conception of endogenous versus exogenous input might be further clarified in the future, specifying where to locate interoception, for instance.

In a multi-agent world, recognizing the source of generation matters. In this study, we thus vary whether the input is endogenous or exogenous, as well as the quantity of input received, to investigate how they influence three markers of volition. Overall, our results show that a simple recommendation can change how voluntary an action feels, yet the key implicit and neural markers of volition mostly track who makes the choice, the person or an outside source. An operant keypress still looks "voluntary," but when its reason is imposed, that reduces both the reported feeling that it is self-generated, as well as the Readiness-Potential and Temporal-Binding. What matters, then, is *who decides*, not just *who moves*. By showing that the self-generation of the decision, not merely the self-initiation of movement, is critical, we refine neuroscientific accounts of volition, which might perhaps in the future also serve in working towards giving a clearer basis for judging moral responsibility. Follow-up studies should investigate whether strengthening endogenous processing, for example, through metacognitive training, can restore both the experience and the implicit markers of volition when external pressures are unavoidable.

Acknowledgements

A.C. is a Research Director with the F.R.S.-FNRS (Belgium) and a fellow of the Canadian Institute for Advanced Research (Brain,Mind,and Consciousness program).

Conflict of inte est

None declared.

FUNDING INFORMATION

G.P.P. has been funded by the Fonds National de la Recherche Scientifique (F.R.S.-FNRS).

AUTHORS CONTRIBUTIONS

Guillaume P. Pech (Conceptualization [lead], Data curation [lead], Formal analysis [lead], Methodology [lead], Visualization [lead], Writing — original draft [lead]), Eva Nicolay (Conceptualization [equal], Methodology [equal], Investigation [lead], Writing — review & editing [equal]), Paulius Rimkevičius (Methodology [equal], Writing — review & editing [equal]), Uri Maoz (Writing — review & editing [equal]), Axel Cleeremans (Project administration [lead], Supervision[lead],Writing — review & editing [equal]).

DATA AVAILABILITY STATEMENT

The behavioural data, analysis scripts, and Bayesian models supporting the findings of this study are available in an OSF repository at the following link: https://osf.io/myprojects/

References

Barlas, Z. (2019). When robots tell you what to do: Sense of agency in human- and robot-guided actions. *Consciousness and Cognition*, *75*, 102819. https://doi.org/10.1016/j.concog.2019.102819

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. https://doi.org/10.1016/j.actpsy.2017.09.004

Barlas, Z., & Obhi, S. (2013). Freedom, choice, and the sense of agency. *Frontiers in Human Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00514

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*. https://www.frontiersin.org/articles/10.3389/fninf.2015.00016

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, Doing, and Controlling: Implicit Sense of Agency Over Somatosensory Events. *Psychological Science*, *28*(7), 882–893. https://doi.org/10.1177/0956797617697693

Brunia, C. H. M., Boxtel, G. J. M. van, & Böcker, K. B. E. (2011, December 15). *Negative Slow Waves as Indices of Anticipation: The Bereitschaftspotential, the Contingent Negative Variation, and the Stimulus-Preceding Negativity*. The Oxford Handbook of Event-Related Potential Components. https://doi.org/10.1093/oxfordhb/9780195374148.013.0108

Brysbaert, M. (2019). How Many Participants Do We Have to Include in Properly Powered Experiments? A Tutorial of Power Analysis with Reference Tables. *Journal of Cognition*, *2*(1), 16. https://doi.org/10.5334/joc.72

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), Article 1. https://doi.org/10.5334/joc.10

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Bussche, E. V. den, Alves, M., Murray, Y. P. J., & Hughes, G. (2020). The effect of cognitive effort on the sense of agency. *PLOS ONE*, *15*(8), e0236809. https://doi.org/10.1371/journal.pone.0236809

Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion Changes the Sense of Agency in the Human Brain. *Current Biology*, *26*(5), 585–592. https://doi.org/10.1016/j.cub.2015.12.067

Caspar, E. A., Cleeremans, A., & Haggard, P. (2018). Only giving orders? An experimental study of the sense of agency when giving or receiving commands. *PLOS ONE*, *13*(9), e0204027. https://doi.org/10.1371/journal.pone.0204027

Caspar, E. A., Desantis, A., Dienes, Z., Cleeremans, A., & Haggard, P. (2016). The Sense of Agency as Tracking Control. *PLOS ONE*, *11*(10), e0163892. https://doi.org/10.1371/journal.pone.0163892

Caspar, E. A., Lo Bue, S., Magalhães De Saldanha da Gama, P. A., Haggard, P., & Cleeremans, A. (2020). The effect of military training on the sense of agency and outcome processing. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18152-x

Charles, L., & Haggard, P. (2020). Feeling free: External influences on endogenous behaviour. *Quarterly Journal of Experimental Psychology*, *73*(4), 568–577. https://doi.org/10.1177/1747021819888066

Choi, B. C. K., & Pak, A. W. P. (2005). A catalog of biases in questionnaires. *Preventing Chronic Disease*, *2*(1), A13.

Cohen, M. X., & Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *Journal of Neurophysiology*, *110*(12), 2752–2763. https://doi.org/10.1152/jn.00479.2013

Di Russo, F., Berchicci, M., Bozzacchi, C., Perri, R. L., Pitzalis, S., & Spinelli, D. (2017). Beyond the "Bereitschaftspotential": Action preparation behind cognitive functions. *Neuroscience & Biobehavioral Reviews*, *78*, 57–81. https://doi.org/10.1016/j.neubiorev.2017.04.019

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

Dienes, Z. (2019). How Do I Know What My Theory Predicts? *Advances in Methods and Practices in Psychological Science*, *2*(4), 364–377. https://doi.org/10.1177/2515245919876960

Engbert, K., Wohlschläger, A., & Haggard, P. (2008). Who is causing what? The sense of agency is relational and efferent-triggered. *Cognition*, *107*(2), 693–704. https://doi.org/10.1016/j.cognition.2007.07.021

Fanselow, M. S., & Wassum, K. M. (2016). The Origins and Organization of Vertebrate Pavlovian Conditioning. *Cold Spring Harbor Perspectives in Biology*, *8*(1), a021717. https://doi.org/10.1101/cshperspect.a021717

Filevich, E., Vanneste, P., Brass, M., Fias, W., Haggard, P., & Kühn, S. (2013). Brain correlates of subjective freedom of choice. *Consciousness and Cognition*, *22*(4), 1271–1284. https://doi.org/10.1016/j.concog.2013.08.011

Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, *20*(2), 303–315. https://doi.org/10.1086/209351

Galang, C. M., Akan, A., Pfister, R., & Brass, M. (2025). Temporal binding during deliberate rule breaking. *Consciousness and Cognition*, *130*, 103851. https://doi.org/10.1016/j.concog.2025.103851

Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford University Press.

Gavenas, J., Schurger, A., & Maoz, U. (2025). Probing for intentions: The early readiness potential does not reflect awareness of motor preparation. *Imaging Neuroscience*, *3*, imag_a_00465. https://doi.org/10.1162/imag_a_00465

Goldman, A., & de Vignemont, F. (2009). Is social cognition embodied? *Trends in Cognitive Sciences*, *13*(4), 154–159. https://doi.org/10.1016/j.tics.2009.01.007

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00267

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207. https://doi.org/10.1038/nrn.2017.14

Haggard, P. (2019). The Neurocognitive Bases of Human Volition. *Annual Review of Psychology*, *70*, 9–28. https://doi.org/10.1146/annurev-psych-010418-103348

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

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association*, *81*(396), 991–999. https://doi.org/10.2307/2289073

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), e12843. https://doi.org/10.1111/cogs.12843

Howard, E. E., Edwards, S. G., & Bayliss, A. P. (2016). Physical and mental effort disrupts the implicit sense of agency. *Cognition*, *157*, 114–125. https://doi.org/10.1016/j.cognition.2016.08.018

Jahanshahi, M., Jenkins, I. H., Brown, R. G., Marsden, C. D., Passingham, R. E., & Brooks, D. J. (1995). Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. *Brain: A Journal of Neurology*, *118 (Pt 4)*, 913–933. https://doi.org/10.1093/brain/118.4.913

Jeannerod, M. (2009). *Le Cerveau volontaire*. Odile Jacob.

Jenkins, I. H., Jahanshahi, M., Jueptner, M., Passingham, R. E., & Brooks, D. J. (2000). Self-initiated versus externally triggered movements: II. The effect of movement predictability on regional cerebral blood flow. *Brain*, *123*(6), 1216–1228. https://doi.org/10.1093/brain/123.6.1216

Jones, P. R. (2019). A note on detecting statistical outliers in psychophysical data. *Attention, Perception, & Psychophysics*, *81*(5), 1189–1196. https://doi.org/10.3758/s13414-019-01726-3

Kaiser, J., Buciuman, M., Gigl, S., Gentsch, A., & Schütz-Bosbach, S. (2021). The Interplay Between Affective Processing and Sense of Agency During Action Regulation: A Review. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.716220

Khalighinejad, N., Brann, E., Dorgham, A., & Haggard, P. (2019). Dissociating Cognitive and Motoric Precursors of Human Self-Initiated Action. *Journal of Cognitive Neuroscience*, *31*(5), 754–767. https://doi.org/10.1162/jocn_a_01380

Khalighinejad, N., Schurger, A., Desantis, A., Zmigrod, L., & Haggard, P. (2018). Precursor processes of human self-initiated action. *NeuroImage*, *165*, 35–47. https://doi.org/10.1016/j.neuroimage.2017.09.057

Kong, G., Aberkane, C., Desoche, C., Farnè, A., & Vernet, M. (2024). No evidence in favor of the existence of “intentional” binding. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(6), 626–635. https://doi.org/10.1037/xhp0001204

Kruschke, J. K. (2018). Rejecting or Accepting Parameter Values in Bayesian Estimation. *Advances in Methods and Practices in Psychological Science*, *1*(2), 270–280. https://doi.org/10.1177/2515245918771304

Lau, S., Hiemisch, A., & Baumeister, R. F. (2015). The experience of freedom in decisions – Questioning philosophical beliefs in favor of psychological determinants.

Consciousness and Cognition, *33*, 30–46. https://doi.org/10.1016/j.concog.2014.11.008

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of Effect Existence and Significance in the Bayesian Framework. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02767

Maoz, U., Yaffe, G., Koch, C., & Mudrik, L. (2019). Neural precursors of decisions that matter-an ERP study of deliberate and arbitrary choice. *eLife*, *8*, e39787. https://doi.org/10.7554/eLife.39787

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC. https://doi.org/10.1201/9781315372495

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*(1), 546–561. https://doi.org/10.1016/j.concog.2011.12.002

Nachev, P., & Husain, M. (2010). Action and the fallacy of the 'internal': Comment on Passingham et al. *Trends in Cognitive Sciences*, *14*(5), 192–193. https://doi.org/10.1016/j.tics.2010.03.002

Nann, M., Cohen, L. G., Deecke, L., & Soekadar, S. R. (2019). To jump or not to jump—The Bereitschaftspotential required to jump into 192-meter abyss. *Scientific Reports*, *9*(1), 2243. https://doi.org/10.1038/s41598-018-38447-w

Obhi, S. S. (2012). The troublesome distinction between self-generated and externally triggered action: A commentary on Schüür and Haggard. *Consciousness and Cognition*, *21*(1), 587–588. https://doi.org/10.1016/j.concog.2011.09.014

Osborne, J., & Overbay, A. (2019). The power of outliers (and why researchers should ALWAYS check for them). *Practical Assessment, Research, and Evaluation*, *9*(1). https://doi.org/10.7275/qf69-7k43

Pan, W., Lei, Y., Lin, J., & Li, H. (2024). Action selection and simultaneously presented emotional sound and reward: Differential effects on implicit and explicit sense of agency. *Consciousness and Cognition*, *126*, 103784. https://doi.org/10.1016/j.concog.2024.103784

Parés-Pujolràs, E., Kim, Y.-W., Im, C.-H., & Haggard, P. (2019). Latent awareness: Early conscious access to motor preparation processes is linked to the readiness potential. *NeuroImage*, *202*, 116140. https://doi.org/10.1016/j.neuroimage.2019.116140

Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2010a). Is it fallacious to talk of self-generated action?: Response to Nachev and Husain. *Trends in Cognitive Sciences*, *14*(5), 193–194. https://doi.org/10.1016/j.tics.2010.03.004

Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2010b). Medial frontal cortex: From self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences*, *14*(1), 16–21. https://doi.org/10.1016/j.tics.2009.11.001

Pavlov (1927), P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, *17*(3), 136–141. https://doi.org/10.5214/ans.0972-7531.1017309

Pech, G. P., & Caspar, E. A. (2022). Can a Video Game with a Fictional Minority Group Decrease Intergroup Biases towards Non-Fictional Minorities? A Social Neuroscience

Study. *International Journal of Human–Computer Interaction*, *0*(0), 1–15. https://doi.org/10.1080/10447318.2022.2121052

Pech, G. P., & Caspar, E. A. (2023). Does the cowl make the monk? The effect of military and Red Cross uniforms on empathy for pain, sense of agency and moral behaviors. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1255835

Pech, G. P., & Caspar, E. A. (2025). A cross-cultural EEG study of how obedience and conformity influence reconciliation intentions. *Social Cognitive and Affective Neuroscience*, *20*(1), nsaf038. https://doi.org/10.1093/scan/nsaf038

Pech, G. P., Caspar, E. A., Pacherie, E., Cleeremans, A., & Maoz, U. (2025). A multi-measurement study of the relation between deliberation and volition. *Neuroscience of Consciousness*, *2025*(1), niaf023. https://doi.org/10.1093/nc/niaf023

Pech, G. P., Gishoma, D., & Caspar, E. A. (2024). A novel electroencephalography-based paradigm to measure intergroup prosociality: An intergenerational study in the aftermath of the genocide against Tutsis in Rwanda. *Journal of Experimental Psychology: General*, *153*(1), 241–254. https://doi.org/10.1037/xge0001480

Rizopoulos, D. (2007). ltm: An R Package for Latent Variable Modeling and Item Response Analysis. *Journal of Statistical Software*, *17*, 1–25. https://doi.org/10.18637/jss.v017.i05

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

Schmidt, S., Jo, H.-G., Wittmann, M., & Hinterberger, T. (2016). ‘Catching the waves’ – slow cortical potentials as moderator of voluntary action. *Neuroscience & Biobehavioral Reviews*, *68*, 639–650. https://doi.org/10.1016/j.neubiorev.2016.06.023

Schurger, A., Hu, P. ‘Ben’, Pak, J., & Roskies, A. L. (2021). What Is the Readiness Potential? *Trends in Cognitive Sciences*, *25*(7), 558–570. https://doi.org/10.1016/j.tics.2021.04.001

Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences*, *109*(42), E2904–E2913. https://doi.org/10.1073/pnas.1210467109

Schüür, F., & Haggard, P. (2011). What are self-generated actions? *Consciousness and Cognition*, *20*(4), 1697–1704. https://doi.org/10.1016/j.concog.2011.09.006

Schüür, F., & Haggard, P. (2012). On capturing the essence of self-generated action: A reply to Obhi (2012). *Consciousness and Cognition*, *21*(2), 1070–1071. https://doi.org/10.1016/j.concog.2012.03.003

Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The Brain in (Willed) Action: A Meta-Analytical Comparison of Imaging Studies on Motor Intentionality and Sense of Agency. *Frontiers in Psychology*, *10*, 804. https://doi.org/10.3389/fpsyg.2019.00804

Semmelmann, K., & Weigelt, S. (2017). Online psychophysics: Reaction time effects in cognitive experiments. *Behavior Research Methods*, *49*(4), 1241–1260. https://doi.org/10.3758/s13428-016-0783-4

Shapiro, L. A. (2019). Flesh matters: The body in cognition. *Mind & Language*, *34*(1), 3–20. https://doi.org/10.1111/mila.12203

Shi, H., & and Yin, G. (2021). Reconnecting p-Value and Posterior Probability Under One- and Two-Sided Tests. *The American Statistician*, *75*(3), 265–275. https://doi.org/10.1080/00031305.2020.1717621

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *117*(11), 2341–2356. https://doi.org/10.1016/j.clinph.2006.04.025

Silvetti, M., & Verguts, T. (2012). Reinforcement Learning, High-Level Cognition, and the Human Brain. In P. Bright (Ed.), *Neuroimaging—Cognitive and Clinical Neuroscience*. InTech. https://doi.org/10.5772/23471

Skinner, B. F. (1984). Selection by consequences. *Behavioral and Brain Sciences*, *7*(4), 477–481. https://doi.org/10.1017/S0140525X0002673X

Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional Binding Without Intentional Action. *Psychological Science*, *30*(6), 842–853. https://doi.org/10.1177/0956797619842191

Tapal, A., Oren, E., Dar, R., & Eitam, B. (2017). The Sense of Agency Scale: A Measure of Consciously Perceived Control over One's Mind, Body, and the Immediate Environment. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01552

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thaler, D., Chen, Y.-C., Nixon, P. D., Stern, C. E., & Passingham, R. E. (1995). The functions of the medial premotor cortex. *Experimental Brain Research*, *102*(3), 445–460. https://doi.org/10.1007/BF00230649

Thompson, E., & Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences*, *5*(10), 418–425. https://doi.org/10.1016/S1364-6613(00)01750-2

Travers, E., Friedemann, M., & Haggard, P. (2021). The Readiness Potential reflects planning-based expectation, not uncertainty, in the timing of action. *Cognitive Neuroscience*, *12*(1), 14–27. https://doi.org/10.1080/17588928.2020.1824176

Travers, E., & Haggard, P. (2021). The Readiness Potential reflects the internal source of action, rather than decision uncertainty. *The European Journal of Neuroscience*, *53*(5), 1533–1544. https://doi.org/10.1111/ejn.15063

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(1), 80–91. https://doi.org/10.1037/0096-1523.31.1.80

Verbaarschot, C., Farquhar, J., & Haselager, P. (2019). Free Wally: Where motor intentions meet reason and consequence. *Neuropsychologia*, *133*, 107156. https://doi.org/10.1016/j.neuropsychologia.2019.107156

Wen, W., & Imamizu, H. (2022). The sense of agency in perception, behaviour and human–machine interactions. *Nature Reviews Psychology*, *1*(4), Article 4. https://doi.org/10.1038/s44159-022-00030-6

Wiesing, M., & Zimmermann, E. (2024). Intentional binding – Is it just causal binding? A replication study of Suzuki et al. (2019). *Consciousness and Cognition*, *119*, 103665. https://doi.org/10.1016/j.concog.2024.103665

Winkler, I., Debener, S., Müller, K.-R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4101–4105. https://doi.org/10.1109/EMBC.2015.7319296

Yao, D. (2001). A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiological Measurement*, *22*(4), 693. https://doi.org/10.1088/0967-3334/22/4/305

Yao, D., Qin, Y., Hu, S., Dong, L., Bringas Vega, M. L., & Valdés Sosa, P. A. (2019). Which Reference Should We Use for EEG and ERP practice? *Brain Topography*, *32*(4), 530–549. https://doi.org/10.1007/s10548-019-00707-x

Yoshie, M., & Haggard, P. (2013). Negative emotional outcomes attenuate sense of agency over voluntary actions. *Current Biology: CB*, *23*(20), 2028–2032. https://doi.org/10.1016/j.cub.2013.08.034